



MODELAGEM PROBABILÍSTICA DE PARÂMETROS DE ONDA PARA
ANÁLISE DE ESTRUTURAS MARÍTIMAS

Plínio Bueno Andrade Silva

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Civil, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Civil.

Orientadores: Luís Volnei Sudati Sagrilo
Luiz Paulo de Freitas Assad

Rio de Janeiro
Setembro de 2018

MODELAGEM PROBABILÍSTICA DE PARÂMETROS DE ONDA PARA ANÁLISE DE
ESTRUTURAS MARÍTIMAS

Plinio Bueno Andrade Silva

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM
ENGENHARIA CIVIL.

Examinada por:

Prof. Luís Volnei Sudati Sagrilo, D.Sc.

Prof. Luiz Landau, D.Sc.

Dr. Valdir Innocentini, Ph.D

RIO DE JANEIRO – RJ, BRASIL
SETEMBRO DE 2018

Silva, Plinio Bueno Andrade

Modelagem Probabilística de Parâmetros de Onda para
Análise de Estruturas Marítimas/ Plinio Bueno Andrade Silva – Rio
Janeiro: UFRJ/COPPE, 2018.

XIII, 58 p.: il.: 29.7 cm.

Orientadores: Luís Volnei Sudati Sagrilo

Luiz Paulo de Freitas Assad

Dissertação (mestrado) – UFRJ/COPPE/Programa de
Engenharia Civil, 2018.

Referências Bibliográficas: p. 57-58.

1. Modelagem probabilística de ondas. 2. Distribuição
Conjunta de Probabilidade. 3. Modelo Condicional. 4.
Transformada de Nataf. 5. Distribuição de Weibull, Lognormal,
Gumbel, Extremos. I. Sagrilo, Luís Volnei Sudati *et al.* II.
Universidade Federal do Rio de Janeiro, COPPE, Programa de
Engenharia Civil. III. Título.

DEDICATÓRIA

A todos os professores, cientistas e pessoas que se dedicam a livrar o mundo da ignorância.

AGRADECIMENTOS

Especialmente ao meu orientador Prof. Luís Volnei Sudati Sagrilo, pelo incansável esforço e conhecimentos tão generosamente oferecidos a mim, sem os quais o presente trabalho jamais poderia ter sido concluído.

Ao meu orientador Prof. Luiz Paulo de Freitas Assad, equipes dos laboratórios LAMCE e LACEO, professores e funcionários do programa de pós-graduação da COPPE.

Aos meus pais, Carlos Alberto e Sandra Maria, pela dedicação extrema na minha formação como cidadão.

Aos meus irmãos, ao círculo de confiança Stefoni e amigos.

As empresas Gardline, Ambidados, suas equipes e gerentes, por proporcionar o tempo necessário para que eu pudesse realizar o presente trabalho.

Uma saudação especial para: Prof. José Antônio Fontes Santiago, Carlos Fonseca, Grupo Glider, Caio Fonteles, Felipe Santos, Roberto Freires, Fagner Magalhães, Tiago Torres, André Vitta, Raquel Toste, Vivian Azeredo, Leonardo Kuniyoshi, Eduardo D'Azevedo, Juliana Rangel.

A comunidade de desenvolvimento e usuários da linguagem Python.

E finalmente, a minha amada e musa inspiradora, Dr. Carina Stefoni Bock.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

MODELAGEM PROBABILÍSTICA DE PARÂMETROS DE ONDA PARA
ANÁLISE DE ESTRUTURAS MARÍTIMAS

Plínio Bueno Andrade Silva

Setembro/2018

Orientadores: Luís Volnei Sudati Sagrilo

Luiz Paulo de Freitas Assad

Programa: Engenharia Civil

O objetivo desta dissertação é analisar os modelos de distribuição conjunta baseados na Transformada de Nataf e no Modelo Condicional para representar os parâmetros oceanográficos de onda: altura significativa de onda (H_s) e período de pico (T_p). No estudo são utilizados os resultados numéricos gerados pelo modelo de ondas WAVEWATCH III para a região da Baía de Santos (costa brasileira) e para o Mar da Noruega (costa norte da Noruega) como dados de entrada para a criação dos modelos de distribuição conjunta. São analisadas algumas distribuições de probabilidades que melhor se ajustam a distribuições marginais de parâmetros de onda bem como o desempenho dos modelos de distribuição conjunta para representá-los. Para cada uma das localidades, as análises foram realizadas para dois conjuntos de dados: um composto por 11 anos de resultados numéricos (2006 até 2016) e outro com apenas 2 anos (2015 à 2016). Os resultados indicam que ambos os modelos se mostraram capazes de representar os dados de ondas nas regiões avaliadas, porém o modelo da Transformada de Nataf apresentou maior facilidade em ser utilizado. Entretanto, em determinadas situações e com os ajustes corretos, o modelo Condicional é capaz de representar melhor certos aspectos das características das ondas.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

PROBABILITY MODEL OF WAVE PARAMETERS FOR MARINE STRUCTURES
ANALYSIS

Plinio Bueno Andrade Silva

September/2018

Advisors: Luís Volnei Sudati Sagrilo
Luiz Paulo de Freitas Assad

Department: Civil Engineering

The aim of this dissertation is to analyze the joint distribution models based on the Nataf's Transformation and on the Conditional Model in order to represent wave parameters such as significant wave height and peak period. The study uses numerical results from the WAVEWATCH III wave model for the Bacia de Santos area (Brazilian coast) and for the Norway Sea (Norway coast) as input data to create joint distribution models. It analyses the probabilities distribution which best fit into marginal distribution wave parameters as well as the performance of joint distribution models to represent them. For each of the sites, two sets of data have been analyzed: one set studies 11 years of numerical results (2006 to 2016) and the other studies just 2 years (2015 to 2016). Findings show that both models are able to represent wave data for the considered sites, but that Nataf's Transformation is easier to use. In certain situations, however, and with the correct adaptations, the Conditional model can better represent certain aspects of wave characteristics.

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Revisão Bibliográfica.....	4
1.2	Objetivos	5
2	MODELAGEM PROBABILÍSTICA DE VARIÁVEIS ALEATÓRIAS	6
2.1	Função distribuição de probabilidade (PDF) e cumulativa de probabilidade (CDF)	6
2.1.1	Valores característicos de uma variável aleatória.....	7
2.1.2	Modelos de distribuição de probabilidades	8
2.1.3	Distribuição de probabilidades Normal ou Gaussiana	8
2.1.4	Distribuição de probabilidades Lognormal	9
2.1.5	Distribuição de probabilidades de Weibull	10
2.1.6	Distribuição de probabilidades de Gumbel ou <i>Tipo I</i>	11
2.2	Distribuição de valores extremos	12
2.3	Técnicas de ajuste de distribuição de probabilidades a dados observados ..	14
2.3.1	Funções densidade e cumulativa de probabilidades empíricas	14
2.3.2	Possíveis distribuições candidatas.....	15
2.3.3	Parâmetros das distribuições candidatas.....	15
2.3.4	Método dos Momentos	16
2.3.5	Método da Máxima Verossimilhança	17
2.3.6	Verificação do Ajuste.....	18
2.4	Modelo de distribuição conjunta ou “ <i>Joint Probability Model</i> ” (JPM).....	18
2.5	Modelagem por Distribuição Condicional (MDC)	20
2.6	Modelo Baseado na Transformada de Nataf	22
2.7	Métodos para análise dos modelos de distribuição conjunta	23
2.7.1	Diagrama de Taylor	25
2.7.2	Comparação Visual	26
3	AJUSTES DE DISTRIBUIÇÕES CONJUNTAS DE H_s E T_P	28

3.1	Fonte dos dados.....	28
3.2	Áreas de estudo	29
3.2.1	Campo de Libra - Bacia petrolífera de Santos (BR).....	29
3.2.2	Mar da Noruega.....	30
3.3	Detalhes sobre o ajuste da função de T_p condicionada a H_s	32
3.4	Sequência usada na modelagem probabilística de H_s e T_p	35
3.5	Resultados obtidos.....	37
3.5.1	Bacia de Santos – Costa Brasileira.....	37
3.5.2	Mar da Noruega.....	47
4	CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS.....	55
5	REFERÊNCIAS BIBLIOGRÁFICAS.....	57

LISTA DE FIGURAS

Figura 1- Registro da elevação da superfície do mar durante um estado de mar.....	2
Figura 2- Densidade espectral ou espectro do mar com o pico do espectro em destaque	2
Figura 3- Funções densidade e cumulativa de probabilidades	7
Figura 4- Diversos tipos de diagrama de dispersão e os correspondentes valores de correlação.....	20
Figura 5- Histograma bidimensional.....	24
Figura 6- Exemplo Diagrama de Taylor.....	26
Figura 7- Representação 3-D do modelo de distribuição conjunta de probabilidades .	26
Figura 8- Representação 2-D da distribuição conjunta: a) empírica e b) modelo	27
Figura 9- Curvas de diferenças entre distribuições conjuntas: empírica e modelada ..	27
Figura 10- Posicionamento do ponto utilizado no estudo (em vermelho). O polígono (linha cinza) delimita a área do Campo de Libra.	29
Figura 11- Posicionamento do ponto utilizado no estudo (em vermelho) no mar da Noruega, região Tromsøflaket.	31
Figura 12- Histograma de frequência relativa de ocorrências de $H_s - T_p$	33
Figura 13- PDFs de T_p para os intervalos de H_s	33
Figura 14- Valores discretos ξ e λ e suas funções ajustadas $\xi(h)$ e $\lambda(h)$	34
Figura 15- Santos 11 anos: Ajuste de CDF do tipo Lognormal e Weibull para série de H_s	38
Figura 16- Distribuição de extremos anuais ajustado pela PDF de Gumbel.....	39
Figura 17- Distribuição do H_s centenário a partir da Lognormal	39
Figura 18- Distribuição do H_s centenário a partir da Weibull	39

Figura 19- Ajuste de CDFs Lognormal e Weibull para T_p .Costa brasileira.	40
<i>Figura 20- Ajuste da CDF condicional de T_p: teste com Weibull e Lognormal. Costa brasileira.</i>	<i>40</i>
Figura 21- Distribuição conjunta pelo método da distribuição condicional – costa brasileira (11 anos de dados)	41
Figura 22- Distribuição conjunta pelo método da distribuição condicional – costa brasileira (11 anos de dados). Intervalo de H_s alterado.	42
Figura 23- Distribuição conjunta pela metodologia baseada no modelo de Nataf – costa brasileira (11 anos de dados)	43
Figura 24- Diagrama de Taylor modelos condicional e Nataf - costa brasileira (11 anos de dados).....	43
Figura 25- Distribuição conjunta pelo método da distribuição condicional – costa brasileira (2 anos de dados)	45
Figura 26- Distribuição conjunta pela metodologia baseada no modelo de Nataf – costa brasileira (2 anos de dados)	45
Figura 27- Diagrama de Taylor modelo condicional e Nataf - costa brasileira (2 anos de dados).....	46
Figura 28- Ajuste de CDFs Lognormal e Weibull para H_s .Costa da Noruega.....	48
Figura 29- Ajuste de CDFs Lognormal e Weibull para T_p .Costa da Noruega.....	48
Figura 30 - Ajuste da CDF condicional de T_p : teste com Weibull e Lognormal. Costa da Noruega.....	49
Figura 31- Distribuição conjunta pelo método da distribuição condicional – costa da Noruega (11 anos de dados)	49
Figura 32- Distribuição conjunta pelo método Nataf – costa da Noruega (11 anos de dados).....	50

Figura 33- Diagrama de Taylor modelo condicional e Nataf - costa da Noruega (11 anos de dados).....	51
Figura 34- Distribuição conjunta pelo método da distribuição condicional – costa da Noruega (2 anos de dados).....	52
Figura 35- Distribuição conjunta pelo método Nataf – costa da Noruega (2 anos de dados).....	53
Figura 36- Diagrama de Taylor modelo condicional e Nataf - costa da Noruega (2 anos de dados).....	54

LISTA DE TABELAS

Tabela I- Valores Característicos de uma Variável Aleatória X.....	8
Tabela II- Dados estatísticos - Campo de Libra	30
Tabela III- Dados estatísticos - Região Tromsøflaket.....	32
Tabela IV - Comparação entre os parâmetros estatísticos: dados brutos e modelos condicional e Nataf (Costa brasileira – 11 anos de dados)	44
Tabela V- Comparação entre os parâmetros estatísticos: dados brutos e modelos condicional e Nataf (Costa brasileira – 02 anos de dados)	46
Tabela VI- Comparação entre os parâmetros estatísticos: dados brutos e modelos condicional e Nataf (Costa da Noruega – 11 anos de dados).....	51
Tabela VII- Comparação entre os parâmetros estatísticos: dados brutos e modelos condicional e Nataf (Costa da Noruega – 2 anos de dados).....	53

1 Introdução

A exploração de recursos no ambiente marítimo por vezes necessita que estruturas sejam instaladas e fiquem expostas ao mar durante grandes períodos de tempo. O ambiente marinho é extremamente sensível às condições meteorológicas e a variação destas podem gerar em uma mesma área tanto ondas com baixos valores de alturas, como tormentas com ondas de enorme poder destrutivo, ou seja, grandes variações no estado de mar.

