

Article

# Multivariate Weibull Distribution for Wind Speed and Wind Power Behavior Assessment

Daniel Villanueva \*, Andrés Feijóo and José L. Pazos

Departamento de Enxeñaría Eléctrica, Universidade de Vigo, Maxwell s/n, Vigo 36301, Spain;  
E-Mails: afeijoo@uvigo.es (A.F.); pazos@uvigo.es (J.L.P.)

\* Author to whom correspondence should be addressed; E-Mail: dvillanueva@uvigo.es;  
Tel.: +34-986-818-745; Fax: +34-986-814-014.

Received: 19 July 2013; in revised form: 5 August 2013 / Accepted: 23 August 2013 /

Published: 3 September 2013

---

**Abstract:** The goal of this paper is to show how to derive the multivariate Weibull probability density function from the multivariate Standard Normal one and to show its applications. Having Weibull distribution parameters and a correlation matrix as input data, the proposal is to obtain a precise multivariate Weibull distribution that can be applied in the analysis and simulation of wind speeds and wind powers at different locations. The main advantage of the distribution obtained, over those generally used, is that it is defined by the classical parameters of the univariate Weibull distributions and the correlation coefficients and all of them can be easily estimated. As a special case, attention has been paid to the bivariate Weibull distribution, where the hypothesis test of the correlation coefficient is defined.

**Keywords:** wind speed; wind power; bivariate Weibull distribution; multivariate Weibull distribution; correlation; inference

---

## 1. Introduction

The Weibull distribution is a continuous probability distribution that was described by Waloddi Weibull in 1951 [1]. It can be applied in a wide range of fields such as survival analysis, reliability engineering or weather forecasting, to mention just a few [1–5]. Mathematically, it can be written as a function of three parameters, *i.e.*, location parameter  $\gamma$ , scale parameter  $C$  and shape parameter  $k$ . The

location parameter gives information about the minimum value of the set, so it can be substituted by 0 in the case of wind speed sets.

When more than one Weibull-described variable is being considered and the dependence among them has certain relevance, the so called multivariate distribution [2,3] has to be applied. The multivariate distribution describes the outcome of a number of variables at the same time. Therefore, if the simultaneous behavior of a number of dependent variables, each of them described by a Weibull distribution, is being evaluated, a multivariate Weibull distribution needs to be used. Obviously, when the variables are independent, they can be evaluated separately.

So far, most of the multivariate Weibull expressions for Cumulative Distribution Function (CDF) or Probability Distribution Function (PDF) are based on models [4,5] that approximately describe the joint probability distribution of a group of variables. The most noteworthy one was introduced by Marshall and Olkin in 1967 [6], but others are also interesting such as those presented by Lee [7], Roy and Mukherjee [8], Crowder [9], Lu and Bhattacharyya [10] and Patra and Dey [11]. Bivariate Weibull models have also been developed based on the following Copula functions [4]: Farlie-Gumbel-Morgenstern, Clayton, Ali-Mikhail-Haq, Gumbel-Hougaard, Gumbel-Barnett, Nelsen Ten. However, most of these models are based on the estimation of parameters with no significance in order to define the model completely, *i.e.*, in most cases those parameters have no direct relationship to the univariate Weibull model parameters and with the correlation coefficients which are used as a measurement of the dependence.

In this paper, a model is proposed for the multivariate Weibull PDF, based on the classic parameters used in the definition of a univariate Weibull model and on the correlation coefficients among the marginal distributions. It develops the change of variables from Normal to Weibull used in [12,13]. It is applied to wind speed and wind power, as there seems to be some agreement on the importance of Weibull distribution for the modeling and simulation of these in wind turbines or wind farms [14–16]. Therefore, this paper presents the relationship of the simultaneous behavior between wind speed and power in a pair or more locations. The bivariate case is analyzed on its own due to its importance, and the correlation coefficient inference procedure is established.

The structure of the paper is as follows: Section 2 derives the multivariate Weibull PDF from the Standard Normal one and outlines its application to wind speed and wind power, Section 3 deals with the bivariate Weibull PDF and Section 4 states the conclusions.

## 2. Multivariate Weibull Distribution

The proposed multivariate Weibull distribution is obtained by means of a Normal to Weibull change of variables. The procedure to derive it is as follows:

- (1) The key point of the procedure is to define a change of variable from a Standard Normal to a Weibull distributed one. This transformation must be differentiable and the inverse function has to exist. The Standard Normal distributed variable is created in order to use its known features and transfer them to the Weibull one;
- (2) Establish as many changes as the number of variables,  $n$ , considering the different Weibull parameters for each case;
- (3) Obtain the multivariate Weibull PDF from the multivariate Standard Normal PDF applying the change of variable.

Finally, the relationship between the correlation coefficients in the multivariate Standard Normal PDF and in the Weibull one has to be checked in order to establish a multivariate Weibull PDF that depends on marginal Weibull parameters and the correlation coefficients between pairs of Weibull variables.

### 2.1. Normal to Weibull Change of Variables

In order to define a change of variables from a Normal distributed variable to a Weibull one, the Probability Integral Transform is applied [17], based on the obtaining of Uniform distributed variables. The Uniform distributed variables derived from Normal and Weibull ones are then equalized and a relationship between them is established.

The CDF of a univariate Weibull distribution [2–4] with scale parameter  $C$  and shape parameter  $k$  is defined in Equation (1):

$$F_u(u) = 1 - \exp(-(u/C)^k) \quad (1)$$

where  $\exp(\cdot)$  is the exponential function;  $u$  is the Weibull variable; and  $F_u(u)$  is the CDF of  $u$ .

On the other hand, the CDF of the univariate Normal distribution [2–4] with mean value  $\mu$  and standard deviation  $\sigma$  is defined in Equation (2):

$$F_x(x) = \left(1 + \operatorname{erf}\left(\frac{x - \mu}{\sqrt{2}\sigma}\right)\right)/2 \quad (2)$$

where  $\operatorname{erf}(\cdot)$  is the error function, defined in [18];  $x$  is the Normal variable; and  $F_x(x)$  is the CDF of  $x$ .

As mentioned earlier, the Probability Integral Transform states that variables from any given continuous distribution can be converted into variables having a uniform distribution, *i.e.*, applying its CDF to a variable provides a new variable, uniformly distributed. So,  $y_u = F_u(u)$  and  $y_x = F_x(x)$  are uniform distributed variables.

