A JOINT PROBABILITY MODEL FOR ENVIRONMENTAL PARAMETERS: WAVE, WIND AND CURRENT USING NATAF TRANSFORM

Irvin Alberto Mosquera Mosquera

Luís Volnei Sudati Sagrilo

Edison Castro Prates de Lima

iamosqueira@laceo.coppe.ufrj.br

sagrilo@coc.ufrj.br

edison@coc.ufrj.br

LACEO/COPPE/UFRJ

Programa de Engenharia Civil, COPPE/UFRJ

Avenida Pedro Calmon, 21941-596, Rio de Janeiro, Rio de Janeiro, Brazil

Abstract. *This paper investigates the joint probabilistic modeling of the wave, wind and current metocean parameters using a joint probability model based on Nataf's transform. The model considers both the intensity and the direction of environmental parameters. Statistical dependence between variables is modeled by means of their correlation coefficient. A procedure is proposed for calculating the equivalent correlation coefficient for the Nataf model when one of the variables is circular. Using a database gathered in a Gulf of Mexico (GoM) location. The accuracy of the model is verified through artificial simulations using a Monte Carlo-based simulation, where the statistical characteristics of the data numerically generated are compared with the original sample data.*

Keywords: *Metocean data, Joint probability model, Nataf transformation, Circular statistical analysis, Long-term analysis.*

1 INTRODUCTION

Joint probabilistic models (JPM) have been used to describe natural phenomena or behavior of random natural events. Nowadays there is an increasing demand for an effective joint probabilistic model which is able of representing the joint statistical behavior of the metocean parameters associated to wave, wind and current since it is used in several engineering applications (BITNER-GREGERSEN *et al.*, 1996), (BITNER-GREGERSEN and HAVER, 1991), (SOUKISSIAN, 2014), (PAPALEO, 2009). The most of the available models take into account just the wave parameters: significant wave height (Hs) and zero crossing period (Tz) or peak period (Tp). In some cases the wind velocity is also included in these models (BRODTKORB, 2005) ,(JOHANNESSEN, MELING and HAVER, 2001). The main reason for this limitation is that JPMs are difficult to be fitted when more than two variables are taken into account or when the data set available is limited.

The direction of the variable is extremely important for some analyses that are heavily dependent on the directionality of the environmental actions. Most of the JPMs available not consider the directionality of the variable modeled due to the fact that it is not easy to establish a JPM when one of the variables is circular. Circular data need special treatment since the circular random variables are limited to the circle, i.e., they are limited in the interval $[0,2\pi]$ or $[-\pi, \pi]$ where both limits are physically the same point. Long-term response based design of turret-moored floating, production, storage and offloading vessels (FPSOs) are very sensitive to directionality of the wave, wind and current. Other projects, such the best location study for the installation of marine power generation turbines, have a strong dependence on waves, winds and currents. For example, in the design of Ocean Wave Energy Converters (OWECs), which are devices for generating energy from waves, the selection and specification of the device depends on the direction of propagation of the waves (SOUKISSIAN, 2014).

The JPMs are usually based on three statistical models: Conditional Modelling Approach (CMA), Copula model and the model based on the Nataf transformation (Nataf Model), each with its own features and limitations. The CMA uses the concept of conditional distributions (BITNER-GREGERSEN and HAVER, 1991), however, due to the practical difficulty of obtaining conditional distributions when the dependence needs to be expressed on more than two parameters; this model is frequently limited to the modelling of up to three parameters. In the Nataf model (KIUREGHIAN and LIU, 1986) the JPM is represented using only the marginal distributions of each variable and the correlation matrix. In principle, this model has no limitations on the number of variables to be represented, as will be seen later, some difficulties occur when the variables are different in nature, such as angular variables mixed with circular variables. The Copula model (LEBRUN and DUTFOY, 2009) consists of a multidimensional model for the complete modeling of stochastic dependency of various statistical variables in which the multi-dimensional functions are transformed into uniform marginal product of one-dimensional functions.

A methodology based on the Nataf transformation (NATAF, 1962) to create a JPM of the environmental parameters taking in account the statistical correlations between the intensity and direction of variable modeled has been recently proposed (PAPALEO, 2009; SAGRILO, PRATES DE LIMA and PAPALEO, 2011). That model has been developed to represent ten metocean parameters.

In this paper the Nataf-based model developed by (PAPALEO, 2009) is improved in order to make its practical application simpler when creating a JPM for environmental parameters including their directions. The main improvement is related to the calculation of the equivalent correlation for the Nataf model when one of the variable is circular. Some concepts related to linear-linear, linear-circular, circular-circular correlation, circular statistic and Nataf transformation are presented in this work in order to develop the model. The model is developed using 21807 registers of 7 simultaneous short-term metocean data of wave, wind and current measured in a Gulf of Mexico location. The data were measured between 2010-2013 at the buoy 42020 (*Corpus Christi, TX)* and are available online in the website of the *National Oceanic and Atmospheric Administration (*NOAA). The numerical example considers seven variables: the significant wave height (*Hs*), zero-up crossing period wave (*Tz*), direction wave (θ_w), amplitude (V_v) and direction (θ_v) of the 1-h average wind velocity and the amplitude (V_c) and direction (θ_c) of the surface current velocity. The Table 1 shows the statistical properties: mean, standard deviation and others parameters for linear variables.

Table 1. Statistic parameters for linear data

Var.	Mean	Median	Standard deviation	Coef. of Variation	Skewness	Kurtosis	Maximum Observ.
H_s	1.300	1.15	0.672	0.517	1.210	5.752	6.940
T_Z	4.742	4.71	0.789	0.166	0.435	3.614	11.450
V_V	6.576	6.5	2.781	0.423	0.300	2.945	17.900
V_C	0.206	0.18	0.135	0.657	1.117	4.553	0.950

2 JOINT PROBABILITY MODEL

Using the Nataf transformation (KIUREGHIAN and LIU, 1986), a joint probability density function can be established as a function of the marginal distribution of each variate and the correlation coefficient between each pair of them. By mean of statistical theory it is possible to obtain a set of uncorrelated standard normal $\mathbf{U}$ variables associated a set of $\mathbf{X}$ random variables by the following transformation:

$$\mathbf{U} = \mathbf{\Gamma}_0 \cdot \mathbf{Z} = \mathbf{\Gamma}_0 \cdot \begin{Bmatrix} \Phi^{-1}\left[F_{X_1}(x_1)\right] \\ \vdots \\ \Phi^{-1}\left[F_{X_N}(x_N)\right] \end{Bmatrix} \tag{1}$$

where $\mathbf{Z}$ are a correlated standard normal variables, $\Phi^{-1}[.]$ is the inverse of the standard normal cumulative distribution, $\mathbf{\Gamma}_0 = \mathbf{L}^{-1}$, $\mathbf{L}$ is the lower triangular matrix obtained from the Cholesky decomposition of the matrix $\mathbf{R}$, where $\mathbf{R}$ is the matrix of correlation coefficients of the variables $\mathbf{Z}$. Each element $r_{i,j}$ of the matrix $\mathbf{R}$ is also known as Nataf correlation coefficient, for each pair of variables it is calculated by solving the following equation (KIUREGHIAN and LIU, 1986).