As estruturas marítimas estão submetidas à situações extremas de carregamento gerados pelos efeitos ambientais bem como à fadiga estrutural. A fadiga pode levar ao crescimento de defeitos indesejáveis na estrutura devido ao efeito cíclico dos carregamentos. Esses carregamentos de fadiga usualmente não têm capacidade para gerar um colapso instantâneo da estrutura, no entanto a ação contínua dos mesmos pode levar a essa situação.

Usualmente, as ondas marinhas são responsáveis pelos maiores danos de fadiga e por respostas extremas. As ondas do mar constituem-se de processos aleatórios e para fins de modelagem, assume-se que tais processos apresentem uma pseudo-estacionariedade durante determinados curtos períodos de tempo, denominados de estados de mar. O período de duração dos estados de mar é usualmente considerado na prática como 3-h. No levantamento de dados de ondas, em cada estado de mar elas são caracterizadas por três parâmetros: altura significativa de onda H_s , período de cruzamento zero T_z (ou o período de pico T_p) e uma direção principal de propagação θ_w , que por convenção da oceanografia refere-se sempre como a direção de onde a onda vem. Tomando-se como base a Figura 1, a altura significativa de onda H_s é definida como sendo a média do último terço das ondas individuais identificadas no registro ordenadas de forma crescente de altura, i.e.

$$H_s = \frac{3}{N} \sum_{i=2/3N}^N H'_i$$

onde N é o número total de ondas individuais e H'_i é a altura da i -ésima onda individual na amostra ordenada em valores crescentes de altura das ondas individuais.

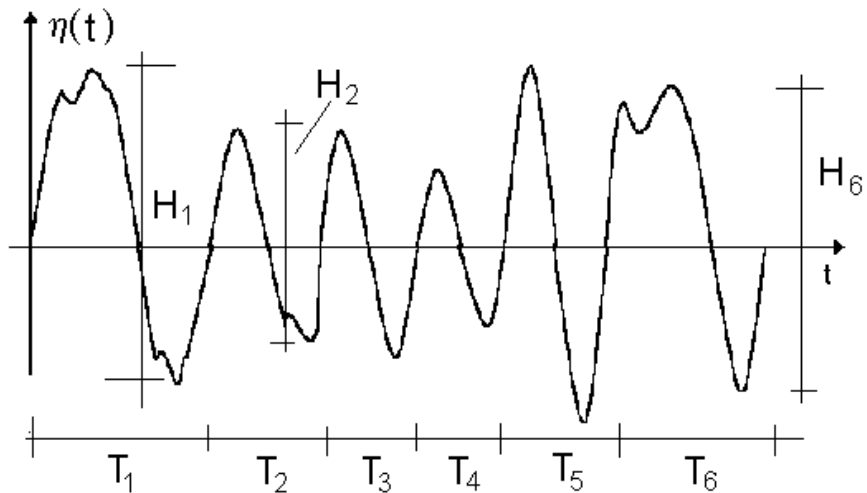


Figura 1- Registro da elevação da superfície do mar durante um estado de mar

Por sua vez, o período de cruzamento zero corresponde ao valor médio de todas as ondas individuais identificadas no registro, i.e.,

$$T_z = \frac{1}{N} \sum_{i=1}^N T_i$$

É comum também fazer uma análise de Fourier do sinal de elevações do mar e obter a densidade espectral do registro, como mostra a Figura 2. Caracteriza-se por período de pico T_p o período correspondente a frequência de pico do espectro ω_p , i.e.,

$$T_p = \frac{2\pi}{\omega_p}$$

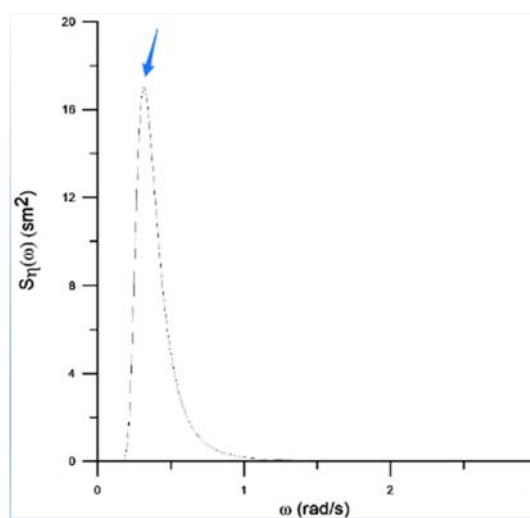


Figura 2- Densidade espectral ou espectro do mar com o pico do espectro em destaque

Existem na literatura relações matemáticas que relacionam T_p a T_z e vice-versa (HOLTHUIJSEN, 2007). Desta forma, na prática escolhe-se um ou outro para caracterizar um estado de mar. Neste trabalho será usado o T_p .

Através de uma série histórica de observações os dados de onda são agrupados por direção de incidência e, assim, os dados simultâneos de H_s e T_p para cada direção podem ser tratados estatisticamente.

Métodos modernos de análise de fadiga e de estimativa da resposta extrema de estruturas requerem a definição da distribuição conjunta de probabilidades dos parâmetros ambientais de onda para a locação onde a estrutura está ou será instalada. Por exemplo, o dano anual de fadiga é calculado por

$$D_{1yr} = 2920 \sum_{i=1}^{N_\theta} \left(\int_0^\infty \int_0^\infty d^i(h, t) f_{H_s, T_p}^i(h, t) dh dt \right) p_i \quad (1-1)$$

onde N_θ é o número de direções de incidência de onda, $d^i(h, t)$ é o dano à fadiga para uma estado de mar com $H_s = h$ e $T_p = t$ incidindo na i -ésima direção (obtido através de uma análise estrutural), $f_{H_s, T_p}^i(h, t)$ é a correspondente distribuição conjunta dos dois parâmetros de onda, p_i é a frequência relativa de ondas na i -ésima direção e 2920 é o número de estados de mar de 3-h durante o período de um ano (MONSALVE-GIRALDO *et al.* 2016). A vida útil a fadiga é calculada como

$$VU = \frac{1}{D_{1yr}} \quad (1-2)$$

O foco do presente trabalho concentra-se na investigação de procedimentos para obtenção da distribuição conjunta de probabilidades de H_s e T_p para futuras aplicações em métodos modernos de análise de estruturas marítimas, que buscam resolver de forma eficiente a equação do dano anual de fadiga.

1.1 Revisão Bibliográfica

A modelagem estatística de parâmetros ambientais vem recebendo atenção de vários pesquisadores ao redor do mundo. MATHISEN e BITNER-GREGERSEN (1990) utiliza dados de onda obtidos na região do Mar da Noruega e faz a comparação entre diferentes métodos para o ajuste de distribuições de probabilidade, tanto para a altura significativa (H_s) quanto para o período de cruzamento zero (T_z) de ondas. Neste estudo, para a locação investigada, verificou-se que T_z foi melhor representado por uma distribuição tipo Lognormal e a distribuição de Weibull de 3 parâmetros foi a mais apropriada para modelar H_s . Para representar a distribuição conjunta destes parâmetros foram comparados os seguintes modelos: *Bivariate Log-Normal Distribution*, *Fang and Ilogben's Distribution*, *Bivariate Weibull Distribution* e *Marginal Weibull and Conditional Log-Normal* (MDC). Foi identificado que o modelo condicional (MDC) é aquele que apresentou melhor performance e sugerido que ele se tornasse padrão de modelo para aquela região.

BITNER-GREGERSEN (1991) introduz outros parâmetros ambientais no modelo anterior (MATHISEN e BITNER-GREGERSEN, 1990), tais como: velocidade do vento, correntes e nível de maré. Com esse novo modelo, BITNER-GREGERSEN (1991) consegue descrever melhor a correlação entre todos esses parâmetros e melhorar a compreensão da importância relativa destes parâmetros no cálculo de fadiga.

SAGRILO *et al.* (2011) e SILVA-GONZÁLEZ *et al.* (2013) destacam a dificuldade do modelo proposto por MATHISEN e BITNER-GREGERSEN (1990) em gerar uma distribuição conjunta utilizando curtos períodos de dados históricos ou com amostras não uniformemente distribuídas ao longo do tempo. SAGRILO *et al.* (2011) apresentam um método baseado na Transformada de Nataf, que utiliza o valor da correlação entre os parâmetros e não uma função condicional para a criação da distribuição condicional.

BITNER-GREGERSEN (2015) utilizou dados de previsão pretérita (*hindcast*) do modelo de ondas da Oceanweather Inc. para gerar distribuições conjuntas baseadas no modelo condicional de probabilidades em 4 regiões: Nigéria, Mar do Norte, Austrália e Ilhas Shetland (pertencentes a Escócia). Os períodos de dados gerados, dependendo da região, foram de 10 a 30 anos, com períodos de estado de mar que foram de 1 hora para a região da Austrália e de 3 horas para as demais regiões.

É interessante observar que a escolha de um modelo para representar estatisticamente um conjunto de dados envolve uma série de observações de decisões. Neste sentido, o artigo de TAYLOR (2001) foi bastante útil no desenvolvimento do presente trabalho. TAYLOR (2001) apresenta um método de comparação de resultados entre modelos que consegue condensar em um único gráfico três parâmetros estatísticos: erro médio quadrático (RMSE), desvio padrão e correlação dos dados empíricos e modelados. O intuito desse gráfico é facilitar a comparação entre os resultados de uma coleção de modelos e o fenômeno modelado, ou então, acompanhar a melhoria do desempenho de um único modelo em função de mudanças em seu desenvolvimento.

1.2 Objetivos

Como já mencionado anteriormente, o foco do trabalho é a modelagem de distribuições conjuntas dos parâmetros ambientais de onda. Serão investigados dois modelos: um baseado na distribuição condicional (MDC) (MATHISEN e BITNER-GREGERSEN, 1990) e outro baseado na Transformada de Nataf.

Serão aplicados aos modelos os dados de uma base histórica de 11 anos gerados pelo modelo WAVEWATCH III. Serão analisadas duas localidades:

- a) Bacia de Santos (costa brasileira)
- b) Mar do Noruega (costa norte da Noruega)

A performance dos métodos será investigada com o conjunto completo de dados e com uma base reduzida para verificar a possível influência do tamanho da amostra de dados na modelagem das distribuições conjuntas destas duas regiões.

2 Modelagem Probabilística de Variáveis Aleatórias

Na literatura relativa a variáveis aleatórias existem modelos associados à variáveis contínuas e variáveis discretas. Neste capítulo somente serão abordados os conceitos e modelos relacionados a variáveis contínuas porque os parâmetros ambientais de ondas, foco deste trabalho, constituem-se intrinsecamente de variáveis contínuas no eixo dos números reais.

2.1 Função distribuição de probabilidade (PDF) e cumulativa de probabilidade (CDF)

A função distribuição de probabilidade ou *probability density function* (PDF), por vezes chamada de função de densidade de probabilidades, é uma função que descreve a densidade relativa de probabilidade de uma variável aleatória X associada a um valor específico x desta variável. Se X é uma variável aleatório contínua e $f_X(x)$ é a função densidade de probabilidades dessa variável aleatória, a probabilidade de ocorrência de valores desta variável aleatória num intervalo particular $[a \leq X \leq b]$ é dada por:

$$P[a \leq X \leq b] = \int_a^b f_X(x) dx \quad (2-1)$$

Já a função cumulativa de probabilidade ou *cumulative distribution function* (CDF) é a função matemática que indica a probabilidade da variável aleatória X assumir um valor menor ou igual a x , ou seja: $F_X(x) = P(X \leq x)$. A CDF e a PDF se relacionam da seguinte forma:

$$F_X(x) = \int_{-\infty}^x f_X(x) dx \quad (2-2)$$

A Figura 3 traz uma apresentação genérica destas duas funções mencionadas acima.