Both variables  $y_u$  and  $y_x$  can be matched in order to establish a relationship between a Weibull and a Normal distributed variable. The whole process can be understood in two steps: First, the conversion of a Normal distributed variable into a Uniform one and then the conversion of this one into a Weibull one. Therefore, if both CDFs are matched, then the Weibull distributed variable,  $u$ , can be expressed as a function of the Normal one  $x$ , such as in Equation (3):

$$u = C \left( -\log \left( \left( 1 - \operatorname{erf} \left( \frac{x - \mu}{\sqrt{2}\sigma} \right) \right) / 2 \right) \right)^{1/k} \quad (3)$$

where  $\log(\cdot)$  is the natural logarithm.

Thus, from a Normal distributed variable with given parameters  $(\mu, \sigma)$ , a Weibull one can be derived, with the desired parameters  $(C, k)$ . Notice that this transformation can be applied to other types of variables to obtain an equation that relates two variables with different distributions.

In order to derive further results, the transformation given in Equation (3) will be referred to as  $\text{ntw}(x; C, k)$  where  $x$  is a Standard Normal variable ( $\mu = 0$  and  $\sigma = 1$ ), the Weibull parameters are  $C$  and  $k$ . For a single variable,  $x_i$ , the notation is expressed in Equation (4):

$$\text{ntw}(x_i; C_i, k_i) = C_i \left( -\log \left( \left( 1 - \operatorname{erf}(x_i/\sqrt{2}) \right) / 2 \right) \right)^{1/k_i} \quad (4)$$

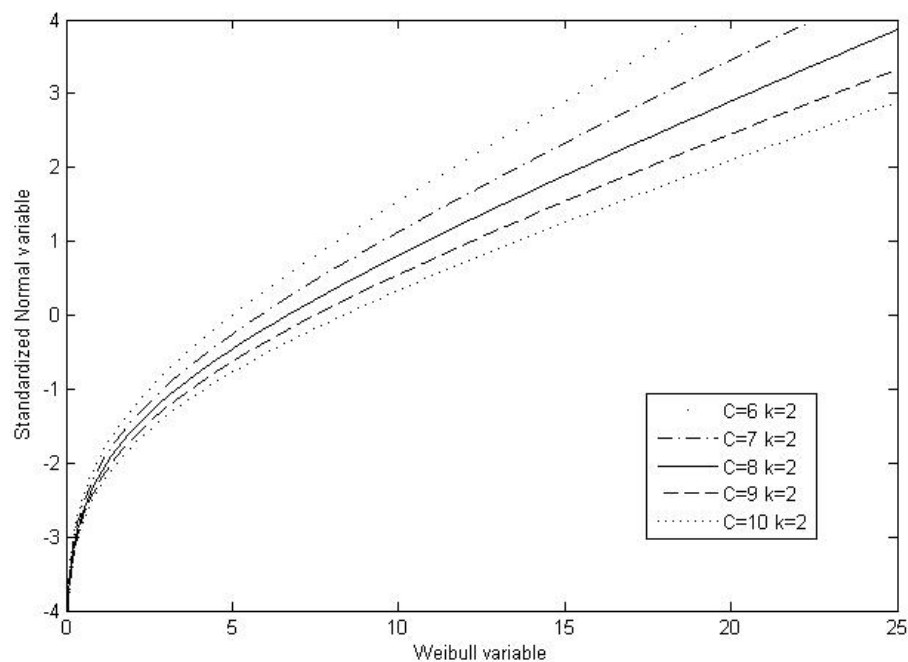
where  $C_i$  and  $k_i$  are the parameters corresponding to the  $i$ th variable.

The inverse transformation is also needed and denoted as  $\text{ntw}^{-1}(u;C,k)$  where  $u$  is the Weibull variable. For a single variable,  $u_i$ , that function is shown in Equation (5):

$$\text{ntw}^{-1}(u_i;C_i,k_i) = \sqrt{2} \text{erf}^{-1} \left( 1 - 2 \exp(-(u_i/C_i)^{k_i}) \right) \quad (5)$$

where  $\text{erf}^{-1}(\cdot)$  is the inverse of the error function [18]. The function  $\text{ntw}^{-1}(u;C,k)$  is shown in Figure 1 for various values of the parameter  $C$ .

**Figure 1.** Standardized Normal variable as a function of a Weibull one, with various values of the scale parameter  $C$ .



The derivative of  $\text{ntw}(\cdot)$  is denoted as  $\text{ntw}'(x;C,k)$  and expressed in Equation (6), also for a single variable.

$$\text{ntw}'(x_i;C_i,k_i) = \sqrt{\frac{2}{\pi}} \frac{C_i}{k_i} \frac{\exp(-x_i^2/2)}{(1 - \text{erf}(x_i/\sqrt{2}))} \left( -\log \left( \frac{(1 - \text{erf}(x_i/\sqrt{2}))}{2} \right) \right)^{1/k_i - 1} \quad (6)$$

In Equation (6) the derivative of the error function is used [18], which is shown in Equation (7):

$$\text{erf}'(x) = \frac{2}{\pi} \exp(-x^2) \quad (7)$$

By using Equations (4–6), the change of variables can be extended to multiple variables.

## 2.2. Multivariate Normal to Weibull Change

In order to broaden the transformation to several variables, the multivariate Standard Normal distribution has to be considered. Its PDF for when the covariance matrix is positive definite is shown in Equation (8):

$$f_{\mathbf{x}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \det(\mathbf{\Sigma})^{1/2}} \exp \left( -\frac{1}{2} \mathbf{x}^t \mathbf{\Sigma}^{-1} \mathbf{x} \right) \quad (8)$$

where  $\mathbf{x}$  is a vector formed by several Standard Normal variables;  $n$  is the number of them;  $\det(\cdot)$  means the determinant of a matrix,  $\Sigma$  is the covariance matrix;  $\Sigma = (\text{Var}_{ij})$   $i, j = 1, \dots, n$ ;  $\text{Var}_{ii} = 1$ ; and  $\text{Var}_{ij} = \rho_{ij}$  ( $i \neq j$ );  $\rho_{ij}$  is the correlation coefficient between variables  $x_i$  and  $x_j$ ;  $\Sigma^{-1}$  means inversion of matrix  $\Sigma$ ; and  $\mathbf{x}^t$  means transposition of vector  $\mathbf{x}$ .