$$\bar{\rho}_{ij} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{F_{X_i}^{-1}(\Phi(z_i)) - \mu_i}{\sigma_i} \right) \left(\frac{F_{X_j}^{-1}(\Phi(z_j)) - \mu_j}{\sigma_j} \right) \phi_2(z_i, z_j, r_{i,j}) dz_i dz_j \tag{2}$$

where $\bar{\rho}_{ij}$ is the linear correlation coefficient between the parameter X_i and X_j, $F_{X_i}^{-1}(.)$, $F_{X_j}^{-1}(.)$, μ_i, μ_j, σ_i, σ_j are the associated inverse cumulative probability function, mean, and standard deviation of each variable, $\phi_2(z_i, z_j, r_{i,j})$ is the joint probability density function of two standard normal variables z_i and z_j with correlation coefficient $r_{i,j}$. Some approximated solutions for Eq. (2) are presented in (KIUREGHIAN and LIU, 1986). Nevertheless, it is straightforward to calculate the equivalent Nataf coefficient for any pair of linear variables. Using the rules of the probability transformation, the joint probability density function for **N** variables $\mathbf{X}$ is

$$f_{\mathbf{X}}(\mathbf{x}) = \phi(\mathbf{z}, \mathbf{R}) \frac{f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_N}(x_N)}{\phi(z_1)\phi(z_2)\cdots\phi(z_N)} \tag{3}$$

where $f_{Xi}(x_i)$ is the marginal probability density function for the random variable X_i, $\phi(\mathbf{z},\mathbf{R})$ is the joint probability density function of ***N*** standard normal variables $\mathbf{Z}$ and $\phi(.)$ is the probability density function of standard normal random variable.

For the vector $\mathbf{X}$ made up of the metocean parameters studied in this paper

$$\mathbf{X} = (x_1 = Hs, x_2 = Tz, \ldots, x_7 = \theta_C) \tag{4}$$

the joint probability density function using Nataf transform is given by

$$f_X(x) = \frac{\prod_{i=1}^{7} f_{X_i}(x_i)}{\prod_{i=1}^{7} \left(\phi\left(\Phi^{-1}\left(F_{X_i}(x_i) \right) \right) \right)} \phi_7 \left[\left(\Phi^{-1}\left(F_{X_1}(x_1) \right) \right), \ldots, \left(\Phi^{-1}\left(F_{X_7}(x_7) \right) \right), \mathbf{R}_{7x7} \right] \tag{5}$$

Lognormal and Weibull (with two and three parameters) distributions were fitted to the linear data available. The parameters of each distribution were determined using the method of moments (ANG and TANG, 1975). Table 2 shows the distribution fitted for each linear variable and Figs. 1-4 show the fitted marginal distribution and the histogram for all linear variables.

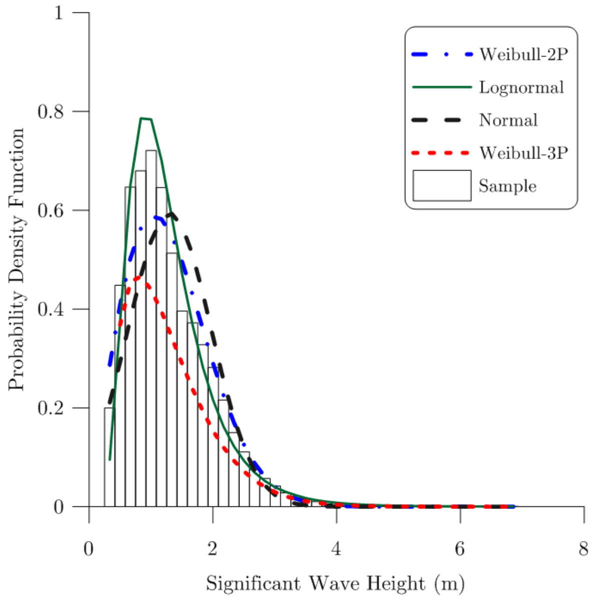

Figure 1. Marginal probability for significant wave height

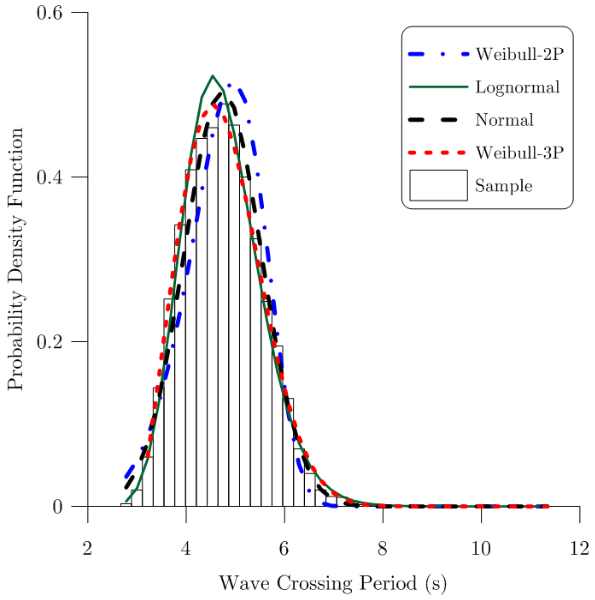

Figure 2. Marginal probability for wave zero crossing-up period

Table 2. Fitted distribution for linear variable

	Distribution function.
H_S	Lognormal
T_Z	Lognormal
V_V	Weibull-3P
V_C	Weibull-2P

Concerning the circular variables, it is not an easy task to obtain their marginal distributions and the Nataf equivalent coefficient for circular-circular and linear-circular variables. Any circular variable requires a special treatment and there is a field in statistics that deals with this topic: Circular statistics. Some probability distributions and correlation measures for circular data are available (MARDIA,1972; FISHER,1993). The Von Mises , Wrapped Normal and Cardioid distributions are the most used distributions for circular data. The C-linear and T-Linear associations are the most common correlation measures of circular-linear and circular-circular variables, respectivelly. Further details about marginal and correlation measures for circular variables will be discussed in the next section.