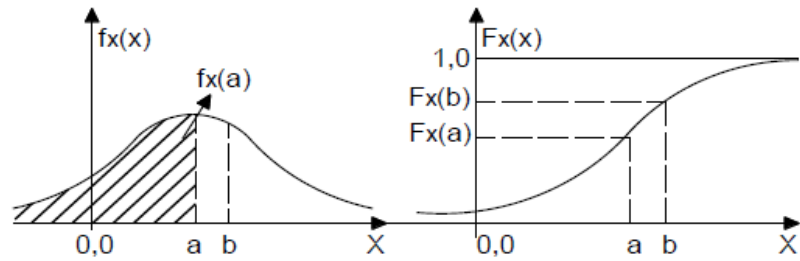


Figura 3- Funções densidade e cumulativa de probabilidades

2.1.1 Valores característicos de uma variável aleatória

Associados a uma variável aleatória existem diversos valores característicos que ajudam a descrevê-la, dentre os quais a média (μ_X), o desvio padrão (σ_X), o coeficiente de assimetria (γ_X^1) e a curtose (γ_X^2) sendo os principais (ANG e TANG, 1975). Além destes, outro valor característico importante é o coeficiente de variação que se constitui do valor do desvio padrão (σ_X) dividido pela média (μ_X).

Todos estes parâmetros podem ser estimados a partir da função densidade de probabilidades $f_X(x)$ da variável aleatória, como mostra a Tabela I. Entretanto, estimativas destes parâmetros podem ser obtidas a partir de uma amostra de tamanho finito N da variável em observação, isto é, $X = (x_1, x_2, \dots, x_N)$. Os valores amostrais dos parâmetros característicos também se encontram na Tabela I.

Outro valor característico muito usado na prática é o valor mais provável ou valor modal de uma distribuição de probabilidades. Este valor u corresponde ao valor de X onde a função densidade de probabilidades assume seu valor máximo, i.e.,

$$\left. \frac{df_X(x)}{dx} \right|_{x=u} = 0 \tag{2-3}$$

Tabela I- Valores Característicos de uma Variável Aleatória X

Parâmetro	Definição	Estimativa amostral
Média	$\mu_X = \int_{-\infty}^{+\infty} x f_X(x) dx$	$m_X = \sum_{i=1}^N \frac{x_i}{N}$
Valor médio quadrático	$\mu_{X^2} = \int_{-\infty}^{+\infty} x^2 f_X(x) dx$	$m_{X^2} = \sum_{i=1}^N \frac{(x_i)^2}{N}$
Desvio Padrão	$\sigma_X = \sqrt{\mu_{X^2} - (\mu_X)^2}$	$s_X = \sqrt{\sum_{i=1}^N \frac{(x_i - m_X)^2}{N}}$
Coefficiente de Assimetria	$\gamma_X^1 = \frac{\int_{-\infty}^{+\infty} (x - \mu_X)^3 f_X(x) dx}{(\sigma_X)^3}$	$s_X^1 = \sum_{i=1}^N \frac{(x_i - m_X)^3}{N(s_X)^3}$
Curtose	$\gamma_X^2 = \frac{\int_{-\infty}^{+\infty} (x - \mu_X)^4 f_X(x) dx}{(\sigma_X)^4}$	$s_X^2 = \sum_{i=1}^N \frac{(x_i - m_X)^4}{N(s_X)^4}$
Coefficiente de Variação (CoV)	$\delta_X = \frac{\sigma_X}{\mu_X}$	$d_X = \frac{s_X}{m_X}$

2.1.2 Modelos de distribuição de probabilidades

Na prática qualquer função matemática que tenha área unitária e não assuma valores negativos pode ser considerada, do ponto de vista teórico, como uma função densidade de probabilidades. Entretanto na literatura existem diversas funções propostas para tal finalidade. A seguir serão descritas algumas delas que são utilizadas ao longo do trabalho.

2.1.3 Distribuição de probabilidades Normal ou Gaussiana

Uma variável X é do tipo Normal quando sua função densidade de probabilidades pode ser descrita pela seguinte função

$$f_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - \mu_X}{\sigma_X}\right)^2\right) \quad -\infty \leq x \leq +\infty$$

(2-4)

sendo que os parâmetros desta distribuição são a média μ_X e o desvio padrão σ_X . A função cumulativa de probabilidades não tem uma função analítica explícita e é definida como:

$$F_X(x) = \Phi\left(\frac{x - \mu_X}{\sigma_X}\right)$$

(2-5)

sendo $\Phi(\cdot)$ a função de distribuição acumulada (CDF) da distribuição do Normal padrão (média 0 e desvio padrão 1). Os valores são tabelados em livros de estatística ou calculada automaticamente em programas de cálculo matemático.

Em função desta distribuição incluir no seu domínio valores negativos e os parâmetros de onda serem sempre positivos, ela praticamente não é usada em ajustes de distribuições para tais parâmetros.

2.1.4 Distribuição de probabilidades Lognormal

A distribuição de probabilidade de uma variável X é do tipo Lognormal quando o $\ln(X)$ possui uma distribuição do tipo Normal. A PDF do tipo Lognormal é dada pela seguinte função:

$$f_X(x) = \frac{1}{\xi x \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln(x) - \lambda}{\xi}\right)^2\right) \quad 0 \leq x \leq +\infty$$

(2-6)

onde λ e ξ representam os parâmetros da distribuição. A CDF é dada pela seguinte função:

$$F_X(x) = \Phi\left(\frac{\ln(x) - \lambda}{\xi}\right)$$

(2-7)

A média e o desvio padrão da variável X se relacionam com parâmetros λ e ξ através das seguintes equações:

$$\xi = \sqrt{\left(\log\left(1 + \left(\frac{\sigma}{\mu}\right)^2\right)\right)}$$

(2-8)

$$\lambda = \log(\mu) - \frac{1}{2}\xi^2$$

(2-9)

2.1.5 Distribuição de probabilidades de Weibull

A função de densidade de probabilidades (PDF) da distribuição de Weibull com 3 parâmetros é dada pela seguinte equação:

$$f_X(x) = \frac{\kappa}{\lambda} \left(\frac{x - \theta}{\lambda}\right)^{\kappa-1} e^{-\left(\frac{x-\theta}{\lambda}\right)^\kappa}$$

(2-10)

onde os parâmetros κ , θ e λ são respectivamente os parâmetros de *forma*, *locação* e *escala* da distribuição. A função cumulativa de probabilidades é dada por

$$F_X(x) = 1 - e^{-\left(\frac{x-\theta}{\lambda}\right)^\kappa}$$

(2-11)

No caso de $\theta = 0$, a distribuição toma o formato da distribuição de Weibull de 2 parâmetros.

Os parâmetros da distribuição de Weibull se relacionam com a média, desvio padrão e coeficiente de assimetria da variável X através das seguintes equações:

$$\begin{aligned}\mu_X &= \theta + \kappa \Gamma\left(1 + \frac{1}{\lambda}\right) \\ \sigma_X &= \kappa \sqrt{\Gamma\left(1 + \frac{2}{\lambda}\right) - \Gamma^2\left(1 + \frac{1}{\lambda}\right)} \\ \gamma_X^1 &= \frac{\Gamma\left(1 + \frac{3}{\lambda}\right) - 3\Gamma\left(1 + \frac{1}{\lambda}\right)\Gamma\left(1 + \frac{2}{\lambda}\right) + 2\Gamma^3\left(1 + \frac{1}{\lambda}\right)}{\left[\Gamma\left(1 + \frac{2}{\lambda}\right) - \Gamma^2\left(1 + \frac{1}{\lambda}\right)\right]^{\frac{3}{2}}}\end{aligned}\tag{2-12}$$

onde $\Gamma(\cdot)$ significa a função Gamma

2.1.6 Distribuição de probabilidades de Gumbel ou *Tipo I*

A função densidade de probabilidades $f_X(x)$ da distribuição de Gumbel (ou distribuição Tipo I) é dada por

$$f_X(x) = \alpha e^{-\alpha(x-u)-e^{-\alpha(x-u)}} \quad -\infty \leq x \leq +\infty\tag{2-13}$$

onde α é o parâmetro de escala da distribuição e u é o parâmetro de locação. A função cumulativa de probabilidades é dada por

$$F_X(x) = e^{-e^{-\alpha(x-u)}}\tag{2-14}$$

Os parâmetros α e u se relacionam com a média e o desvio padrão da variável aleatória X através das seguintes equações:

$$\mu_X = u + \frac{0.5772}{\alpha}\tag{2-15}$$

$$\sigma_X = \frac{\pi}{\sqrt{6}\alpha}\tag{2-16}$$

Adicionalmente, é possível também demonstrar que o parâmetro u corresponde ao valor mais provável desta distribuição. Esta distribuição tem uma importância muito grande na análise de valores extremos de uma análise aleatória.

2.2 Distribuição de valores extremos

Muitos projetos de engenharia devem ser baseados em valores extremos de uma dada variável aleatória X associados a um número N de observações da mesma, que no caso de parâmetros ambientais pode ser relacionado a um dado período de tempo ou período de retorno. No caso de existirem dados extremos observados, por exemplo, uma amostra com os valores máximos de onda numa dada locação, a distribuição de extremos seria a distribuição de probabilidades que se ajusta a esta amostra de dados. Entretanto, esta situação nem sempre ocorre, i.e., este tipo específico de dado não está disponível. Nesta situação é possível obter a distribuição de valores extremos através da Estatística de Ordem (ANG e TANG, 1984). Como será ilustrado a seguir, esta teoria baseia-se na distribuição ordinária da variável $F_X(x)$ (e sua correspondente $f_X(x)$), chamada aqui de distribuição inicial.

Nesta teoria emprega-se a lógica que, numa amostra de tamanho N da variável X , cada valor antes de sua observação constitui-se na verdade de uma variável aleatória, i.e,

$$X = \{X_1, X_2, X_3, \dots, X_i\} \quad (2-17)$$

$$\text{e } F_{X_1}(x) = F_{X_2}(x) = F_{X_3}(x) \dots F_{X_N}(x) = F_X(x) \quad (2-18)$$

onde X_i corresponde ao i -ésimo valor observado. Como cada valor observado pertence a população de X , é possível identificar que $F_{X_i}(x) = F_X(x)$. Supondo que um determinado valor $X = x$ pertença à população dos valores de X_e , este valor deve ser no mínimo maior ou igual a todos os valores da amostra X . A probabilidade deste evento é descrita por

$$P(X_e \leq x) = P(X_1 \leq x; X_2 \leq x; X_3 \leq x \dots X_N \leq x) \quad (2-19)$$

Assumindo-se que as observações na amostra X sejam estatisticamente independentes, a equação (2-19) pode ser escrita como

$$P(X_e \leq x) = P(X_1 \leq x)P(X_2 \leq x)P(X_3 \leq x) \dots P(X_N \leq x) \quad (2-20)$$

ou

$$F_{X_e}(x) = F_{X_1}(x)F_{X_2}(x)F_{X_3}(x) \dots F_{X_N}(x) \quad (2-21)$$

que resulta em

$$F_{X_e}(x) = F_X(x)^N \quad (2-22)$$

onde $F_{X_e}(x)$ é a CDF do valor extremo em N vezes que o evento se repete. Por exemplo, no caso da altura significativa de ondas H_s , se as medições deste parâmetro são realizadas a cada 3-h em uma direção de incidência θ_{w_j} , $N = 2920p_j$ para a distribuição do valor extremo anual e $N = 29200p_j$ para o valor centenário, sendo p_j a frequência relativa de ondas na direção considerada. A correspondente função densidade de probabilidades do valor extremo é dada por:

$$f_{X_e}(x) = NF_X(x)^{N-1}f_X(x) \quad (2-23)$$

Observa-se que através desta formulação é possível obter a distribuição do valor extremo para qualquer período de retorno com base unicamente no parâmetro e na distribuição inicial da variável, que é obtida com a base de dados disponível de dados de X .