By using Equation (8), the PDF corresponding to the multivariate Weibull distribution is obtained through Equation (9):

$$f_u(\mathbf{u}) = f_x(\mathbf{x}(\mathbf{u})) / |\det(J)| \quad (9)$$

where  $\mathbf{u}$  is the vector formed by several Weibull variables;  $J$  is the Jacobian matrix; and  $||$  means absolute value. The elements of the Jacobian matrix are shown in Equation (10):

$$J = \left( \frac{\partial u_i}{\partial x_j} \right) \quad \begin{cases} \frac{\partial u_i}{\partial x_i} = \text{ntw}'(x_i; C_i, k_i) \\ \frac{\partial u_i}{\partial x_j} = 0 & i \neq j \end{cases} \quad (10)$$

Thus, the Jacobian matrix has non-zero elements in its diagonal. The determinant of this matrix is obtained through the equation expressed in Equation (11):

$$\det(J) = \prod_{i=1}^n \text{ntw}'(x_i; C_i, k_i) \quad (11)$$

where  $\Pi$  means the product of a sequence of terms featuring the  $i$  index.

The multivariate Weibull PDF of a group of variables is shown in Equation (12) as a function of the multivariate Standard Normal PDF and  $\text{ntw}(\cdot)$ .

$$f_u(u_1, u_2, \dots, u_n) = \frac{f_x(\text{ntw}^{-1}(u_1; C_1, k_1), \text{ntw}^{-1}(u_2; C_2, k_2), \dots, \text{ntw}^{-1}(u_n; C_n, k_n))}{\left| \prod_{i=1}^n \text{ntw}'(\text{ntw}^{-1}(u_i; C_i, k_i); C_i, k_i) \right|} \quad (12)$$

Equation (12) can be expressed as in Equation (13) if Equation (8) is taken into account.

$$f_u(u_1, u_2, \dots, u_n) = \frac{\exp\left(-\frac{1}{2}(\text{ntw}^{-1}(u_1; C_1, k_1), \dots, \text{ntw}^{-1}(u_n; C_n, k_n))\Sigma^{-1}(\text{ntw}^{-1}(u_1; C_1, k_1), \dots, \text{ntw}^{-1}(u_n; C_n, k_n))^t\right)}{(2\pi)^{n/2} \det(\Sigma)^{1/2} \left| \prod_{i=1}^n \text{ntw}'(\text{ntw}^{-1}(u_i; C_i, k_i); C_i, k_i) \right|} \quad (13)$$

Notice that Equation (13) depends on two parameters per variable ( $C_i$  and  $k_i$ ) and one for every pair of them ( $\rho_{ij}$ ). However, the correlation coefficient  $\rho$ , between pairs of variables is established according to Standard Normal distributed variables, so the relationship between both parameters has to be obtained. By making a numerical approach based on the Cholesky decomposition (see Appendix), an interval can be obtained for each parameter where the correlation coefficients in Equation (13) can be considered according to Weibull variables, *i.e.*, where the elements of the covariance matrix,  $\Sigma$ , in Equation (13) correspond to the correlation coefficients between pairs of variables  $u_i$  and  $u_j$ . The intervals are shown in Equation (14):

$$C_i \in (0; +\infty) \quad k_i \in [1; +\infty) \quad (14)$$

Therefore, in many cases the parameters used in Equation (13) are defined by the behavior of the group of Weibull variables.

Even though Equation (13) seems a trifle complex, it should be emphasized that, when introducing it in a software application, its complexity does not depend on the number of variables,  $n$ . Moreover, some models referred to in section I serve only the bivariate case and those applied to the multivariate one do not consider the correlation coefficients, which are the classical measurements of dependence.

### 2.3. Multivariate Wind Speed Distribution

According to [14–16] the cumulative behavior of the wind speed at location  $i$ ,  $v_i$ , can be described by a Weibull distribution with parameters  $C_i$  and  $k_i$ , and the relationship between every two distributions can be described by its correlation coefficient [19–22]. So, the multivariate wind speed PDF, or multilocation wind speed PDF, for a group of  $n$  variables  $(v_1, \dots, v_n)$ , with Weibull parameters  $C_i$  and  $k_i$  and correlation matrix  $R_v$  referred to the pairs of Normally distributed variables, is shown in Equation (15):

$$f_v(v_1, \dots, v_n) = \frac{\exp\left(-\frac{1}{2}(\text{ntw}^{-1}(v_1; C_1, k_1), \dots, \text{ntw}^{-1}(v_n; C_n, k_n))\Sigma^{-1}(\text{ntw}^{-1}(v_1; C_1, k_1), \dots, \text{ntw}^{-1}(v_n; C_n, k_n))^t\right)}{(2\pi)^{n/2} \det(\mathbf{R}_v)^{1/2} \left| \prod_{i=1}^n \text{ntw}'(\text{ntw}^{-1}(v_i; C_i, k_i); C_i, k_i) \right|} \quad (15)$$

where  $\text{ntw}^{-1}()$  and  $\text{ntw}'()$  are defined in Equations (5) and (6) respectively.

On the other hand, in most cases the Weibull parameters of the wind speed distributions lie in the intervals expressed in Equation (14), so, as it is explained in Section 2.1,  $R_v$  can also represent the correlation matrix of the  $v_i$  variables.

### 2.4. Multivariate Wind Power Distribution

The relationship between wind speed and wind power [14,16] is defined by Equation (16):

$$P_i = \frac{1}{2} A_i d_i u_i^3 \quad (16)$$

where  $P_i$  is the power contained in an airstream that is flowing through a surface of area  $A_i$  and  $d_i$  is the air density at location  $i$ .