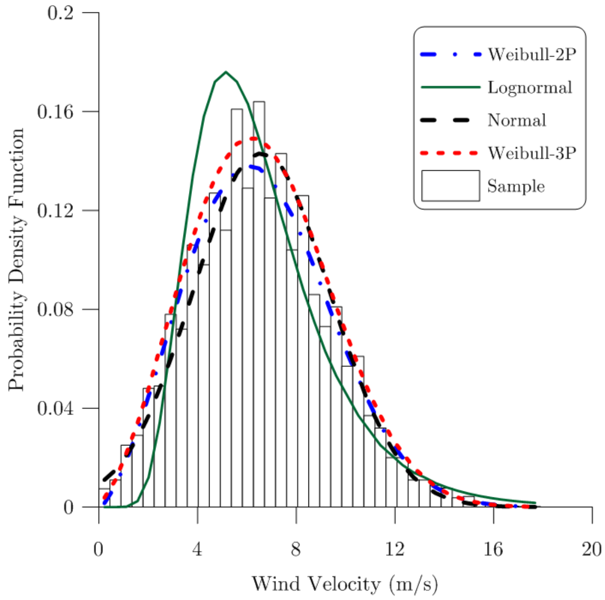

Figure 3. Marginal probability for wind velocity

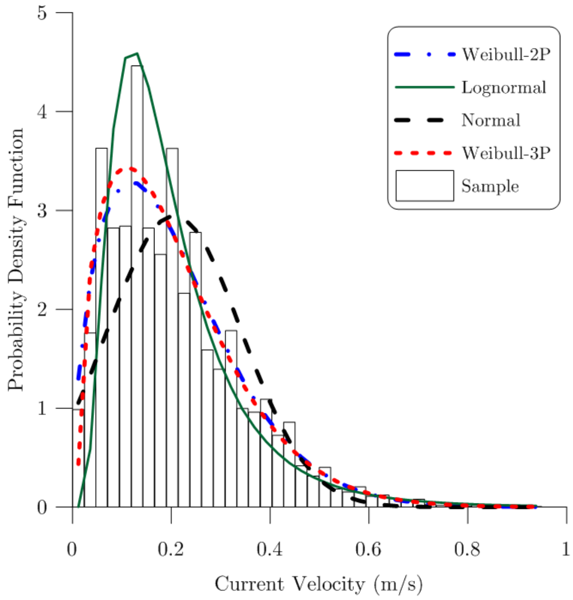

Figure 4. Marginal probability for superficial current velocity

2.1 Probability Distributions for Circular Data

The direction of wave, wind and superficial current are circular variables since the all data are enclosed in a unitary circle. Circular statistics takes into account the natural periodicity of the circle, i.e., $0^o = 360^o$. Any temping of using the conventional measures for the line could produce unrelistic results when applied to a circular variable. To describe the circular variable there are specific statistical measures such as: mean, variance, correlation, etc. In the same way there are specific forms of probability distributions applied only to circular data. Let $\boldsymbol{\theta} = (\theta_1, \theta_2 \cdots \theta_N)$ be a sample of a circular variable, ten the mean direction is given by (FISHER, 1993):

$$\bar{\theta} = a\cos\left(\frac{C}{R}\right) \quad \bar{\theta} = a\sin\left(\frac{S}{R}\right) \tag{6}$$

where R is the the resultant length defined as

$$R^2 = C^2 + S^2 \quad (R \geq 0) \tag{7}$$

and

$$C = \sum_{i=1}^{N} \cos(\theta_i) \qquad S = \sum_{i=1}^{N} \sin(\theta_i) \tag{8}$$

The sample standard deviation is defined as

$$\upsilon = \left\{ -2\log\left(\frac{R}{N}\right) \right\}^{\frac{1}{2}} \tag{9}$$

Table 3 presents the statistical parameters for all circular variables investigated in this work.

Table 3. Circular parameter data sample

Var.	Mean Direction	Mean resultant length	Standard deviation
θ_W	115.686	0.745	0.506
θ_V	121.734	0.546	0.725
θ_C	-6.747	0.313	1.004

Any probability density function (pdf) for circular data is a 2π-periodic function. There are many pdf models for circular variables such as: Uniform, Wrapped Normal, Von Mises, Cardioid, among others. The Von Mises and the Wrapped Normal are the most widely used (MARDIA, 1972). The Wrapped Normal distribution is closely approximated by the Von Mises distribution, so in practice one uses whichever is more convenient (FISHER, 1993). In this work will be use the Wrapped Normal because is more convenient to the task at hand and it has many properties related to the Normal distribution for data represented on the line. The Wrapped Normal distribution is given by

$$f(\theta) = \frac{1}{2\pi}\left(1 + 2\sum_{p=1}^{\infty}\left(\rho_\theta^{\,p^2} \cos p(\theta - \mu_\theta)\right)\right) \qquad 0 \le \theta \le 2\pi \tag{10}$$

where ρ_θ and μ_θ are the resultant length mean and circular population mean (assuming the methods of moments $\rho_\theta = \frac{R}{N}$ and $\mu_\theta = \bar{\theta}$), respectively. The Wrapped Normal is a symmetric two-parameter distribution that can be obtained by wrapping the Normal or Gaussian distribution (on the line) around the circle. If is thought of as being obtained by wrapping a Normal distribution with variance σ^2 and mean μ, the the circular and linear parameters are related as

$$\sigma = \{-2\log(\rho_\theta)\}^{\frac{1}{2}} \tag{11}$$

and

$$\mu_\theta = \mu \tag{12}$$

If X is any random variable on the real line being wrapped on the circle, the corresponding angular variable can be obtained by

$$\theta = X[\mathrm{mod}\, 2\pi] \tag{13}$$

Some circular data present several modes. In order to fit a circular pdf for these cases it is necessary to employ a multimodal distribution. A multimodal function could be obtained by a mixture of unimodal distributions in the following way

$$f(\theta) = \sum_{i=1}^{k} \alpha_i f_i(\theta) \qquad \left[0 \leq \alpha_i \leq 1 \quad \sum_{i=1}^{k} \alpha_i = 1 \right] \tag{13}$$

$$F(\theta) = \sum_{i=1}^{k} \alpha_i F_i(\theta) \qquad \left[0 \leq \alpha_i \leq 1 \quad \sum_{i=1}^{k} \alpha_i = 1 \right] \tag{14}$$

where k is the number of modes, α_i are the model weights, $F_i(\theta)$and $f_i(\theta)$ are the unimodal circular probability density and cumulative distributions centered at the individual ith mode. Some numerical methodologies to calculate the parameters for each mode when the unimodal circular distributions are described by the von Mises distribution have been investigated by (HORNIK and GRUN, 2014). In this work just Wrapped normal distributions were used. The parameters for each mode were fitted using an iterative visual methodology:

- draw the sample data frequency histogram;
- identify the number of modes or peaks;
- for each mode to associate a Wrapped Normal unimodal distribution; the mean of each unimodal distribution has to be the center of the mode;
- set a initial standard deviation to Wrapped Normal unimodal for each mode;
- set a relative weight for each mode; as an initial estimate can be take the highest frequency value of each mode. The sum of the weights must be equal to 1;
- vary the standard deviation and weights iteratively until to obtain a good data fit.

Figs. 5-7 show the data frequency histogram and its corresponding fitted distribution for wave, wind and current data directions, respectively. Table 4 presents the modes and respective parameters for each directional data analyzed.