A distribuição do valor extremo apresenta, para $N \rightarrow \infty$, um comportamento assintótico (ANG e TANG, 1984). Para a maioria das distribuições de interesse prático, $F_X(x)^N$ converge para uma distribuição de Gumbel cujos parâmetros podem ser definidos em função dos parâmetros de $F_X(x)$ e de N . Por exemplo, caso $F_X(x)$ seja descrita por uma distribuição de Weibull os parâmetros de escala e locação da Tipo I são dados por:

$$u = \theta + \lambda(\ln(N))^{\frac{1}{\kappa}} \quad (2-24)$$

$$\alpha = \frac{\kappa}{\lambda}(\ln(N))^{\frac{\kappa-1}{\kappa}} \quad (2-25)$$

2.3 Técnicas de ajuste de distribuição de probabilidades a dados observados

Na prática, a partir de uma amostra de valores observados, busca-se representar uma variável aleatória por um modelo de distribuição de probabilidade como, por exemplo, um dos descritos anteriormente neste trabalho. Este procedimento é denominado de “ajuste de uma distribuição” e consiste basicamente nas seguintes etapas:

- a) Obtenção dos valores característicos da amostra, tais como: média, desvio padrão, etc. (Tabela I).
- b) Obtenção da função densidade de probabilidades (ou da função cumulativa) empírica dos dados;
- c) Escolha das possíveis distribuições candidatas para modelar a variável em estudo;
- d) Obtenção dos parâmetros das distribuições candidatas e;
- e) Verificação do ajuste.

2.3.1 Funções densidade e cumulativa de probabilidades empíricas

A função empírica da função densidade de probabilidades pode ser obtida a partir de um histograma de frequência relativas dos dados da amostra. O valor aproximado da PDF no ponto médio do i -ésimo intervalo do histograma é dada por

$$f_X(x_i) \approx \frac{n_i}{N\Delta x} \quad (2-26)$$

onde n_i é número de valores da amostra dentro i-ésimo intervalo do histograma, N é o número de valores de X contidos na amostra observada e Δx é o tamanho dos intervalos do histograma.

A função cumulativa de probabilidades empírica é estabelecida a partir do ordenamento em ordem crescente dos valores da amostra. Para o i-ésimo valor de X na amostra “ordenada”, i.e., x'_i , a correspondente função cumulativa empírica de probabilidades é estimada aproximadamente como:

$$F_X(x'_i) \approx \frac{i}{N + 1} \tag{2-27}$$

2.3.2 Possíveis distribuições candidatas

Dentre os vários modelos teóricos existentes, alguns deles podem apresentar características que intrinsicamente não se aplicam à variável aleatória em questão. Por exemplo, uma variável aleatória que assume valores negativos e positivos não pode ser representada por distribuição Lognormal porque essa só representa dados que sejam sempre positivos. Este tipo de informação assim como, por exemplo, escolher como candidata a distribuição de Gumbel para representar uma amostra de valores extremos, servem para limitar o conjunto de modelos que serão testados no ajuste de dados.

2.3.3 Parâmetros das distribuições candidatas

Para cada uma das distribuições de probabilidade escolhidas como candidatas a representarem a amostra em análise existe a necessidade de se calcular os seus respectivos parâmetros. Por exemplo, se a distribuição candidata for uma Gumbel, então, devem ser definidos os correspondentes parâmetros α e u , conforme item 2.1.6.

Existem várias maneiras de se estimar tais parâmetros a partir de uma amostra de dados, dentre as quais o método dos momentos e método da máxima verossimilhança empregados neste trabalho e que serão descritos a seguir.

2.3.4 Método dos Momentos

O método dos momentos usa as estimativas de valores característicos amostrais da variável aleatória X , obtidos conforme descrito na Tabela I. Estes valores são representativos da amostra em questão e, portanto, podem não representar a população completa da variável X , exceto no caso em que a amostra seja suficientemente grande. Em outras palavras os parâmetros definidos anteriormente são apenas uma aproximação dos parâmetros reais da variável aleatória X .

No método dos momentos assume-se que os valores característicos da amostra da variável aleatória sejam iguais ao da sua população, i.e.,

$$E(X) = \mu_x \approx m_x \tag{2-28}$$

$$\text{Var}(X) = \sigma_x^2 \approx s_x^2 \tag{2-29}$$

e assim sucessivamente.

Como estas grandezas estão diretamente relacionadas aos parâmetros das distribuições de probabilidades, estes últimos podem ser facilmente obtidos. Por exemplo, para uma distribuição Normal os parâmetros μ_x e σ_x correspondem exatamente à média e o desvio padrão da variável e no caso de uma distribuição de Gumbel os parâmetros u e α são dados por

$$\alpha = \frac{\pi}{\sqrt{6}s_x} \tag{2-30}$$

$$u = m_x - \frac{0.5772}{\alpha} = m_x - \frac{0.5772\sqrt{6}s_x}{\pi} \tag{2-31}$$

O número de momentos (valores característicos da amostra) necessários corresponde ao número de parâmetros da distribuição.

2.3.5 Método da Máxima Verossimilhança

Supondo que seja necessário obter o parâmetro θ de uma distribuição cuja PDF é definida por $f_X(x, \theta)$ para verificar se a mesma se ajusta ou não à amostra observada $X = (x_1, x_2, \dots, x_n)$ (note que, por enquanto, a distribuição tem um só parâmetro, i.e., θ). Baseado nesta amostra, a seguinte pergunta pode ser feita: Qual o valor mais provável de θ que produz o conjunto de observações x_1, x_2, \dots, x_n ? Ou em outras palavras, qual é o valor de θ que maximiza a possibilidade de obter a sequência x_1, x_2, \dots, x_n ? A possibilidade de se obter tal sequência pode ser assumida como sendo proporcional ao valor da função densidade de probabilidades (PDF) calculada em cada x_i . Assim, uma função de “probabilidades” pode ser definida como

$$P(x_1, x_2, \dots, x_N, \theta) = f_X(x_1, \theta) f_X(x_2, \theta) \dots f_X(x_N, \theta) \quad (2-32)$$

O valor de θ que maximiza esta função pode ser obtido resolvendo-se a seguinte expressão

$$\frac{\partial P(x_1, x_2, \dots, x_N, \theta)}{\partial \theta} = 0 \quad (2-33)$$

Para distribuições dependentes de mais de um parâmetro, a função de probabilidades torna-se

$$P(x_1, x_2, \dots, x_n, \theta_1, \dots, \theta_m) = f_X(x_1, \theta_1, \dots, \theta_m) f_X(x_2, \theta_1, \dots, \theta_m) \dots f_X(x_N, \theta_1, \dots, \theta_m) \quad (2-34)$$

onde m é número de parâmetros da distribuição e os mesmos são obtidos através da solução do seguinte sistema de equações:

$$\frac{\partial P(x_1, x_2, \dots, x_N, \theta_1, \dots, \theta_m)}{\partial \theta_1} = 0 \quad (2-35)$$

$$\vdots$$

$$\frac{\partial P(x_1, x_2, \dots, x_N, \theta_1, \dots, \theta_m)}{\partial \theta_m} = 0 \quad (2-36)$$

Conforme descrito na literatura (ANG e TANG, 1975), dependendo da situação, torna-se conveniente operar com o $\log P(x_1, \dots, x_N, \dots, \theta_1, \dots, \theta_m)$.

2.3.6 Verificação do Ajuste

A escolha da distribuição é feita usualmente por meio de testes de aderência e/ou comparações visuais. Dentre as distribuições candidatas escolhe-se para modelar a variável aleatória em análise aquela em que a sua PDF (ou CDF) apresenta uma melhor aderência a PDF (ou CDF) empírica.

2.4 Modelo de distribuição conjunta ou “*Joint Probability Model*” (JPM)

Em muitos casos práticos existe a necessidade de se modelar o comportamento randômico simultâneo (ou conjunto) de duas ou mais variáveis aleatórias, como, por exemplo, altura significativa de onda H_s e o período de cruzamento zero T_z de vários estados de mar. Neste caso recorre-se a distribuição conjunta de probabilidades. No caso de duas variáveis aleatórias X e Y , a função de densidade de probabilidades conjunta $f_{X,Y}(x, y)$ tem a seguinte interpretação:

$$P(x \leq X \leq x + dx \cap y \leq Y \leq y + dy) = f_{X,Y}(x, y) dx dy \quad (2-37)$$

ou seja, a probabilidade da variável aleatória X assumir um valor no intervalo $a \leq X \leq b$ e simultaneamente a variável Y assumir um valor no intervalo $c \leq Y \leq d$ é dada por

$$P(a \leq X \leq b \cap c \leq Y \leq d) = \int_c^d \int_a^b f_{X,Y}(x, y) dx dy \quad (2-38)$$

Usualmente uma distribuição conjunta de probabilidades é definida utilizando o conceito de distribuição condicional (ANG e TANG, 1975). Neste caso

$$f_{X,Y}(x, y) = f_X(x) f_{Y|X}(y|x) \quad (2-39)$$

onde $f_X(x)$ é a função densidade de probabilidades marginal da variável X e $f_{Y|X}(y|x)$ é a função densidade de probabilidades da variável aleatória Y condicionada a valores da variável aleatória X . Um aspecto importante no estudo do comportamento simultâneo de duas (ou mais) variáveis aleatórias é o conceito de independência estatística. Duas variáveis são ditas estatisticamente independentes se uma não tem nenhuma influência (ou correspondência) sobre a outra e vice-versa. Neste caso a distribuição condicional de Y é a mesma para quaisquer valores de X , i.e., $f_{Y|X}(y|x) = f_Y(y)$ e conseqüentemente:

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad (2-40)$$

Outra grandeza relacionada à dependência (ou não) entre duas variáveis aleatórias X e Y é o coeficiente de correlação, dado por:

$$\rho = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} \quad (2-41)$$

onde σ_X e σ_Y são os correspondentes desvios padrões das variáveis aleatórias X e Y e $Cov(X,Y)$ é a covariância entre as duas variáveis, que é definida por:

$$Cov(X,Y) = E(XY) - E(X)E(Y) = \iint_{-\infty}^{+\infty} xyf_{X,Y}(x,y) - \mu_X\mu_Y \quad (2-42)$$

No caso de um conjunto de dados com N valores simultâneos (x_i, y_i) , o coeficiente de correlação é estimado por:

$$r = \frac{\sum_{i=1}^N \frac{x_i y_i}{N} - m_x m_y}{s_x s_y} \quad (2-43)$$

onde m_x , m_y , s_x e s_y são definidos na Tabela I.

O coeficiente de correlação varia entre -1 e +1 (ANG e TANG, 1975). Os casos extremos (-1 e +1) significam uma dependência completa entre as variáveis, ou seja, conhecendo-se o comportamento de uma delas também se define o comportamento da

outra. No caso de variáveis estatisticamente independentes o coeficiente de correlação é nulo, porém, a recíproca não é verdadeira. O coeficiente de correlação mede uma dependência linear entre as variáveis, assim um valor nulo de correlação não exclui a possibilidade de haver outro tipo de dependência (não linear) entre as variáveis. A Figura 4 ilustra, genericamente, o comportamento do coeficiente de correlação para diversos tipos de dados.

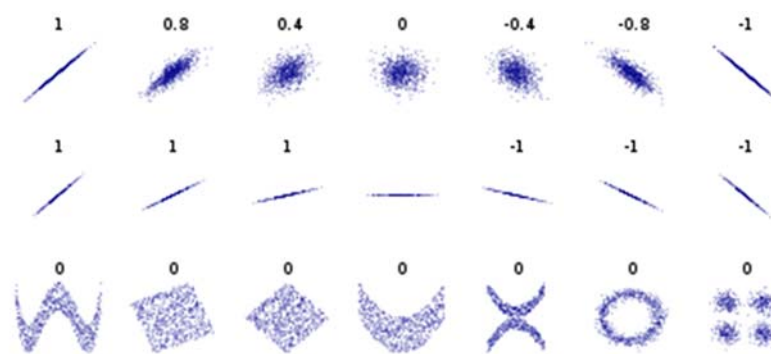


Figura 4- Diversos tipos de diagrama de dispersão e os correspondentes valores de correlação

A obtenção de uma distribuição conjunta de probabilidades baseada no conceito de distribuição condicional não é tarefa simples. Uma modelagem alternativa para a distribuição conjunta é o modelo baseado na Transformada de Nataf (NATAF, 1962). Este último modelo utiliza apenas as distribuições marginais de probabilidades das variáveis e seus correspondentes coeficientes de correlação. Como o objetivo deste trabalho é o de obter um modelo de distribuição conjunta (JPM) para os parâmetros de onda H_s e T_z , ambos os modelos serão investigados.