Therefore, applying a change of variables,  $P_i$  can be described by a Weibull distribution of parameters  $C_i'$  and  $k_i'$  [14], expressed in Equation (17):

$$C_i' = A_i d_i C_i^3 / 2 \quad k_i' = k_i / 3 \quad (17)$$

where  $C_i$  and  $k_i$  are the Weibull parameters of the wind speed at location  $i$ . Therefore, as defined in Equation (13) the multivariate wind power PDF is expressed in Equation (18) as a function of the parameters  $C_i'$  and  $k_i'$ .

$$f_P(P_1, \dots, P_n) = \frac{\exp\left(-\frac{1}{2}(\text{ntw}^{-1}(P_1; C_1', k_1'), \dots, \text{ntw}^{-1}(P_n; C_n', k_n'))\Sigma^{-1}(\text{ntw}^{-1}(P_1; C_1', k_1'), \dots, \text{ntw}^{-1}(P_n; C_n', k_n'))^t\right)}{(2\pi)^{n/2} \det(\mathbf{R}_P)^{1/2} \left| \prod_{i=1}^n \text{ntw}'(\text{ntw}^{-1}(P_i; C_i', k_i'); C_i', k_i') \right|} \quad (18)$$

And in Equation (19), as a function of the parameters of the wind speed distribution.

$$f_P(P_1, \dots, P_n) = \frac{\exp\left(-\frac{1}{2}(\text{ntw}^{-1}(P_1; A_1 d_1 C_1^3/2, k_1/3), \dots) \Sigma^{-1} (\text{ntw}^{-1}(P_1; A_1 d_1 C_1^3/2, k_1/3), \dots)^t\right)}{(2\pi)^{n/2} \det(\mathbf{R}_P)^{1/2} \left| \prod_{i=1}^n \text{ntw}'(\text{ntw}^{-1}(P_i; A_i d_i C_i^3/2, k_i/3); A_i d_i C_i^3/2, k_i/3) \right|} \quad (19)$$

In both cases, Equations (18) and (19), the matrix  $R_P$  contains the correlation coefficients corresponding to the pairs of normally distributed variables. In most cases the parameters of the Weibull distributions defined in Equation (17) lie outside the intervals expressed in Equation (14), so, as it is explained in Section 2.1,  $R_P$  does not represent the correlation matrix of the  $P_i$  variables.

### 3. Bivariate Weibull Distribution

In many cases the bivariate Weibull distribution is sought in order to describe the wind speed or wind power behavior in a pair of locations. Due to its importance, we have considered it interesting to develop here as a particular case.

#### 3.1. Bivariate Weibull Distribution Applied to Wind Speed

Equation (9) specifically for  $n = 2$  is shown in Equation (20):

$$f_v(v_1, v_2) = f_x(x_1, x_2) \begin{vmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_1}{\partial x_2} \\ \frac{\partial v_2}{\partial x_1} & \frac{\partial v_2}{\partial x_2} \end{vmatrix}^{-1} \quad (20)$$

And the bivariate Standard Normal PDF is expressed in Equation (21):

$$f_x(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(x_1^2 + x_2^2 - 2\rho x_1 x_2)\right) \quad (21)$$

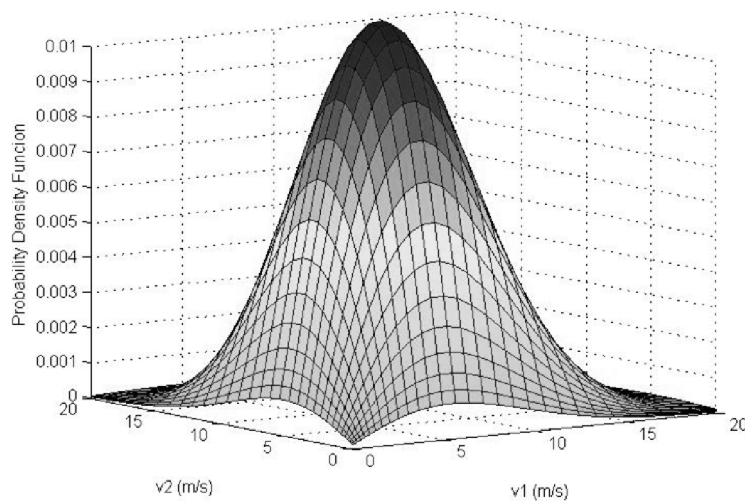
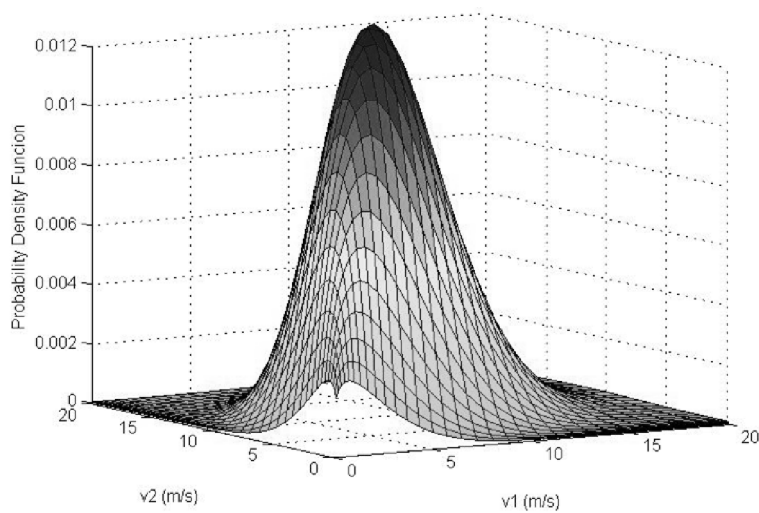
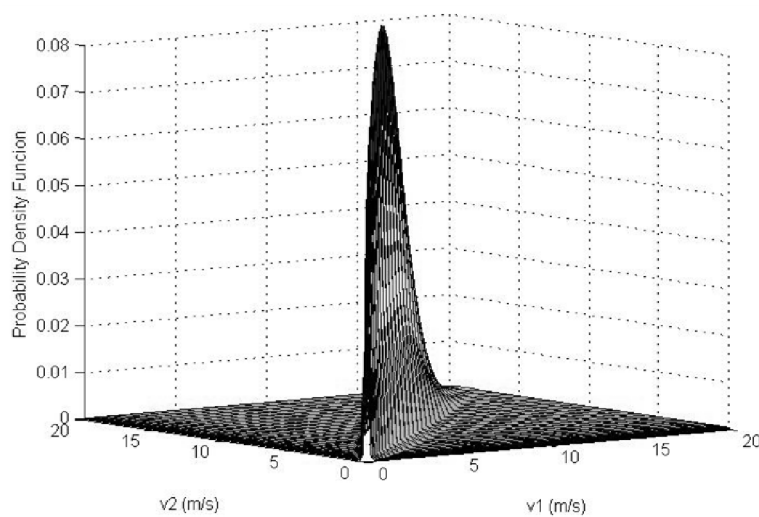
By using Equations (20) and (21) the bivariate Weibull PDF Equation (22) is obtained as a function of  $C_1$ ,  $k_1$ ,  $C_2$ ,  $k_2$  and  $\rho$ , which stands for the correlation coefficient between  $x_1$  and  $x_2$  but, as has been stated above, can be considered as the correlation coefficient between  $v_1$  and  $v_2$ .