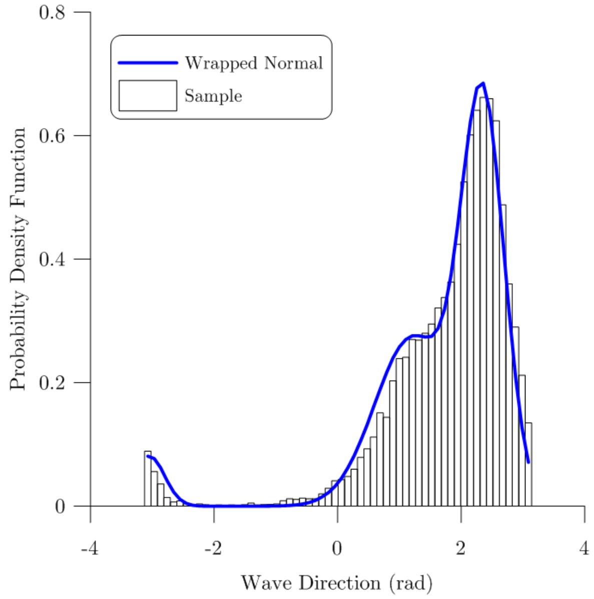

Figure 5. Marginal probability for wave direction

In many cases it is difficult to fit a multi-modal probability function. Then, in this work it is proposed an alternative manner to obtain this function. The circular distribution is represented by means of a real function defined in the interval (0, 2π) taking advantage of the fact that for any circular distribution $F_\theta(0)=0$ and $F_\theta(2\pi)=1$. Using this hypothesis is possible to represent any circular cumulative distribution by means of the empirical distribution of the sample data, avoiding the use of the wrapped distribution. The empirical distribution for the ordered (in ascending order) data sample $\mathbf{\Theta}=[\Theta_1,\Theta_2,\ldots\Theta_N]$ in ascending order is defined by

$$\overline{F}(\Theta_i)=\frac{i}{N+1} \tag{18}$$

The figures 8–10 show the interpolated cumulative function and the cumulative wrapped normal fitted for the wave, wind and superficial current direction.

2.2 Data Correlation

The Nataf model needs the definition of Nataf correlation coefficients matrix $\mathbf{R}$. Since the model developed in thois work involves linear and circular variables, there are three types of correlation to be accounted for: linear-linear, linear-circular and circular-circular. These three types are described in what follows.

2.2.1 Linear-Linear Correlation

The correlation coefficient between two linear variables X_1 and X_2 is described by the Pearson or linear correlation coefficient as

$$\bar{\rho} = E\left[\left(\frac{X_1-\mu_1}{\sigma_1}\right)\left(\frac{X_2-\mu_2}{\sigma_2}\right)\right] \tag{15}$$

where μ_1, μ_2 and σ_1, σ_2 are the mean and standard deviation of both variables, respectively.

The linear and the equivalent correlation are related by means of Eq. (2). The equivalent correlation between pairs of linear variables were calculated using the Eq. (2) and are shown in Table 5. It can be observed a strong correlation between significant wave height and zero crossing period wave and also between the significant wave height and wind velocity. However, no significant correlation is observed between the superficial current velocity and any other linear variable.

Table 4. Parameters for marginal circular distribution

	Modes number´s	Weight	Standard deviation	Mean
		0.55	0.35	2.35
θ_W	3	0.05	0.25	-3.05
		0.40	0.6	1.2
		0.64	0.45	2.6
θ_V	4	0.04	0.12	-3.1
		0.16	0.45	0.0
		0.16	0.50	10
		0.48	0.50	0.1
θ_C	4	0.34	1.20	-2.2
		0.09	0.50	2.9
		0.09	0.40	1.8

Table 5. Correlation Coefficient for Nataf model linear variables

	H_S	T_Z	V_V	V_C
H_S	1.000	0.760	0.716	-0.077
T_Z	0.760	1.000	0.171	-0.065
V_V	0.716	0.171	1.000	-0.034
V_C	-0.077	-0.065	-0.034	1.000