2.5 Modelagem por Distribuição Condicional (MDC)

O MDC é um método que utiliza uma função condicional para definir a distribuição de probabilidade de uma variável em função de uma outra variável. Em geral, os parâmetros da distribuição condicional são modelados (parametrizados) por uma função matemática dependente do valor assumido pela outra variável. Na literatura (ANG e TANG, 1975), um dos poucos casos em que esta parametrização da distribuição condicional é conhecida analiticamente é quando, por exemplo, duas variáveis aleatórias X e Y são normais correlacionadas através de um coeficiente de correlação ρ . Neste caso a distribuição conjunta é dada por

$$f_{X,Y}(x,y) = f_X(x)f_{Y|X}(y|x) \quad (2-44)$$

onde

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - \mu_X}{\sigma_X}\right)^2\right) \quad (2-45)$$

$$f_{Y|X}(y|x) = \frac{1}{\sigma_Y\sqrt{1-\rho^2}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - \mu_Y(x)}{\sigma_Y\sqrt{1-\rho^2}}\right)^2\right) \quad (2-46)$$

sendo

$$\mu_Y(x) = \mu_Y + \rho(\sigma_Y/\sigma_X)(x - \mu_X) \quad (2-47)$$

onde μ_X e μ_Y são as médias individuais das duas variáveis e σ_X e σ_Y são correspondentes desvios padrões. Observe que neste caso quando $\rho = 0$ as variáveis são estatisticamente independentes.

Em geral, a descrição deve ser por meio do ajuste de alguma função matemática em função da amostra dos dados simultâneos das duas variáveis. Por exemplo, supondo que a distribuição condicional de Y seja uma Lognormal, esta pode ser escrita como

$$f_{Y|X}(y|x) = \frac{1}{\xi(x)y\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln(y) - \lambda(x)}{\xi(x)}\right)^2\right) \quad (2-48)$$

Neste modelo acima, o desafio principal é obter as funções analíticas dos parâmetros $\lambda(x)$ e $\xi(x)$ que representem adequadamente os dados analisados. Este tópico será abordado novamente mais adiante neste trabalho na modelagem dos parâmetros de onda H_s e T_z .

2.6 Modelo Baseado na Transformada de Nataf

Após algumas manipulações algébricas, a distribuição Normal conjunta de probabilidades descrita no item anterior pode ser também escrita da seguinte forma (ANG e TANG, 1975):

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right\}\right) \quad (2-49)$$

No caso específico das médias serem 0 e os desvios padrões iguais a 1, ou seja, X e Y são variáveis Normais padrão, este modelo é representado por

$$\phi(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\{x^2 + y^2 - 2\rho xy\}\right) \quad (2-50)$$

Outro conceito bastante importante em probabilidade e estatística é o de equivalência estatística. A equivalência estatística entre duas variáveis aleatórias X e Y é dada por:

$$F_X(x) = F_Y(y) \quad (2-51)$$

ou seja, o valor específico x de uma variável aleatória X é estatisticamente equivalente a um valor y de uma variável aleatória Y quando os valores de suas funções cumulativas, $F_X(x)$ e $F_Y(y)$, respectivamente, são iguais.

Utilizando o conceito de equivalência estatística, onde as variáveis aleatórias X e Y são individualmente transformadas para Normais padrão equivalentes, NATAF (1962) demonstrou que a distribuição de probabilidades conjunta destas duas variáveis aleatórias é representada por:

$$f_{X,Y}(x,y) = \frac{f_X(x)f_Y(y)}{\phi(\Phi^{-1}[F_X(x)])\phi(\Phi^{-1}[F_Y(y)])} \phi_2(\Phi^{-1}[F_X(x)], \Phi^{-1}[F_Y(y)], \rho^N) \quad (2-52)$$

onde $\Phi^{-1}[\cdot]$ e $\phi(\cdot)$ correspondem a inversa da função cumulativa de probabilidades e a função densidade de probabilidades, respectivamente, da distribuição Normal padrão, $\phi_2(\cdot, \cdot, \rho^N)$ é função densidade de probabilidades da distribuição Normal conjunta para 2 variáveis Normais padrão cuja correlação entre elas é ρ^N (veja equação (2-50), $f_X(x)$ e $f_Y(y)$ e $F_{X_i}(x_i)$ correspondem às funções densidade de probabilidades das variáveis X e Y , respectivamente, e $F_X(x)$ e $F_Y(y)$ são as correspondentes funções cumulativas de probabilidade marginais da variável X_i , respectivamente. ρ^N é chamado de correlação equivalente de Nataf e é dado por

$$\rho^N = \iint_{-\infty}^{+\infty} \left(\frac{F_X^{-1}(\Phi(z_1)) - \mu_X}{\sigma_X} \right) \left(\frac{F_Y^{-1}(\Phi(z_2)) - \mu_Y}{\sigma_Y} \right) \phi_2(z_1, z_2, \rho) dz_1 dz_2 \quad (2-53)$$

onde $F_X^{-1}(\cdot)$ e $F_Y^{-1}(\cdot)$ correspondem às inversas da funções cumulativas de probabilidades das variáveis X e Y , respectivamente, μ_X e μ_Y são as médias destas variáveis, σ_X e σ_Y os correspondentes desvios padrões, e ρ é o coeficiente de correlação linear (original) entre as variáveis X e Y . A solução da equação (2-53) pode ser feita numericamente, porém, DER KIUREGHIAN e LIU (1986) apresentam soluções aproximadas para substituir a avaliação numérica desta integral dupla. Porém, em muitos casos práticos é possível assumir $\rho \approx \rho^N$. Esta última hipótese foi adotada neste trabalho.

O modelo de Nataf necessita, para ser definido, apenas das distribuições marginais de probabilidades das variáveis envolvidas e do coeficiente de correlação entre elas. Não há necessidade de parametrizar o comportamento do parâmetro de uma variável em função da outra. Por outro lado, deve se observar que existem situações em que a correlação não consegue expressar completamente a dependência entre duas variáveis aleatórias.

2.7 Métodos para análise dos modelos de distribuição conjunta

A partir de uma amostra simultânea de dados de duas variáveis aleatórias podem ser facilmente obtidas e/ou calculadas, entre outras, as seguintes informações:

- a) Parâmetros estatísticos individuais (média, desvio padrão, etc.) de cada uma das variáveis;

- b) Distribuições de probabilidade marginais de cada uma das variáveis;
- c) Coeficiente de correlação entre as variáveis e;
- d) Função densidade de probabilidade conjunta empírica dos dados.

A função densidade de probabilidade conjunta empírica dos dados pode ser obtida a partir de um histograma bidimensional (no caso de ondas é conhecido como diagrama de dispersão) é definida como:

$$f_{x,y}(x_i, y_j) = \frac{n_{i,j}}{N\Delta x\Delta y} \tag{2-54}$$

onde (x_i, y_j) é o ponto central da célula do histograma considerada (Figura 5), $n_{i,j}$ é o número de observações de pares simultâneos da amostra que pertencem à célula considerada, N é o tamanho total da amostra e Δx e Δy são os correspondentes tamanhos dos intervalos assumidos para montar o diagrama de dispersão bidimensional.

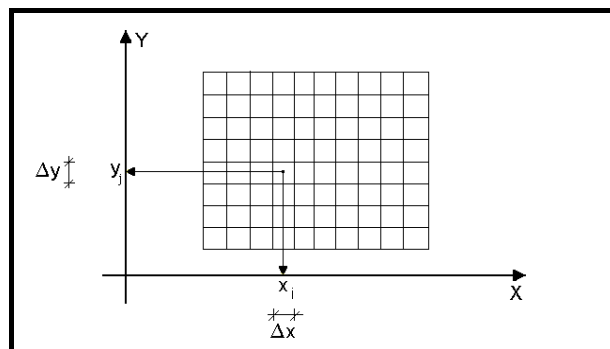


Figura 5- Histograma bidimensional

A distribuição de probabilidades conjunta é o modelo matemático que melhor representa ou se ajusta à distribuição empírica. Esta escolha pode ser feita comparando-se as diferenças entre a função teórica e a função empírica em todas as células do modelo. Além disto, é interessante que o modelo escolhido reproduza razoavelmente os parâmetros característicos das variáveis envolvidas. Neste trabalho foi usado o Diagrama de Taylor (que compara as diferenças entre o modelo numérico e o empírico, descrito a seguir), a comparação visual e a observação das diferenças entre os valores característicos da amostra e do modelo numérico para a escolha do modelo final de distribuição conjunta representativo para a os dados analisada.

2.7.1 Diagrama de Taylor

O diagrama de Taylor é uma forma de comparar o resultado de diferentes modelos numéricos aplicados a uma mesma série de dados. No presente trabalho, a amostra de valores observados são os valores da função densidade de probabilidade empírica dos dados em cada uma das células do histograma bidimensional e os valores “calculados” são os correspondentes valores obtidos pelo modelo numérico no ponto central de cada uma destas células.

Cada ponto plotado no gráfico representa três valores estatísticos da série de resultados de cada modelo. Estes valores são: o desvio padrão dos resultados numéricos, a correlação e a raiz do erro médio quadrático (RMSE) entre o resultado numérico do modelo e os dados empíricos. No eixo das abscissas plota-se o desvio padrão dos dados brutos e depois plota-se um ponto de coordenadas (r, θ) sendo r igual ao desvio padrão dos resultados do modelo e θ um ângulo relacionada a correlação entre as duas amostras, sendo $\theta = 0^\circ$ para a correlação igual a 1 e $\theta = 90^\circ$ para um valor nulo deste parâmetro. Isto é feito para todos os modelos testados na modelagem. Aquele modelo cujo ponto correspondente esteja mais próximo ao ponto no eixo das abscissas (desvio padrão da amostra de dados empíricos) é supostamente o modelo numérico que melhor representa a amostra em análise.

No diagrama de Taylor, os três valores estatísticos característicos de cada uma das séries estão interconectados através de uma relação similar a lei dos cossenos, $c^2 = a^2 + b^2 - 2ab \cos(\gamma)$ como:

$$\left(\frac{RMSE}{\sigma_o}\right)^2 = 1 + \left(\frac{\sigma_m}{\sigma_o}\right)^2 - 2\left(\frac{\sigma_m}{\sigma_o}\right) * \rho$$

(2-55)

onde σ_o é o desvio padrão dos dados empíricos e σ_m é o desvio padrão do resultado numérico (ATLASUL, 2018).

Tomando a Figura 6 como exemplo, os pontos “A”, “B”, “C” e “D” são resultados de modelos numéricos candidatos a representar os dados discretos representados pelo ponto Emp. O ponto C apresenta um desvio padrão muito próximo do desvio padrão dos dados do ponto “Emp” pois o raio r entre esse ponto e o “Emp” é muito próximo, porém,

o valor da correlação entre esses dois pontos é de apenas 0.6, que é dado pelo ângulo θ a partir do eixo das abscissas e o ponto C. Já o ponto “A” apresenta um desvio padrão um pouco superior ao ponto que representa os dados do ponto “Emp”, porém a sua correlação é bem mais próxima, por isso segundo o Diagrama de Taylor ele seria o modelo ideal para representar os dados Empíricos.

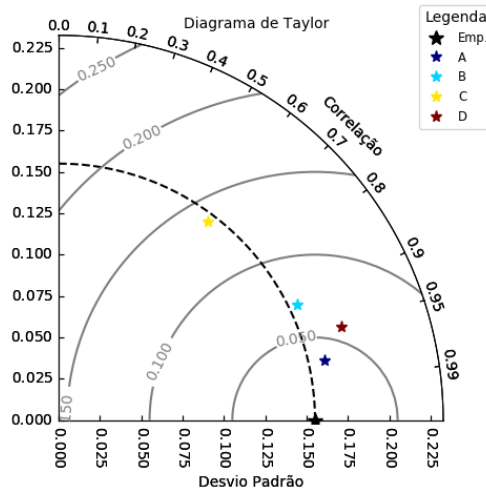


Figura 6- Exemplo Diagrama de Taylor

2.7.2 Comparação Visual

Para visualização dos resultados dos modelos e comparação entre estes e a distribuição de probabilidades conjunta empírica dos dados brutos foram utilizados gráficos bidimensionais com as isolinhas das funções. A Figura 7 ilustra um modelo de distribuição conjunta num gráfico 3-D. A Figura 7 b) ilustra as curvas de isolinhas do modelo e a Figura 7 a) a correspondente as curvas de isolinhas da distribuição conjunta empírica.