$$f_v(v_1, v_2) = \frac{1}{4\sqrt{1-\rho^2}} \cdot \exp\left(-\frac{1}{2(1-\rho^2)}(x_1^2 + x_2^2 - 2\rho x_1 x_2)\right) \cdot \left[ \frac{C_1}{k_1} \frac{\exp(-x_1^2/2)}{(1 - \text{erf}(x_1/\sqrt{2}))} \left(-\log\left(\frac{1 - \text{erf}(x_1/\sqrt{2})}{2}\right)\right)^{\frac{1}{k_1}-1} \cdot \frac{C_2}{k_2} \frac{\exp(-x_2^2/2)}{(1 - \text{erf}(x_2/\sqrt{2}))} \left(-\log\left(\frac{1 - \text{erf}(x_2/\sqrt{2})}{2}\right)\right)^{\frac{1}{k_2}-1} \right] \quad (22)$$

In order to simplify Equation (22), as it depends on  $x_1$  and  $x_2$ , their relationships with  $v_1$  and  $v_2$  are shown in Equation (23):

$$\begin{aligned} x_1 &= \sqrt{2} \text{erf}^{-1}\left(1 - 2\exp(-(v_1/C_1)^{k_1})\right) \\ x_2 &= \sqrt{2} \text{erf}^{-1}\left(1 - 2\exp(-(v_2/C_2)^{k_2})\right) \end{aligned} \quad (23)$$

As stated, in most cases the Weibull parameters to define the wind speed behavior lie in the intervals given in Equation (14), so Normal and wind speed correlation coefficients can be considered equal. The bivariate wind speed PDF for several values of  $\rho$  is shown in Figures 2–4 ( $C_1 = 8$ ,  $k_1 = 2$ ,  $C_2 = 8$ ,  $k_2 = 2$ ).

**Figure 2.** Bivariate wind speed Probability Distribution Function (PDF) for  $\rho = 0.0$ .**Figure 3.** Bivariate wind speed PDF for  $\rho = 0.5$ .**Figure 4.** Bivariate wind speed PDF for  $\rho = 1.0$ .

### 3.2. Correlation Coefficient Inference

In order to perform the correlation coefficient inference between two variables ( $u_1, u_2$ ), the sample correlation coefficient has to be tested with a hypothesis [23–25]. The sample value,  $r$ , is obtained using Equation (24), once the Weibull variables are changed to Normal ones ( $y_1, y_2$ ).

$$r = \frac{\sum (y_{1j} - \bar{y}_1)(y_{2j} - \bar{y}_2)}{\sqrt{\sum (y_{1j} - \bar{y}_1)^2 \sum (y_{2j} - \bar{y}_2)^2}} \quad (24)$$

where  $y_{ij}$  is the  $j$ th sample value of the variable  $y_i$  and  $\bar{y}_i$  is the sample mean of the variable  $y_i$ .

The hypotheses to be checked in this case are the following:

$$\begin{aligned} H_0: \rho_{12}^y &= \rho_0 \\ H_1: \rho_{12}^y &\neq \rho_0 \end{aligned}$$

where  $\rho_0$  is a known value that corresponds to a bivariate Normal distribution, which needs to be tested. In order to do so, a new variable  $z$  is obtained from the sample correlation coefficient  $r$ , as expressed in Equation (25):

$$z = \frac{1}{2} \log \left( \frac{1+r}{1-r} \right) \quad (25)$$

According to [23–25], the variable  $Z$  is Normal distributed with the parameters given in Equation (26):

$$Z \sim N \left( \frac{1}{2} \log \left( \frac{1+\rho_0}{1-\rho_0} \right) + \frac{\rho_0}{2(n-1)}, \frac{1}{\sqrt{n-3}} \right) \quad (26)$$

where  $n$  is the number of elements in the sample.

Depending on the significance level,  $\alpha$ , a confidence interval,  $CI = [z_{\min}, z_{\max}]$ , is established, in which its limits are expressed in an implicit way in Equation (27):

$$\begin{aligned} F_z(z_{\min}) &= \frac{\alpha}{2} \\ F_z(z_{\max}) &= 1 - \frac{\alpha}{2} \end{aligned} \quad (27)$$

where  $F_z()$  is the CDF of the variable  $Z$ , according to the parameters of Equation (26).

So, if  $z$ , obtained through Equation (25), belongs to the CI, the Null Hypothesis  $H_0$  can be accepted with significance level  $\alpha$ , and if not, it cannot be accepted.

If  $H_0$  is accepted, the correlation coefficient corresponding to the bivariate Weibull distribution has to be obtained.

According to previous sections, the correlation coefficients corresponding to the bivariate Normal distribution and to the Weibull one, when it represents a pair of wind speed variables, are approximately the same, so if  $H_0$  is accepted, it can be said that  $\rho_0$  can be taken as an estimation of the correlation coefficient in a bivariate Weibull distribution with significance level of  $\alpha$ .

### 3.3. Bivariate Weibull Distribution Applied to Wind Power

According to Equation (19), the bivariate wind power PDF is expressed in Equation (28):