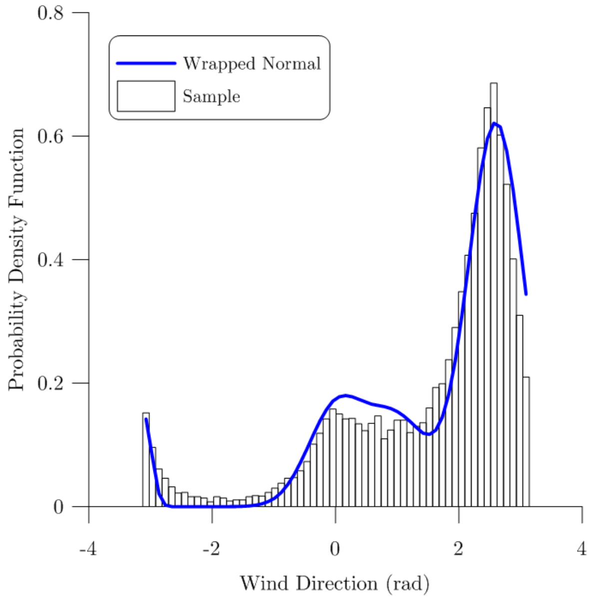

Figure 6. Marginal probability for wind direction

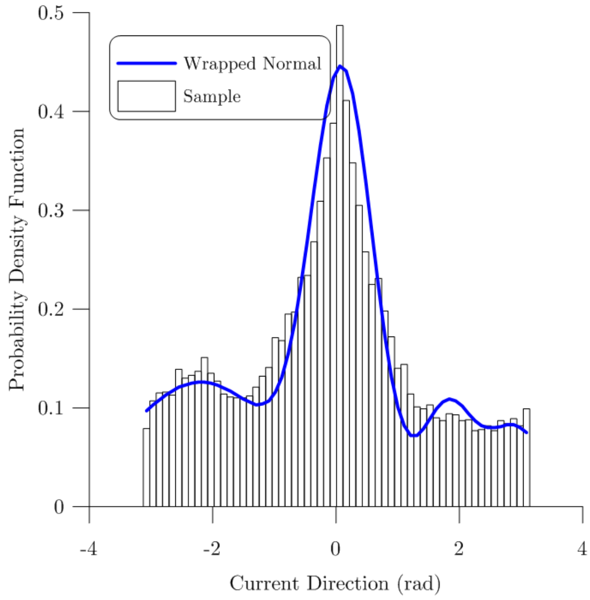

Figure 7. Marginal probability for superficial current direction

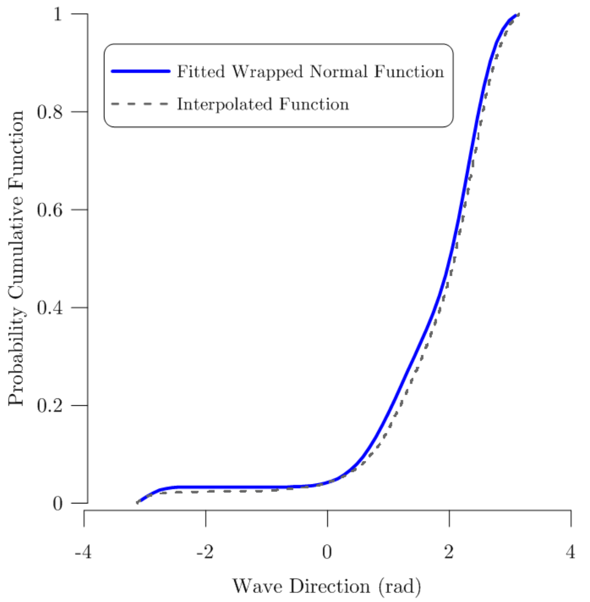

Figure 8. Interpolated and marginal cumulative function for wave direction

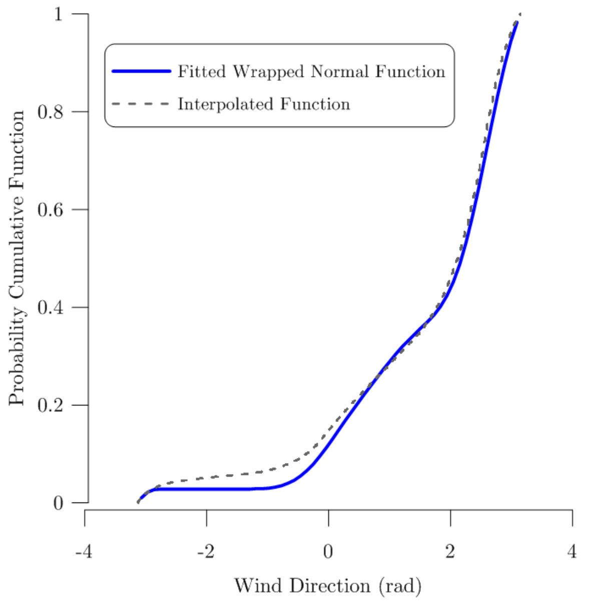

Figure 9. Interpolated and marginal cumulative function for wind direction

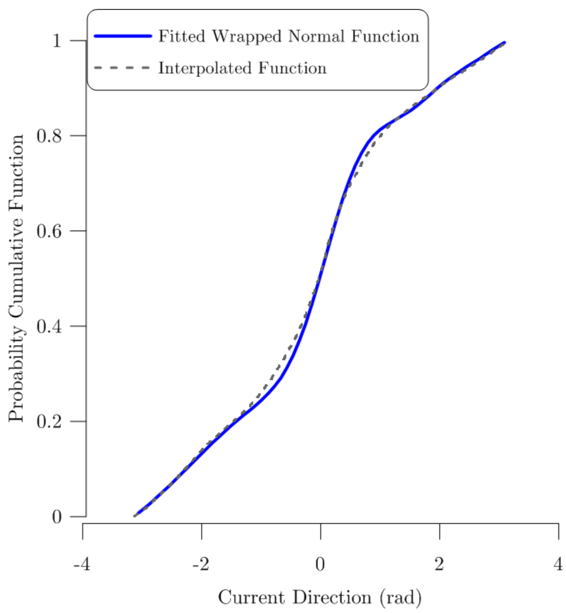

Figure 10 Interpolated and marginal cumulative function for superficial current direction

2.2.2 Circular-Circular

One correlation measure between two circular variables is described by (FISHER, 1993). This correlation describes an association analogue to the linear association between two linear variables. For two circular variables Θ and ϕ with sample simultaneous data, $\mathbf{\Theta} = [\Theta_1, \Theta_2, \ldots \Theta_N]$ and $\phi = [\phi_1, \phi_2, \ldots \phi_N]$, the circular correlation is given by the following equation

$$\rho_{\Theta,\phi} = \frac{4(AB - CD)}{\left((N^2 - E^2 - F^2)(N^2 - G^2 - H^2)\right)^{\frac{1}{2}}} \tag{16}$$

where

$$A = \sum_{i=1}^{N} \cos(\phi_i)\cos(\Theta_i) \qquad B = \sum_{i=1}^{N} \sin(\phi_i)\sin(\Theta_i)$$

$$C = \sum_{i=1}^{N} \cos(\phi_i)\sin(\Theta_i) \qquad D = \sum_{i=1}^{N} \sin(\phi_i)\cos(\Theta_i)$$

$$E = \sum_{i=1}^{N} \cos(2\phi_i) \qquad F = \sum_{i=1}^{N} \cos(2\Theta_i)$$

$$G = \sum_{i=1}^{N} \sin n(2\phi_i) \qquad H = \sum_{i=1}^{N} \sin(2\Theta_i)$$

The circular data needs to be represented on the real line in the Nataf´s model. Then, it is also necessary to obtain the correlation for linear representation of two circular variables. When both Θ and ϕ are unimodal Wrapped Normal variables it is possible (JAMMALAMADAKA and RAMAKRISHANA, 1988) to calculate the linear correlation coefficient $\rho^{L}_{\Theta,\phi}$ by the following equation

$$\rho^{L}_{\Theta,\phi} = \frac{\sinh(2\sigma_{\Theta,\phi})}{\sqrt{\left(\sinh(2\sigma^2_{\Theta})\sinh(2\sigma^2_{\phi})\right)}} \tag{17}$$

where σ_{Θ}, σ_{ϕ} and $\sigma_{\Theta,\phi}$ are the linear standard deviations and covariance of Θ and ϕ respectively.

A numerical procedure to calculate the equivalent correlation coefficient for two circular variables was proposed by (PAPALEO,2009; SAGRILO, PRATES DE LIMA and PAPALEO,2011) which allows to calculate the equivalent correlation coefficient for any type of circular distribution, considering both single or multi-modal cases. This procedure can be summarized in the following steps:

1. calculate the circular correlation $\rho_{\Theta,\phi}$ from available data using Eq. (16);
2. by using Eq. (18) obtain the empirical circular probability functions $F_{\Theta}(\Theta)$ and $F_{\phi}(\phi)$ and represent them by means of an interpolation function (e.g. using splines);
3. set a guess value for the equivalent Nataf correlation coefficient $r_{\Theta,\phi}$;
4. using $F_{\Theta}(\Theta)$, $F_{\phi}(\phi)$ and $r_{\Theta,\phi}$ generate numerically N pairs of value **(**Θ,ϕ) and transform them to the circle by means of Eq. (13). (N must be sufficiently large in order to reduce the statistical uncertainty of the correlation estimates bellow);
5. use Eq. (16) to calculate the circular correlation $\rho^{S}_{\Theta,\phi}$ for the sample generated above;
6. repeat steps 4 to 6 until $\rho^{S}_{\Theta,\phi} = \rho_{\Theta,\phi}$ is nearly equal to $\rho_{\Theta,\phi}$ obtained in step (1).