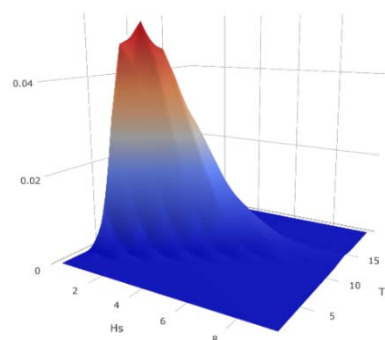


Figura 7- Representação 3-D do modelo de distribuição conjunta de probabilidades

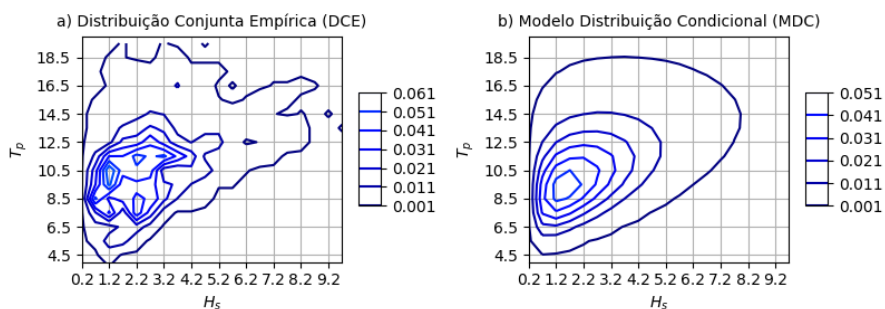


Figura 8- Representação 2-D da distribuição conjunta: a) empírica e b) modelo

Utilizando-se a mesma discretização usada para montar o diagrama de frequências relativas é possível calcular as diferenças entre os modelos matemático e empírico e montar um novo gráfico de isolinhas como mostra a Figura 9. Este tipo de gráfico também é importante para escolher o modelo probabilístico apropriado.

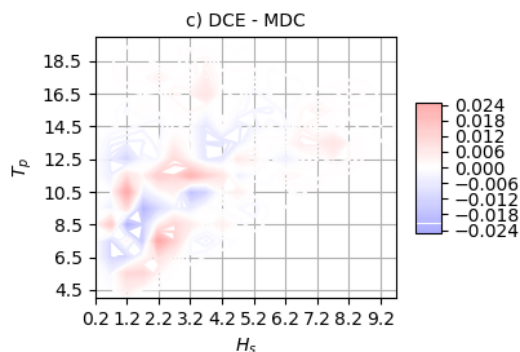


Figura 9- Curvas de diferenças entre distribuições conjuntas: empírica e modelada

3 Ajustes de distribuições conjuntas de H_s e T_p

Neste trabalho o objetivo é investigar o uso dos métodos da distribuição condicional e o baseado na Transformada de Nataf para modelar a distribuição conjunta de parâmetros ambientais de onda: altura significativa H_s e período de pico T_p . Esta distribuição é de extrema importância para o uso de métodos avançados para estimativa de fadiga e de valores extremos da resposta de estruturas marítimas.

3.1 Fonte dos dados

Os dados utilizados no presente trabalho foram extraídos do produto WAVEWATCH III® Production Hindcast Multigrid, disponibilizado pela *National Centers for Environmental Prediction (NCEP) / National Oceanic and Atmospheric Administration (NOAA)* (disponível em http://polar.ncep.noaa.gov/waves/hindcasts/prod-multi_1.php).

O período disponibilizado é do ano de 2005 até o ano da apresentação do presente trabalho, com resolução temporal de 3 em 3 horas. Dentre as diversas grades disponíveis, foi selecionada a grade global que apresenta uma resolução de 0.5° de latitude e longitude. Os dados fornecidos pelo *output* do tipo *Field output in grib2 format* são: H_s , T_p , D_p , componentes U e V de ventos, e H_s , T_p , D_p , de ondas Swell. Destes resultados de modelo, os efetivamente utilizados no presente trabalho foram H_s , T_p , D_p , do período de 2006 a 2016. Estes resultados de modelo serão referidos como dados empíricos ou simplesmente como dados de H_s e T_p no presente trabalho.

Segundo informação dos autores do modelo, “O WAVEWATCH III, é um modelo de terceira geração desenvolvido na NOAA/NCEP. É uma evolução dos modelos WAVEWATCH, desenvolvido na Delft University of Technology (Tolman 1989, 1991), e WAVEWATCH II, desenvolvido na NASA, no Goddard Space Flight Center (e.g., Tolman 1992). No entanto o WAVEWATCH III se diferencia dos seus predecessores em vários importantes aspectos como nas equações governantes, estrutura do modelo, métodos numéricos e parametrizações físicas”. O modelo pode ser obtido gratuitamente em <http://polar.ncep.noaa.gov/waves/wavewatch/>.

3.2 Áreas de estudo

Para os estudos efetuados neste trabalho, foram escolhidos dois pontos posicionados em regiões distintas, um no oceano Atlântico Sul próximo ao Brasil e outro no Mar do Noruega, que é a área estudada por MATHISEN e BITNER-GREGERSEN (1990). As características oceanográficas de cada uma destas regiões, bem como a localização de tais pontos serão apresentadas nos próximos itens e foram propositalmente escolhidas dessa forma por possuírem características bem distintas de estados de mar. Ao aplicar os métodos na área do Mar da Noruega, pretende-se visualizar e entender melhor os resultados obtidos por MATHISEN e BITNER-GREGERSEN (1990) naquela região.

3.2.1 Campo de Libra - Bacia petrolífera de Santos (BR)

A maior bacia sedimentar *offshore* do país é a Bacia de Santos, que se estende de Cabo Frio (RJ) a Florianópolis (SC) (PETROBRAS, 2018). Essa é a bacia onde se encontra a região pré-sal do Brasil, e como é uma região de enorme interesse econômico, foi selecionada como área de estudo no presente trabalho. O trabalho focará nas coordenadas 24.5°S, 42°O, estando localizada dentro da Bacia de Santos, mais especificamente dentro do Campo de Libra. A Figura 10 destaca o contorno do Campo de Libra através do polígono cinza, a coordenada do ponto analisado em vermelho e a batimetria em tons de azul.

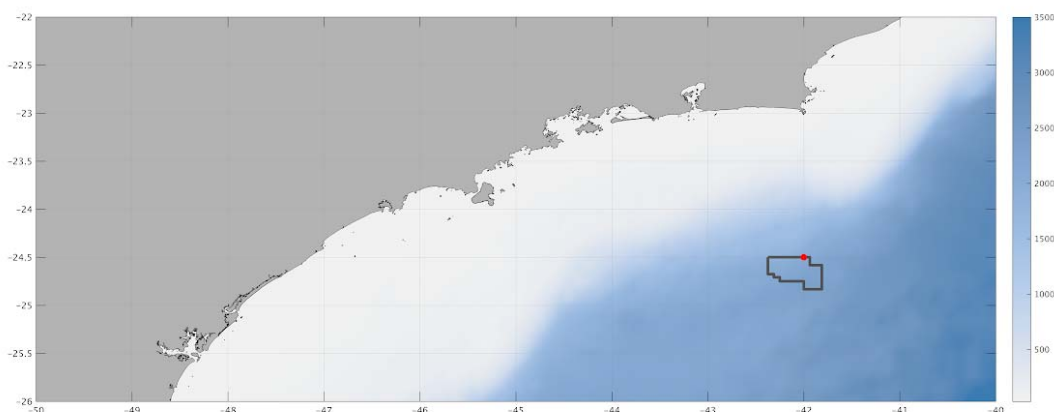


Figura 10- Posicionamento do ponto utilizado no estudo (em vermelho). O polígono (linha cinza) delimita a área do Campo de Libra.

A Tabela II mostra as características estatísticas básicas, com as ondas subdivididas em oito grupos majoritários de direção de propagação dos parâmetros de

onda das coordenadas do ponto analisado. Como mencionado anteriormente, estes parâmetros foram calculados com base na série histórica de dados de fornecidos pela NOAA e gerados pelo modelo WAVEWATCH III. O período dos dados é de 01/2006 a 12/2016, totalizando uma amostra com 32144 estados de mar de 3-h. A direção predominante de propagação das ondas é de Sul para Norte (adotando a convenção usual de oceanografia, que a direção de onde a onda vem, estas ondas seriam na direção S), com uma quantidade insignificante de ondas provenientes das direções Norte e Oeste e ausência completa de ondas provenientes de Noroeste. A região demonstra pouca variação dos parâmetros de ondas ao longo do tempo como se vê através do desvio padrão dos parâmetros de H_s e T_p . Todas as direções com maiores registros de ocorrências de ondas apresentam um baixo valor de correlação entre H_s e T_p .

Tabela II- Dados estatísticos - Campo de Libra

θ_w	H_s			T_p			Número de Ocorrências	ρ_{H_s, T_p}
	Média	Des. Padrão	Max. Obs.	Média	Des. Padrão	Max. Obs.		
N	2,258	0,629	2,98	5,776	0,942	6,93	5	0,489
NE	2,077	0,508	4,88	6,761	0,806	10,14	7408	0,351
E	1,924	0,559	4,92	8,002	1,02	11,75	5162	0,403
SE	1,981	0,645	6,19	9,416	1,766	16,37	4873	0,264
S	2,274	0,704	5,95	11,184	1,937	19,77	12771	0,293
SW	2,873	0,85	5,83	10,308	1,933	17,43	1885	0,343
W	2,702	0,706	4,16	6,633	0,984	8,46	39	0,839
NW	-	-	-	-	-	-	0	-
Todas	2,164	0,681	6,19	9,328	2,396	19,77	32144	0,324

3.2.2 Mar da Noruega

Essa área de estudo, região Tromsøflaket, na costa da Noruega, com coordenadas 71.5°N e 18.5°L, (Figura 11) foi escolhida por ser a área citada no trabalho

de MATHISEN e BITNER-GREGERSEN (1990). O objetivo é comparar os resultados obtidos pelos dois métodos de obtenção de distribuição conjunta abordados no presente trabalho nessa área marítima que possui características de estado de mar bem distintas das observadas na costa brasileira.

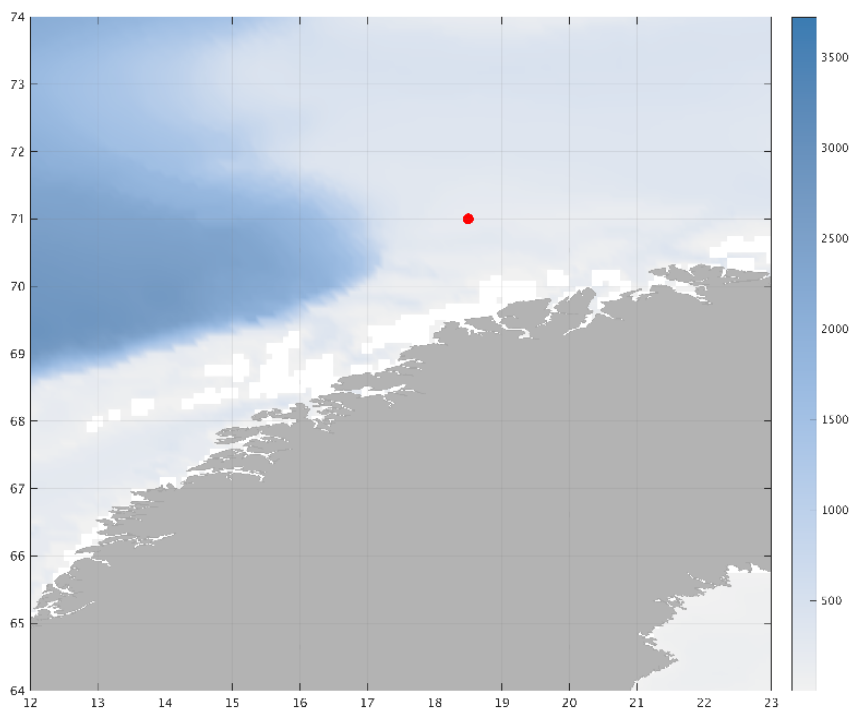


Figura 11- Posicionamento do ponto utilizado no estudo (em vermelho) no mar da Noruega, região Tromsøflaket.