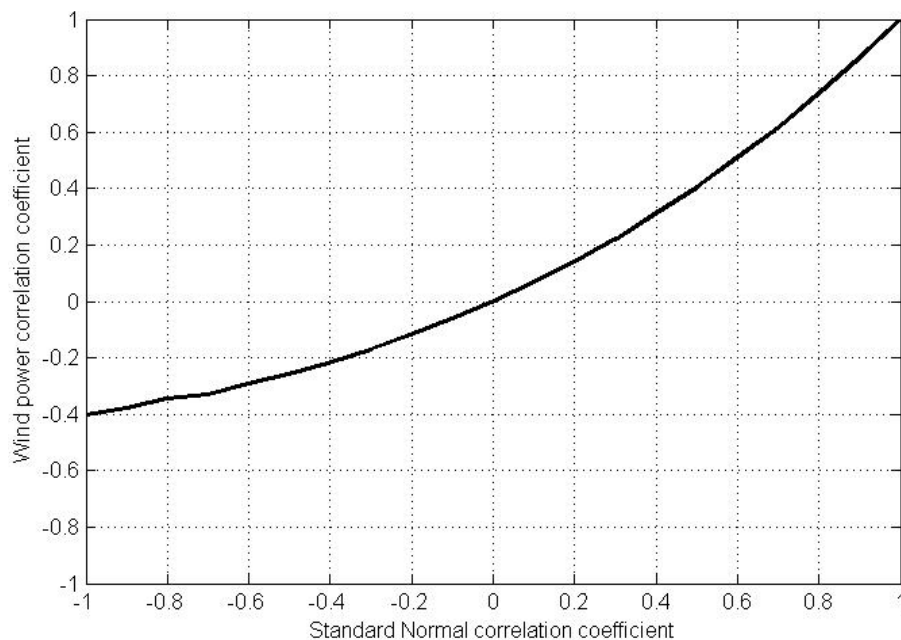
$$f_P(P_1, P_2) = \frac{1}{9\sqrt{1-\rho^2}} \cdot \frac{\exp\left(-\frac{1}{2(1-\rho^2)}(x_1^2 + x_2^2 - 2\rho x_1 x_2)\right)}{\left| \frac{A_1 d_1 C_1^3 \exp(-x_1^2/2)}{k_1 (1 - \operatorname{erf}(x_1/\sqrt{2}))} \left(-\log\left(\left(1 - \operatorname{erf}\left(\frac{x_1}{\sqrt{2}}\right)\right)/2\right)\right)^{\frac{3}{k_1}-1} \frac{A_2 d_2 C_2^3 \exp(-x_2^2/2)}{k_2 (1 - \operatorname{erf}(x_2/\sqrt{2}))} \left(-\log\left(\left(1 - \operatorname{erf}\left(\frac{x_2}{\sqrt{2}}\right)\right)/2\right)\right)^{\frac{3}{k_2}-1} \right|} \quad (28)$$

where  $x_1$  and  $x_2$  are shown in Equation (29):

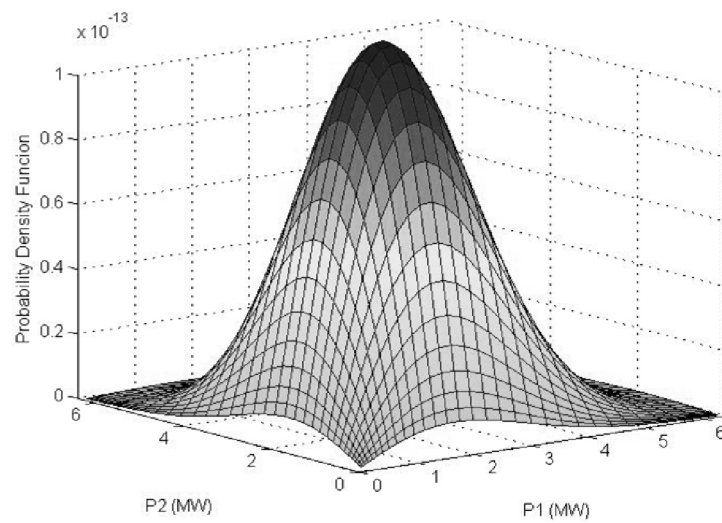
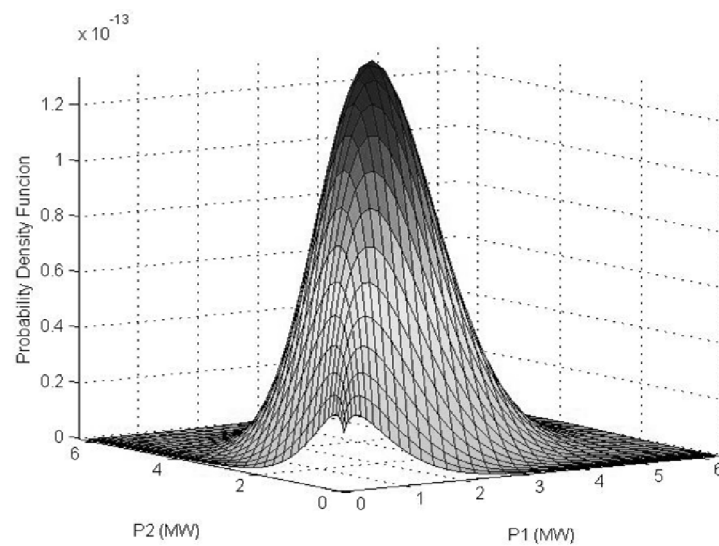
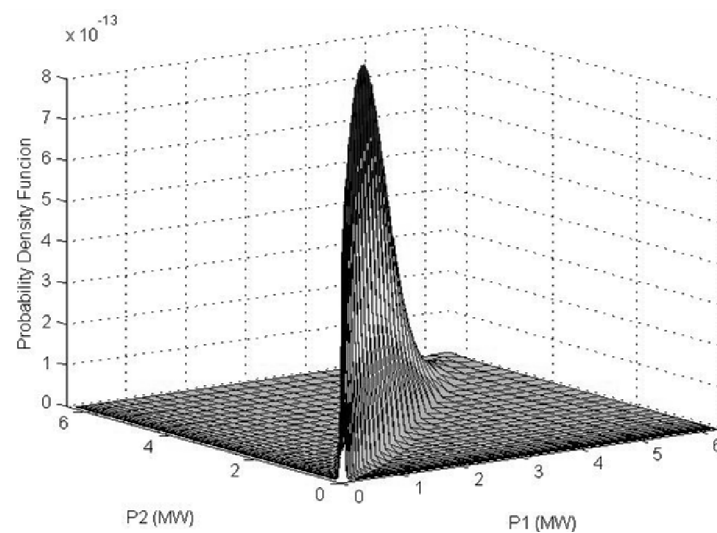
$$\begin{aligned} x_1 &= \sqrt{2} \operatorname{erf}^{-1} \left( 1 - 2 \exp \left( - \left( \sqrt[3]{\frac{2P_1}{A_1 d_1}} / C_1 \right)^{k_1} \right) \right) \\ x_2 &= \sqrt{2} \operatorname{erf}^{-1} \left( 1 - 2 \exp \left( - \left( \sqrt[3]{\frac{2P_2}{A_2 d_2}} / C_2 \right)^{k_2} \right) \right) \end{aligned} \quad (29)$$

The relationship between both parameters is shown in Figure 5.