Tables 6 and 7 show the Circular-Circular sample correlation and the equivalent Nataf correlation between direction the directional variables of the developed model. The strongest correlation is between wave (θ_w) and wind (θ_v) directions.

Table 6. Circular-Circular Correlation

	θ_W	θ_V	θ_C
θ_W	1.000	0.490	6.34E-03
θ_V	0.490	1.000	-3.57E-03
θ_C	6.34E-03	-3.57E-03	1.000

Table 7. Equivalent Nataf Correlation for Circular-Circular data

	θ_W	θ_V	θ_C
θ_W	1.000	0.820	0.010
θ_V	0.820	1.000	-0.009
θ_C	0.010	-0.009	1.000

2.2.3 Circular-Linear

Considering N samples of a linear and a circular variable observed simultaneously, i.e., $\mathbf{X} = [x_1, x_2, \ldots x_N]$ and $\mathbf{\Theta} = [\Theta_1, \Theta_2, \ldots \Theta_N]$, an associated correlation coefficient defined by (FISHER, 1993), known as C association, is computed by

$$\rho^2_{X,\Theta} = \frac{\rho^2_{xc} + \rho^2_{xs} - 2\rho_{xc}\rho_{xs}\rho_{cs}}{1 - \rho^2_{cs}} \tag{19}$$

where $\rho_{xc} = \rho(X, \cos(\Theta))$, $\rho_{xs} = \rho(X, \sin(\Theta))$ and $\rho_{cs} = \rho(\cos(\Theta), \sin(\Theta))$.

A similar procedure as described above can be applied to obtain the equivalent Nataf correlation coefficient when the circular data is represented on the line. The main point is that signal of the equivalent coefficient must be obtained based on visual observations of the original and the artificially generated data samples (PAPALEO,2009). To avoid this problem, a new procedure is developed in this work.

A correlation coefficient analogous to the Pearson (or linear) coefficient for one linear variable and other circular could be adapted from the correlation definition show in (JAMMALAMADAKA and RAMAKRISHANA, 1988) which reads

$$\rho_{X,\Theta} = \frac{E\left(\sin(\Theta - \mu_\theta)(X - \mu_x)\right)}{\left(E\left(\sin(\Theta - \mu_\theta)^2\right)E\left((X - \mu_x)^2\right)\right)^{\frac{1}{2}}} \tag{20}$$

where μ_θ and μ_x are the means of circular variable Θ and linear variable X, respectively. In fact, this correlation measure is a not an appropiated correlation measure for multimodal data. Instead, it just allow to one to obtain the correct sign of the data correlation coefficient defined in Eq. (19). In summary the use of Eq. (20) avoids the visual inspection required in the original work proposed by PAPALEO (2009). The rest of the procedure is very similar.

Figs. 11-13 show a bi-dimensional histogram for the wind velocity (V_V) and wind direction (θ_V) for the original data sample and for two other data samples artificial generated by Monte Carlo simulation with equivalent Nataf correlations $\rho = \pm 0.144$**.** Since the correlation is small, is not easy to determine the correct signal just using a visual inspection. Eq. (20) indicates that the correct value is $\rho = -0.144$ for the equivalent Nataf correlation. Figs. 14-16 show similar analysis for zero crossing period wave (T_Z) and wave direction (θ_W). Using the numerical procedure proposed by (PAPALEO, 2009) the equivalent correlation is ρ = -0.39, since the positive value leads to distorted histogram compared to the one obtained from the original data. The use of Eq. (20) gives a value ρ = -0.21. As mentioned before, despite of the numerical value obtained, Eq. (20) is useful to obtain the correct signal of the correlation. It is clear by the illustrations that the correlation is negative for those variables.

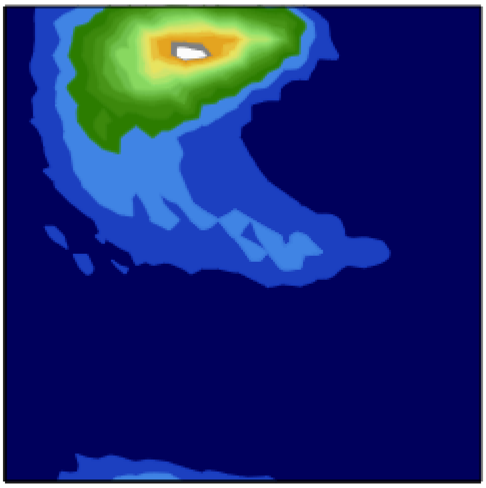

Figure 11 Data sample histogram

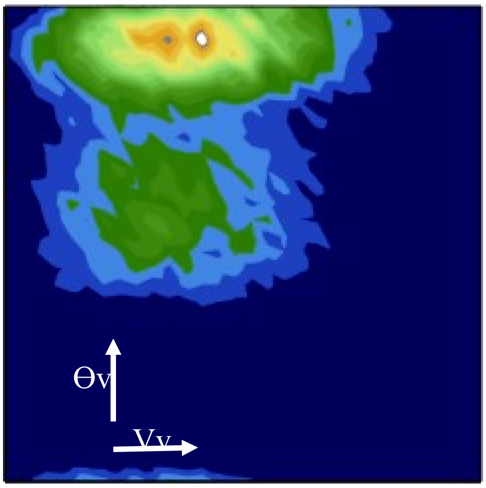

Figure 12 Histogram for positive correlation

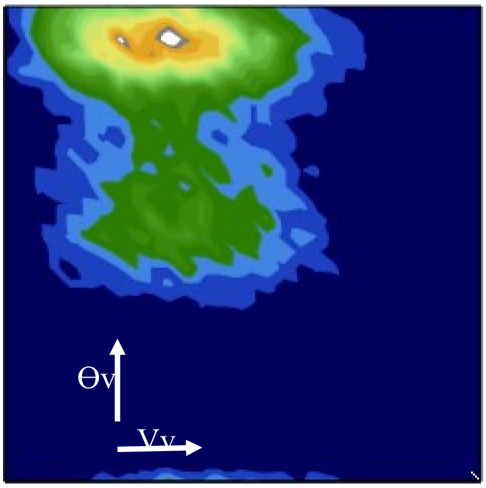

Figure 13. Histogram for negative correlation

Tables 8 and 9 presents the Circular-Linear sample correlations and the equivalent Nataf correlations for the parameters investigated in this work. It is observed that the most significant dependence is between V_C and θ_C, H_S and θ_W and H_S and θ_V

Table 8. Circular-Linear Correlation

	H_S	T_Z	V_V	V_c
θ_W	0.207	0.304	0.258	0.082
θ_V	0.234	0.259	0.121	0.04
θ_C	0.044	0.011	0.055	0.215