A Tabela III mostra as estatísticas básicas do Mar da Noruega, e assim como na Tabela II, as ondas foram agrupadas em 8 direções majoritárias de propagação. O período dos dados é o mesmo do caso anterior.

A área é caracterizada por possuir variados e intensos estados de mar, com ondas que podem atingir altos valores de H_s e que superam em vários metros os valores alcançados na região de estudo da bacia brasileira. O parâmetro T_p também apresenta valores muito superiores aos encontrados na costa brasileira. A correlação entre H_s e T_p apresenta uma grande variação dependendo da direção de incidência das ondas e ao contrário do que acontece na região brasileira, foram registradas ondas de todas as direções.

Tabela III- Dados estatísticos - Região Tromsøflaket

θ_w	H_s			T_p			Número de Ocorrências	ρ_{H_s, T_p}
	Média	Des. Padrão	Max. Obs.	Média	Des. Padrão	Max. Obs.		
N	2,161	1,062	6,96	7,442	1,305	13,43	2693	0,751
NE	2,233	1,072	8,89	7,597	1,507	13,11	5491	0,68
E	2,081	0,973	7,3	6,909	1,7	12,92	3463	0,636
SE	2,165	0,844	5	5,982	1,222	11,92	505	0,533
S	2,046	1,149	5,51	6,365	1,767	11,89	246	0,569
SW	2,625	1,417	9,58	10,374	2,684	21,81	7885	0,298
W	2,681	1,6	10,89	9,878	2,687	21,51	7622	0,326
NW	2,527	1,385	8,68	7,959	1,622	13,46	4239	0,810
Todas	2,449	1,349	10,89	8,745	2,578	21,81	32144	0,445

3.3 Detalhes sobre o ajuste da função de T_p condicionada a H_s .

No modelo condicional optou-se por modelar a função conjunta de probabilidades através da seguinte expressão:

$$f_{T_p, H_s}(t, h) = f_{H_s}(h) f_{T_p|H_s}(t, h) \quad (3-1)$$

A função densidade de probabilidades marginal de H_s , i.e., $f_{H_s}(h)$, é obtida através de uma técnica, como descrito no item 2.3. A função densidade de probabilidades de T_p condicional a valores de H_s , i.e., $f_{T_p|H_s}(t, h)$, requer uma análise mais detalhada. Para determinar esta função, inicialmente monta-se um diagrama de frequências bidimensional onde os dados de H_s são separados em intervalos de classe com tamanho ΔH_s e os dados T_p são também separados em intervalos de tamanho ΔT_p . A partir daí são identificados os números de ocorrências da amostra de dados em que os valores de H_s e T_z ocorrem dentro de cada uma das células. Dividindo-se o número de ocorrências de cada célula pelo número total de amostras obtém-se o diagrama de frequências relativas ilustrado, genericamente, na Figura 12

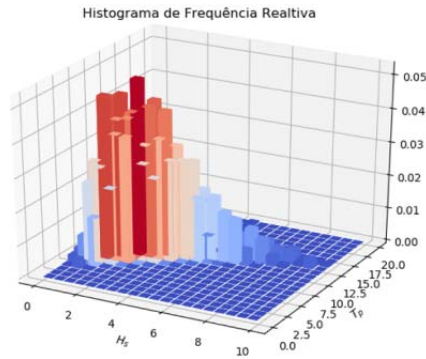


Figura 12-Histograma de frequência relativa de ocorrências de $H_s - T_p$

A partir do diagrama de frequências relativas, é possível ajustar uma distribuição de probabilidades para T_p para cada intervalo de H_s , como ilustra a Figura 13. Para modelagem da distribuição conjunta é necessário utilizar um mesmo modelo de distribuição para todos os intervalos, por exemplo, Lognormal, Weibull, etc. Na prática escolhe-se aquele modelo que mais ocorre considerando todos os intervalos de H_s .

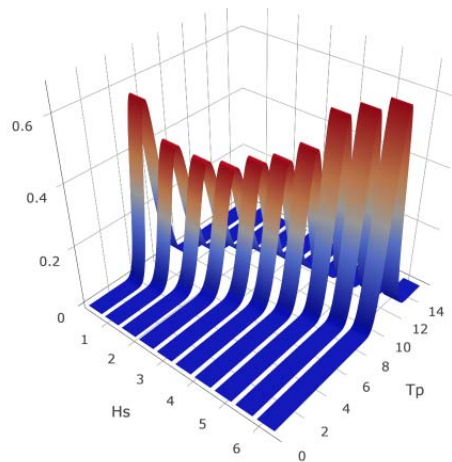


Figura 13- PDFs de T_p para os intervalos de H_s

Com o mesmo modelo de distribuição para intervalo de H_s , o que ocorre é que os parâmetros deste modelo variam de intervalo para intervalo. Tomando-se por exemplo, uma distribuição Lognormal, os seus parâmetros ξ e λ são funções de H_s . Para obter o modelo condicional torna-se necessário obter funções contínuas destes parâmetros em função de H_s . Assim a partir dos pontos discretos representados pelos pontos médios de cada intervalo de H_s é necessário estabelecer ou ajustar as funções $\xi(h)$ e $\lambda(h)$, como mostra a Figura 14

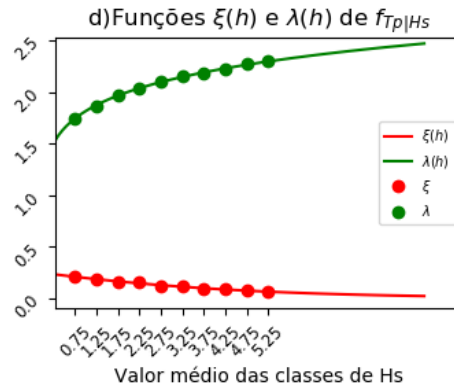


Figura 14- Valores discretos ξ e λ e suas funções ajustadas $\xi(h)$ e $\lambda(h)$

A definição do modelo matemático para representar as curvas do comportamento dos parâmetros da distribuição de probabilidades escolhida é de fundamental importância neste modelo baseado na distribuição condicional. Usualmente, na utilização prática da distribuição conjunta, haverá necessidade de calcular probabilidades associadas a valores de H_s maiores que os observados nos dados, ou seja, haverá necessidade de fazer extrapolação usando as funções ajustadas. Este aspecto requer muito cuidado e é muito relevante na prática, pois, por exemplo, as extrapolações não podem conduzir a valores ou situações físicas inválidas ou irreais. No caso específico de análise de ondas, BITNER-GREGERSEN e HAVER (1991) sugerem que no caso de uma distribuição Lognormal $\xi(h)$ e $\lambda(h)$ podem ser modelados por

$$\xi(h) = a_1 e^{-b_1 h} + a_2 e^{-b_2 h} \tag{3-2}$$

$$\lambda(h) = \alpha_1 + \alpha_2 h^\beta \tag{3-3}$$

onde os parâmetros $\alpha_1, \alpha_2, \beta, a_1, a_2, b_1$ e b_2 podem ser ajustados utilizando técnicas de regressão-linear. Desta forma, no caso de uma distribuição Lognormal ser apropriada, a distribuição de T_p condicional a valores H_s seria dada por:

$$f_{T_p|H_s}(t, h) = \frac{1}{\xi(h) t \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln(t) - \lambda(h)}{\xi(h)}\right)^2\right) \tag{3-4}$$

No caso de a Lognormal não ser a distribuição mais apropriada, tratamento similar deve ser dado aos parâmetros que descrevem o modelo de distribuição escolhido.

3.4 Sequência usada na modelagem probabilística de H_s e T_p .

As séries de valores de H_s e T_p de cada região foram separadas em função da direção de propagação θ_w . No total, os dados foram separados em oito intervalos de direções cujos centros são os 8 pontos cardeais e colaterais da rosa dos ventos. Os dados estatísticos das ondas nestas direções para as duas locações investigadas já foram apresentados nas Tabela I e Tabela II. No entanto, devido à enorme quantidade de dados gerados, apenas algumas direções foram analisadas no presente trabalho. O critério utilizado para fazer tal seleção se baseou no número de ondas observadas na direção, nas alturas significativas dessas ondas, e na correlação entre H_s e T_p de cada área, priorizando a formação de grupos com alta correlação e grupos com baixa correlação a fim de analisar o comportamento dos diferentes métodos de distribuição conjunta com diferentes níveis de dependência entre H_s e T_p .

Para cada direção escolhida para esse estudo, os dados dos parâmetros de ondas foram tratados da seguinte forma:

- Para a amostra de H_s da direção escolhida foram testados 2 tipos de PDF, Weibull e Lognormal, a fim de identificar qual delas melhor representa o conjunto de dados.
- Para amostra de valores T_p foram também testadas as mesmas duas distribuições. Observa-se que esta distribuição é usada na metodologia baseada no modelo de Nataf para obtenção da distribuição conjunta de H_s e T_p .
- Para o ajuste da distribuição de T_p condicional a H_s foi utilizado um diagrama de frequência com intervalos de 0,5m para H_s e 1,0s para T_p na área brasileira e de 0,625m para H_s e 1,25s para T_p na região norueguesa.
- A verificação dos melhores ajustes para H_s foi feita com base na comparação visual baseada na distribuição cumulativa dos dados empíricos e a correspondente do modelo estatístico. Além disto, em determinados casos fez-se uma investigação sobre os valores extremos para definir o melhor ajuste.

- A partir das informações acima foram criadas as distribuições conjuntas pelo modelo da distribuição condicional e pela metodologia baseada no modelo de Nataf.
- Ambos os modelos são verificados em relação a função densidade de probabilidades conjunta empírica através dos parâmetros estatísticos, diagrama de Taylor e curvas de diferenças entre modelo e a empírica (veja item 2.7.2).

3.5 Resultados obtidos

3.5.1 Bacia de Santos – Costa Brasileira

Dois estudos de caso foram feitos com os dados dessa localização, e ambos considerando a mesma direção de ondas, mudando apenas a quantidade de dados disponíveis para cada estudo. A direção escolhida foi a direção Sudoeste (SW) por representar a direção da chegada de tempestades na área, cujos ventos fortes geram altas ondas. Nesta direção a correlação entre H_s e T_p é relativamente baixa (veja Tabela II).

Foram selecionados dois períodos de dados: uma série histórica de 11 anos (01/2006 a 12/2016) e uma série menor de 02 anos (01/2015 a 12/2016). A primeira série contém 1885 estados de mar nesta direção e a segunda apenas 379. Esta segunda série foi selecionada para comparar o desempenho dos modelos quando se dispõe de poucos dados disponíveis, situação muito comum quando a coleta de dados é feita através de equipamentos oceanográficos.

3.5.1.1 Ondas de Sudoeste – Série histórica de 11 anos de dados

Foram obtidos através do método da máxima verossimilhança os parâmetros de uma distribuição Lognormal e de uma distribuição de Weibull para H_s . As funções cumulativas de probabilidades destas duas distribuições juntamente com a CDF empírica dos dados são mostradas na Figura 15. Observando a Figura 15, não é claro qual a função que melhor se ajusta para H_s , embora a distribuição de Weibull pareça aderir melhor aos dados quando o valor da função tende a 1, que é a região mais importante por conter os valores de extremos deste parâmetro.