**Figure 5.** Relationship between Standard Normal and wind power correlation coefficients.



The bivariate wind power PDF for several values of  $\rho$  is shown in Figures 6–8 ( $C_1 = 8$ ,  $k_1 = 2$ ,  $A_1 = 7853 \text{ m}^2$ ,  $d_1 = 1.225 \text{ kg/m}^3$ ,  $C_2 = 8$ ,  $k_2 = 2$ ,  $A_2 = 7853 \text{ m}^2$ ,  $d_2 = 1.225 \text{ kg/m}^3$ ).

**Figure 6.** Bivariate wind power PDF for  $\rho = 0.0$ .**Figure 7.** Bivariate wind power PDF for  $\rho = 0.5$ .**Figure 8.** Bivariate wind power PDF for  $\rho = 1.0$ .

#### 4. Case Study

As a case study, it can be assessed if a certain model of correlation coefficient between a pair of locations can be accepted in order to estimate its value as a function of the distance between them.

The main features of the behavior of the wind in Galicia, in the Northwest of Spain, are that during winter, the winds blow from the Southwest and are very constant and powerful and during the summer, the winds normally blow softly from the Northeast. There are a great number of meteorological stations spread throughout Galicia [26], if wind speed series of data are collected through them, the Weibull distribution parameters for each location can be obtained. Moreover, the correlation coefficients of each pair can be derived according to Equation (24).

In order to estimate the correlation coefficient for locations with a low number of simultaneous sample values of wind speed measures, the relationship between the correlation coefficient and the distance can be analyzed [27]. The great-circle distance between two locations ( $i$  and  $j$ ) can be obtained through Equation (30), as a function of their geographical coordinates.

$$d_{ij} = R_E \cdot \arccos \left( \sin(\text{lat}_i) \sin(\text{lat}_j) + \cos(\text{lat}_i) \cos(\text{lat}_j) \cos(\text{lon}_i - \text{lon}_j) \right) \quad (30)$$

where  $d_{ij}$  is the distance between location  $i$  and  $j$ ;  $R_E$  is the Earth's radius;  $\text{lat}_i$  and  $\text{lon}_i$  are the latitude and longitude coordinates of the location  $i$ ;  $\sin(\ )$  and  $\cos(\ )$  are the sine and cosine functions; and  $\arccos(\ )$  is the inverse of the cosine function.

Therefore, including all the possible pairs (distance, correlation coefficient) in Galicia and by means of the least square method, the relationship Equation (31) is derived.

$$\rho = a \cdot \text{dist} + b \quad (31)$$

where  $\rho$  is the correlation coefficient;  $\text{dist}$  is the distance between locations in km; and  $a$  and  $b$  are parameters obtained for each case. Here,  $a = -0.0007 \text{ km}^{-1}$  and  $b = 0.6589$ .

So, if the correlation coefficient between a location not included in the previous analysis (Coto Muiño), and another one that is included (Melide), needs to be estimated, all that has to be done is to apply Equation (30) between both locations, and then Equation (31), after which a value of  $\rho_0 = 0.6012$  is obtained.

Moreover, by utilizing simultaneous sample data ( $n = 1000$ ) from both locations, the sample correlation coefficient can be obtained through Equation (24),  $r = 0.6202$ , and the change suggested in Equation (25) applied, to obtain  $z = 0.7253$ .

Considering  $\alpha = 0.05$ ,  $Z \sim N(0.6951, 0.0317)$ , the CI obtained is  $CI = [0.6330, 0.7571]$ . Therefore, as explained in the previous sections, as  $z_{\min} < z < z_{\max}$ ,  $\rho_0 = 0.6012$  can be accepted as the correlation coefficient of the bivariate Weibull distributions corresponding to the wind speed data of those locations. The results are shown in Table 1.

**Table 1.** Results of the case study.

Variable	Obtained value	Minimum value	Maximum value
$r$	0.6202	0.5601	0.6394
$z$	0.7253	0.6330	0.7571

## 5. Conclusions

In this paper, a Normal to Weibull change of variables has been defined. It should be noticed that the process can be applied to any type of variables, even inversely. The multivariate Weibull PDF has been obtained and justified, depending on the classic parameters of a single variable and the correlation coefficients between pairs of them. It upgrades former approaches that mainly consist of models based on parameters that have no direct relationship with the univariate parameters and the usual dependence measurement. The function proposed can be easily implemented in a software application regardless of the number of variables. The bivariate case seems a bit complex, compared to other models, but it uses the correlation coefficient between both variables. From the point of view of  $n$  variables, each defined by a Weibull distribution with correlation coefficients between pairs given, the PDF proposed is not an approximation, it provides exact results. Additionally, the application of the multivariate Weibull PDF to wind speed and wind power has been explained and derived. The bivariate case for both has also been specified due to its relevance, and some figures are given for clarification purposes. Moreover, the inference of the correlation coefficient in the bivariate wind speed distribution is explained and applied to a particular case.

## Appendix

The decomposition of a Hermitian, positive-definite matrix into the product of a lower triangular matrix and its conjugate transpose is called the Cholesky decomposition. A Hermitian matrix is a square matrix with complex entries that is equal to its conjugate-transpose.

Therefore, given  $\Omega$ , Hermitian and positive-definite, the Cholesky decomposition consists of obtaining  $L$ , fulfilling Equation (32):

$$\Omega = L \cdot L^* \quad (32)$$

where  $L^*$  means the conjugate transpose of the matrix  $L$ .

The Cholesky decomposition is mainly used for the numerical solution of linear equations, linear least squares problems, non-linear optimization or in Kalman filters. Here it is utilized in another application: the Monte Carlo [28–30] simulation.

Given  $X$ , a matrix of uncorrelated series of samples, and  $\Omega$ , the desired correlation matrix for these series, Equation (33) is applied in order to obtain  $Y$ , a matrix of correlated series of samples according to  $\Omega$ , where  $L$  is the result of the Cholesky decomposition of  $\Omega$ .

$$Y = L \cdot X \quad (33)$$

Moreover, if  $X$  is a matrix of series of samples where each of these series follows a Normal distribution, the resulting matrix in Equation (33),  $Y$ , is also a matrix of series of samples that follows a Normal distribution, which can easily be demonstrated. Equation (33) has the shape shown in Equation (34).

$$\begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ l_{m1} & l_{m2} & \cdots & l_{mm} \end{pmatrix} \cdot \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \quad (34)$$

where  $n$  is the number of samples of each series and  $m$  is the number of series.