Table 9. Equivalent Nataf Correlation for Circular-Linear data

	H_S	T_Z	V_V	V_c
θ_W	-0.240	-0.340	-0.290	0.090
θ_V	-0.331	-0.340	-0.160	0.056
θ_C	0.080	0.001	0.100	0.440

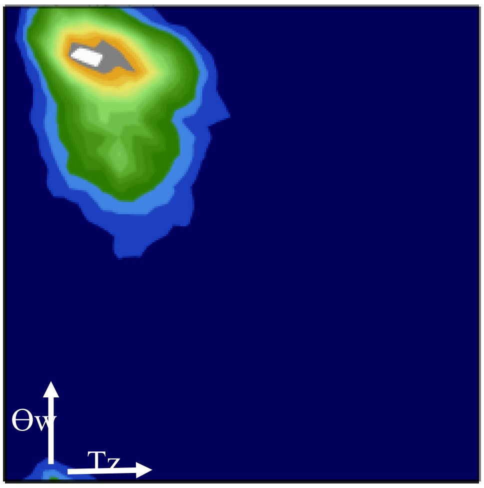

Figure 14. Data sample histogram

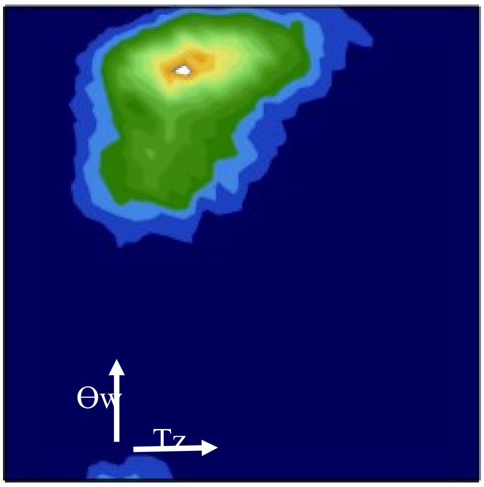

Figure 15. Histogram for positive correlation

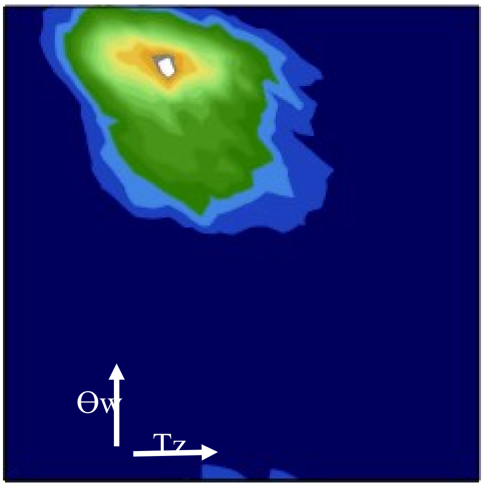

Figure 16. Histogram for negative correlation

2.2.4 Numerical Verification of the Joint Probability Model

Using Monte Carlo simulation the joint probability model described above was used to generate 21807 artificial data samples (the same sample size of the original data) of the metocean parameters. The statistics and some histograms of both samples are compared in order to verify if the joint probability model reproduces properly the main characteristics of the original data sample. Table 10 shows some statistical properties of the linear variables obtained from the arificially generated data sample. The results shown in this table should be compared with those of Table 1. Table 11 presents some statistics of circular data obtained from the simulated data sample. These figures should be compared with those presented in Table 4. Figures 17-19 show the joint histograms obtained using the original data sample and the data sample generated numerically using the Nataf-based joint probability distribution for wave (θ_W) and wind (θ_V) directions, and for significant wave height (H_S) and wave direction (θ_W). Other results and applications can be found in (MOSQUERA, 2015).

In general, it can be observed that the joint probability model developed in this work is able to represent the main characteristics of the original data sample. In other words, the dependence between all variables is reasonably well modeled. Then, the proposed model is suitable for representing the joint distribution of joint environmental parameters.

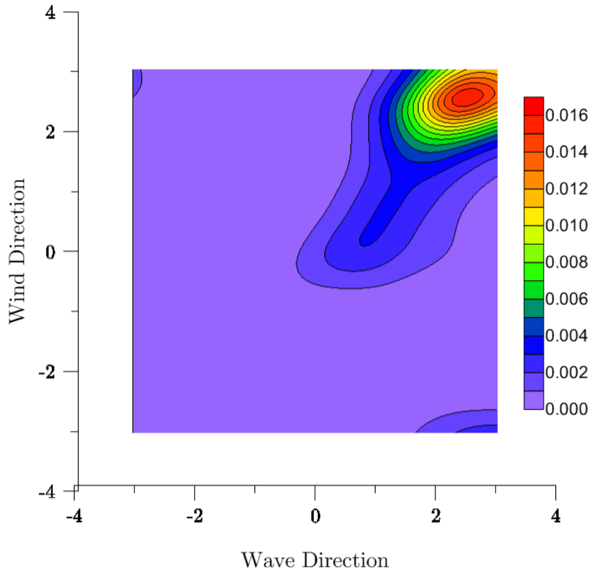

Figure 17. Joint probability distribution for wind direction and wave direction. Original data sample.

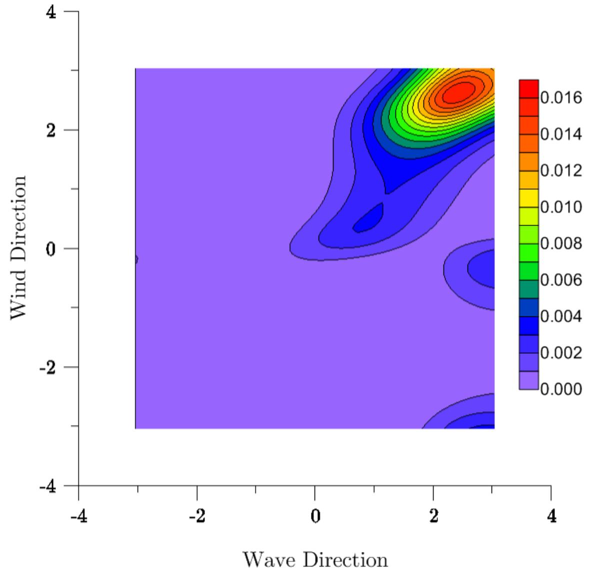

Figure 18. Joint probability distribution for wind direction and wave direction. Simulated data sample.

Table 10. Statistic parameters for artificial linear data

	Mean	Median	Standard deviation	Coef. of Variation	Skewness	Kurtosis
H_s	1.304	1.160	0.672	0.517	1.210	7.528
T_Z	4.744	4.680	0.789	0.166	0.435	3.355
V_V	6.601	6.441	2.781	0.423	0.300	2.823
V_C	0.205	0.180	0.135	0.657	1.117	4.216

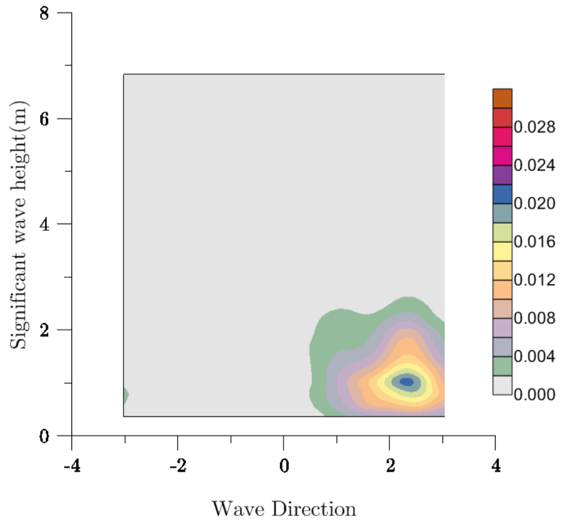

Figure 19. Joint probability distribution for significant wave height and wave direction. Original data sample.

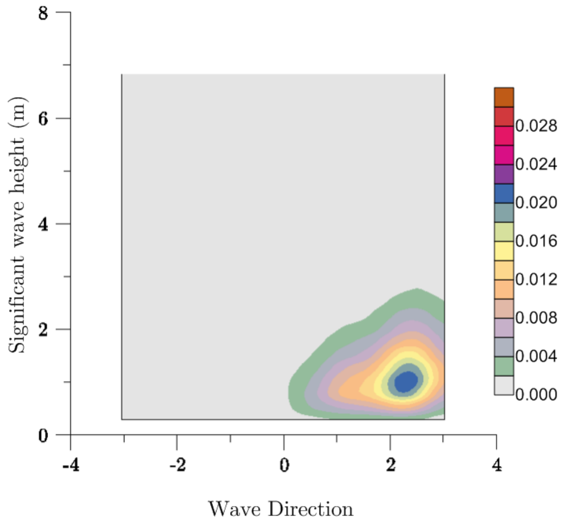

Figure 20. Joint probability distribution for significant wave height and wave direction. Original data sample.

Table 11. Statistic parameters for artificial circular data

Var.	Mean Direction	Mean resultant length	Standard deviation
θ_W	115.799	0.744	0.507
θ_V	121.895	0.546	0.725
θ_C	-7.013	0.309	1.010

3 FINAL REAMARKS

This paper presents some improvements for the Nataf-based model developed by PAPALEO (2009) for representing the joint probability distribution of metocean data where intensity and directions of the environmental parameters are considered. The first improvement was the use of an interpolated distribution in a finite interval $[0,2\pi]$ to represent the marginal distribution of circular data in the model. This aspect avoids the use of multimodal Wrapped Normal distributions and, as a consequence, the modelling of this kind of variable becomes straightforward. The second improvement is associated to the use of an auxiliary equation just to indicate the correct sign of the correlation coefficient between a linear and a circular variable. This step avoids the use of visual inspections to define the correct sign this coefficient and helps a lot for the case where it is small.

The model was tested for an available metocean data set measured at a location in Gulf of Mexico. Through numerical simulations it was observed that the joint probability model is able of representing the statistical behavior of the original data.

REFERENCES

ANG, A. H.-S.; TANG, W. H.,1975, *Probability Concepts in Engineering Planning and Design. Volume I, Basic Principles.* John Wiley & Sons, New York.

BITNER-GREGERSEN, E. et al., 1996."World-Wide Characteristics of Hs and Tz for Long-Term Load Responses of Ship and Structures". *International Offshore and Polar Engineering Conference*, Los Angeles, United States.

BITNER-GREGERSEN, E.; HAVER, S., 1991. "Joint Environmental Model for Reliability Calculations". *International Offshore and Polar Engineering Conference.* Edinburgh, Scotland.

BRODTKORB, P. A., 2005, *The Probability of Occurrence of Dangerous Waves Situations at the Sea*, PhD. Thesis, Departament of Marine Technology, Norwegian of Science and Technology, Trondheim, Norway.

FISHER, N. I., 1993. *Statistical Analysis of Circular Data*. Cambridge University Press, Cambridge.

FISHER, N. I.; LEE, A. J. A., 1983. *Correlation Coefficient for Circular Data*. Biometrika, Volume 70, Pp. 327-332.

HORNIK, K.; GRUN, B., 2014, *movMF: An R Package for Fitting Mixtures of von Mises-Fisher Distribuitions*. Journal of Statistical Software, Volume 40, No13, Pp. 1-31.

JAMMALAMADAKA, R.; RAMAKRISHANA, S., 1998, "A correlation Coefficient for Angular Variables". *Proceedings of the Second Pacific Area Statistical Conference*, North Holland, Netherlands.

JOHANNESSEN, K.; MELING, T. S.; HAVER, S., 2001, "Joint Distribution for Wind and Waves in the Northern North Sea". *Proceeding of the International Offshore and Polar Engineering Conference*, Stavanger, Norway.

JOHNSON, R. A.; WEHRLY, T., 1977, *Measures for Angular Correlation and Angular-Linear Correlation*. Journal of the Royal Statistical Society, Volume 39, No 2, Pp 222-229.

KIUREGHIAN, A. D.; LIU, P.-L., 1986, *Structural Reliability Under Incomplete Probability Information.* Journal Engineering Mechanics, Volume 112, Pp. 85-104.

LEBRUN, R.; DUTFOY, A., 2009, *An Innovating Analysis of the Nataf Transformation from the Copula Viewpoint*. Probabilistic Engineering Mechanics,Volume 24, Issue 3, Pp. 312-320.

MARDIA, K. V., 1972. *Statistics of Directional Data*, Academic Press Inc ,London.

MOSQUERA, I.A.M, 2015.Modelo de Função de Probabilidade Conjunta de Parâmetros Ambientais Usando a Transformada de Nataf. Tese de M. Sc., COPPE/UFRJ, Rio de Janeiro. Brasil.

NATAF, A., 1962, *Détermination des Distribution dont les Marges sont Données*. Comptes Rendus de l´Academie des Sciences, Volume 255, Pp. 42-43.

PAPALEO, A., 2009. *Metodologia para Definição de Casos de Carregamentos Ambientais Equivalentes para o Projeto de Risers em Catenária*. Tese de D. Sc., COPPE/UFRJ, Rio de Janeiro. Brasil.

SAGRILO, L. V.; PRATES DE LIMA, E.; PAPALEO, A., 2011, *A Joint Probability Model for Environmental Parameters.* Journal of Offshore Mechanics and Arctic Engineering, Volume 133, Issue 3.

SOUKISSIAN, T., 2014, *Probabilistic Modeling of Directional and Linear Characteristics of Wind and Sea States*. Ocean Engineering, Volume 91, Pp. 91-110.