The  $j$ th series is formed by  $m$  elements of the type expressed in Equation (35).

$$y_{ji} = \sum_{l=1}^j l_{jl} \cdot x_{li} \quad 1 \leq i \leq n \quad (35)$$

If  $x_{li}$  is distributed according to a  $N(\mu_l, \sigma_l)$  distribution, then  $y_{ji}$  will follow a  $N\left(\sum_{l=1}^j l_{jl} \cdot \mu_l, \sqrt{\sum_{l=1}^j l_{jl}^2 \cdot \sigma_l^2}\right)$ . Moreover, if  $\mu_1 = \mu_2 = \dots = \mu$  and  $\sigma_1 = \sigma_2 = \dots = \sigma$ , then  $y_{ji}$  will follow a  $N\left(\left(\sum_{l=1}^j l_{jl}\right)\mu, \sqrt{\sum_{l=1}^j l_{jl}^2} \sigma\right)$  distribution. And, in the Standard case, if  $\mu = 0$  and  $\sigma = 1$ , then  $y_{ji}$  will be a  $N\left(0, \sqrt{\sum_{l=1}^j l_{jl}^2}\right)$  distribution.

On the other hand, as the condition that  $L$  fulfills is Equation (32), it is always true that  $\sum_{l=1}^j l_{jl}^2 = 1$ , therefore in the particular case  $\mu_1 = \mu_2 = \dots = 0$ ,  $\sigma_1 = \sigma_2 = \dots = 1$ ,  $y_{ji}$  will be distributed by a  $N(0,1)$  distribution.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Weibull, W. A statistical distribution function of wide applicability. *J. Appl. Mech. Trans. ASME* **1951**, *18* (3), 293–297.
2. Devore, J.L. *Probability and Statistics*; Brooks/Cole: Belmont, CA, USA, 2010.
3. Milton, J.S.; Arnold, J.C. *Introduction to Probability and Statistics*; McGraw-Hill: New York, NY, USA, 2002.
4. Pham, H. *Handbook of Engineering Statistics*; Springer: Piscataway, NJ, USA, 2006.
5. Murthy, D.N.P.; Xie, M.; Jiang, R. *Weibull Models*; John Wiley: Hoboken, NJ, USA, 2004.
6. Marshall, A.W.; Olkin, I. A multivariate exponential distribution. *J. Am. Stat. Assoc.* **1967**, *62*, 30–44.
7. Lee, L. Multivariate distributions having Weibull properties. *J. Multivar. Anal.* **1979**, *9* (2), 267–277.
8. Roy, D.; Mukherjee, S.P. Some characterizations of bivariate life distributions. *J. Multivar. Anal.* **1989**, *28*, 1–8.
9. Crowder, M. A multivariate distribution with weibull connections. *J. R. Statist. Soc. B* **1989**, *51* (1), 93–107.
10. Lu, J.; Bhattacharyya, G.K. Some new constructions of bivariate Weibull models. *Ann. Inst. Statist. Math.* **1990**, *42* (3), 543–559.
11. Patra, K.; Dey, D.K. A multivariate mixture of Weibull distributions in reliability modeling. *Stat. Probab. Letters* **1999**, *45*, 225–235.
12. Liu, P.L.; Kiureghian, A.D. Multivariate distribution models with prescribed marginals and covariances. *Probab. Eng. Mech.* **1986**, *1* (2), 105–112.
13. Morales, J.M.; Baringo, L.; Conejo, A.J.; Minguez, R. Probabilistic power flow with correlated wind sources. *IET Gener. Transm. Distrib.* **2010**, *4*, 641–651.

14. Troen, I.; Petersen, E.L. *European Wind Atlas*; Riso National Laboratory: Roskilde, Denmark, 1989.
15. *Wind turbines. Part 1: design requirements*; IEC 61400-1; IEC Standards: Geneva, Switzerland, 2005.
16. Freris, L.L. *Wind Energy Conversion Systems*; Prentice Hall: London, UK, 1990.
17. Stuart, A.; Ord, K. *Kendall's Advanced Theory of Statistics*; Oxford University Press Inc.: New York, NY, USA, 1994.
18. Bronshtein, I.; Semendiaev, K. *Handbook of Mathematics*; Springer: New York, NY, USA, 2007.
19. Lu, X.; McElroy, M.; Kiviluoma, J. Global potential for wind-generated electricity. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10933–10938.
20. Correia, P.F.; Ferreira de Jesús, J.M. Simulation of correlated wind speed and power variates in wind parks. *Electr. Power Syst. Res.* **2010**, *80* (5), 592–598.
21. Segura-Heras, I.; Escrivá-Escrivá, G.; Alcázar-Ortega, M. Wind farm electrical power production model for load flow analysis. *Renew. Energy* **2011**, *36* (3), 1008–1013.
22. Vallée, F.; Lobry, J.; Deblecker, O. System reliability assessment method for wind power integration. *IEEE Trans. Power Syst.* **2008**, *23* (3), 1288–1297.
23. Jobson, J.D. *Applied Multivariate Data Analysis, Volume I: Regression and Experimental Design*; Springer: New York, NY, USA, 1991.
24. Johnson, R.A.; Wichern, D.W. *Applied Mutivariate Statistical Analysis*; Prentice Hall: Upper Saddle River, NJ, USA, 2007.
25. Kleinbaum, D.G.; Kupper, L.L.; Muller, K.E.; Nizam, A. *Applied Regression Analysis and Multivariable Methods*; Thomson Brooks/Cole: Pacific Grove, CA, USA, 1998.
26. Meteogalicia Homepage. Available online: <http://www.meteogalicia.es> (accessed on 1 November 2012).
27. Freris, L.L.; Infield, D. *Renewable Energy in Power Systems*; John Wiley & Sons: Chichester, UK, 2008.
28. Metropolis, N.; Ulam, S. The Monte Carlo Method. *J. Am. Stat. Assoc.* **1949**, *44*, 335–341.
29. Rubinstein, R.Y. *Simulation and the Monte Carlo Method*; Wiley Interscience: Hoboken, NJ, USA, 2008.
30. Gentle, J.E. *Random Number Generation and Monte Carlo Methods*; Springer: New York, NY, USA, 2005.