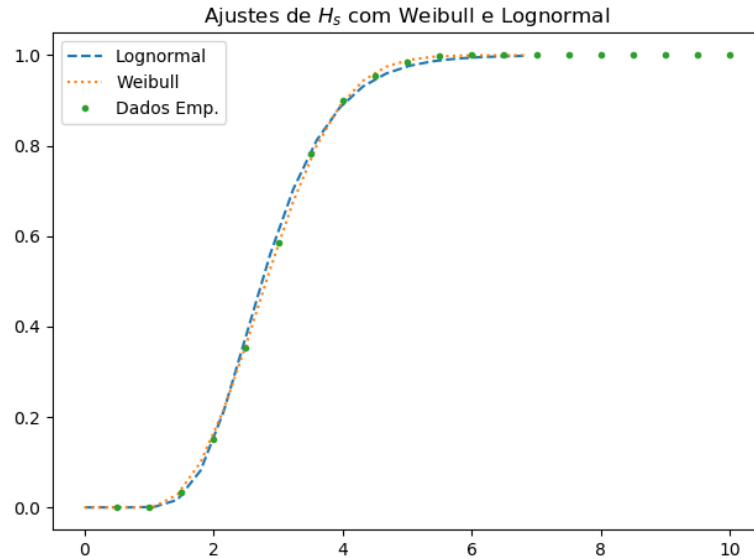


Figura 15- Santos 11 anos: Ajuste de CDF do tipo Lognormal e Weibull para série de H_s

Para definir qual é distribuição que representa H_s foi feita uma análise do valor extremo centenário mais provável usando ambas as distribuições usando a teoria de extremos decrita no capítulo 2.2 e uma estimativa aproximada do mesmo baseado nos valores extremos anuais da amostra como descrito no capítulo 2.1.6. Neste último caso ajustou-se uma distribuição de Gumbel para a amostra de 11 valores extremos anuais observados na amostra e estimou-se o valor centenário mais provável como sendo o valor cuja probabilidade de excedência corresponde a 1/100. As Figura 16 ilustra graficamente este último ajuste assim como o correspondente valor centenário estimado (em torno de 6,9m). As Figura 17 e Figura 18 ilustram os correspondentes valores centenários estimados a partir do ajuste de uma Lognormal e de uma Weibull, respectivamente. Como pode ser observado, a distribuição Lognormal superestima largamente o valor centenário (8,9m contra 6,9m) e, portanto, a distribuição de Weibull é aque melhor representa este conjunto de dados de H_s para esta locação.

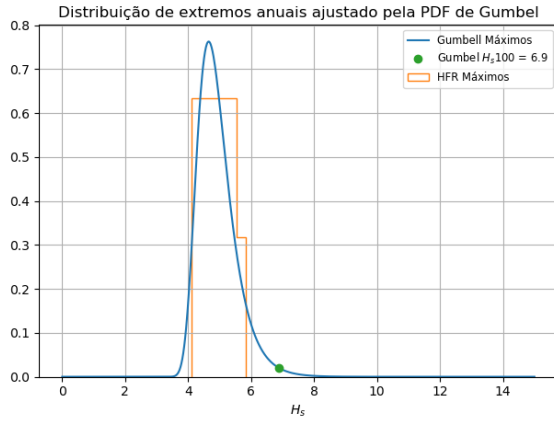


Figura 16- Distribuição de extremos anuais ajustado pela PDF de Gumbel

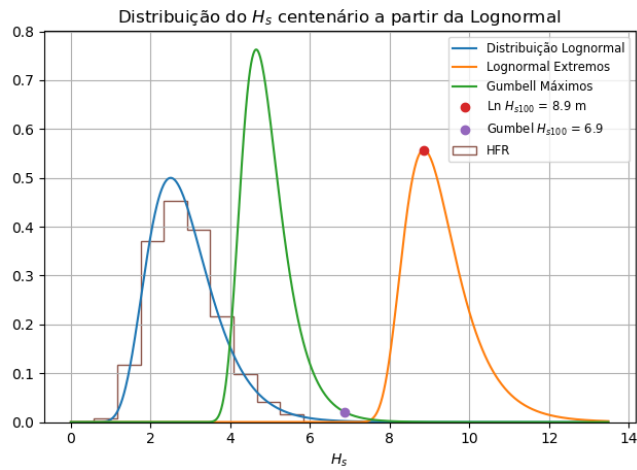


Figura 17- Distribuição do H_s centenário a partir da Lognormal

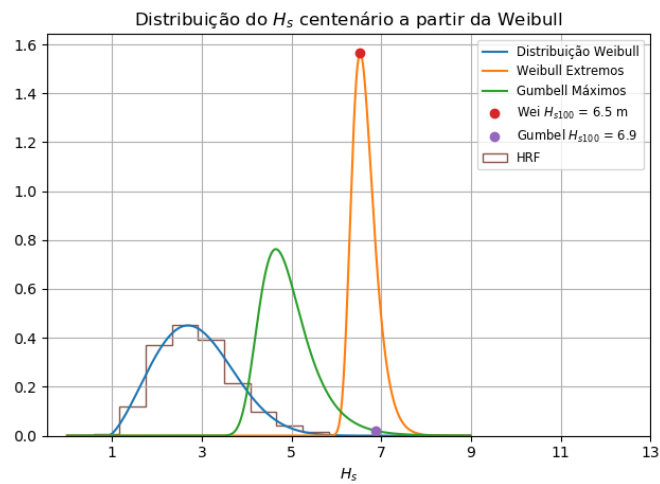


Figura 18- Distribuição do H_s centenário a partir da Weibull

Nos demais resultados apresentados no presente trabalho, a distribuição de Weibull será usada como padrão para representar H_s , pois na maioria de outras análises ela apresentou uma aderência melhor com relação à distribuição empírica e uma melhor coerência com relação aos valores de onda centenária. Esta observação é corroborada também pelo que é sugerido por MATHISEN e BITNER-GREGERSEN (1990).

De forma similar ao que foi realizado para H_s , foram obtidos através do método da máxima verossimilhança os parâmetros de uma distribuição Lognormal e de uma distribuição de Weibull com $\theta = 0$ para modelar T_p . Os resultados obtidos são apresentados na Figura 19.

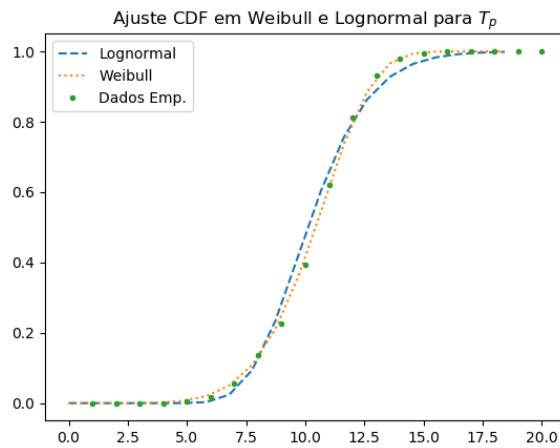


Figura 19- Ajuste de CDFs Lognormal e Weibull para T_p . Costa brasileira.

A Figura 20 demonstra o ajuste das CDFs do tipo Lognormal e Weibull para T_p em função das classes de H_s , que serão utilizados no método MDC.

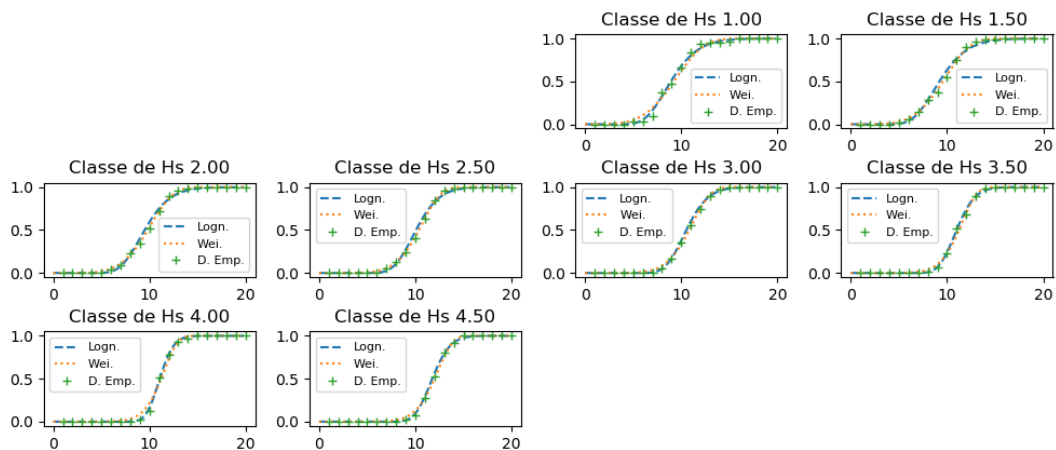


Figura 20- Ajuste da CDF condicional de T_p : teste com Weibull e Lognormal. Costa brasileira.

Embora a CDF de Weibull pareça se ajustar bem aos dados empíricos, a CDF do tipo Lognormal também se adequou bem. O trabalho MATHISEN e BITNER-GREGERSEN (1990) sugere que para parâmetros de T_p seja utilizado a distribuição do tipo Lognormal. Para que haja uma maior semelhança com o presente trabalho, a distribuição Lognormal também foi escolhida tanto para representar a distribuição marginal de T_p bem com a sua função condicional nos demais casos investigados.

Após as definições das distribuições de H_s e T_p (marginal e condicional) foi possível obter a distribuição conjunta destes parâmetros pelos dois métodos investigados neste trabalho. A Figura 21 apresenta um dos resultados obtidos pelo método da distribuição condicional utilizando classes de 0,625m de H_s . O tamanho das classes de H_s foi escolhido através da verificação visual de um histograma de frequências com H_s variando de 0 a 10 metros, subdivido em 16 classes. Através dessa figura, é possível verificar que o modelo se ajusta bem ao histograma da série de dados empíricos, no entanto, é possível verificar pelo gráfico da função $\xi(h)$, que a mesma se comporta de forma completamente anômala para valores de H_s superiores a 5,0 m.

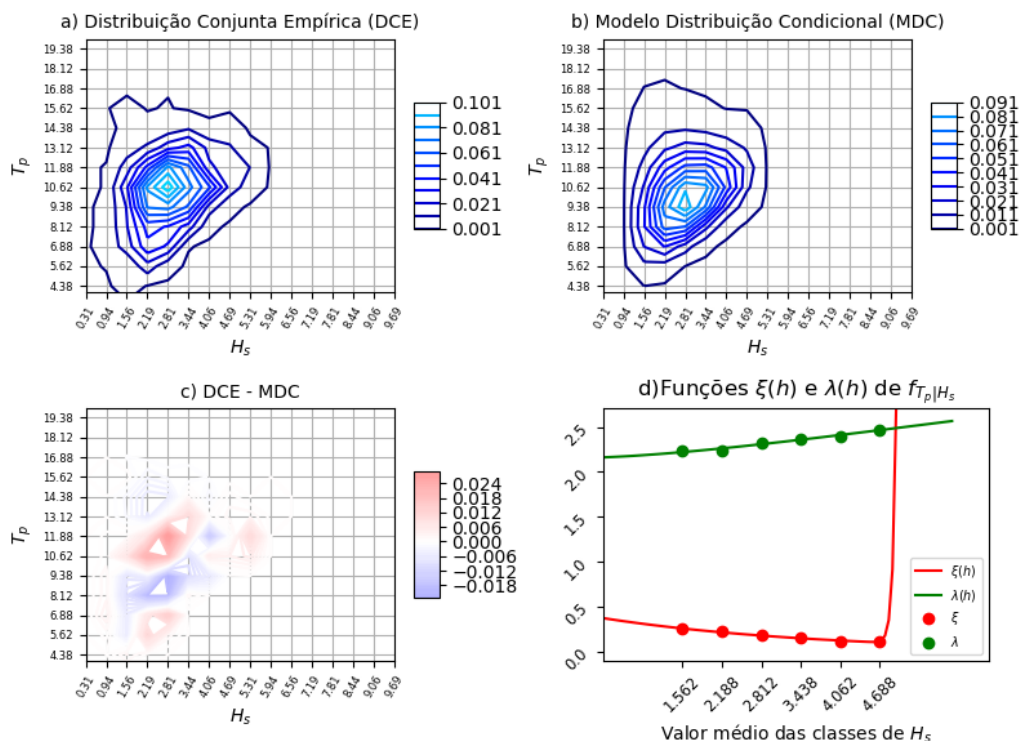


Figura 21- Distribuição conjunta pelo método da distribuição condicional – costa brasileira (11 anos de dados)

Para contornar o problema identificado na função contínua do parâmetro ξ , consequência de uma falha na estimação dos parâmetros α_1, α_2, b_1 e b_2 pelo método de regressão linear, o mesmo modelo foi refeito utilizando uma separação diferente das classes de H_s . Foi utilizado para um intervalo H_s de 0,5m para montar o diagrama de frequência relativa, mantendo-se o intervalo de 1,0 s para T_p . Os resultados obtidos com esta alteração são ilustrados na Figura 22.

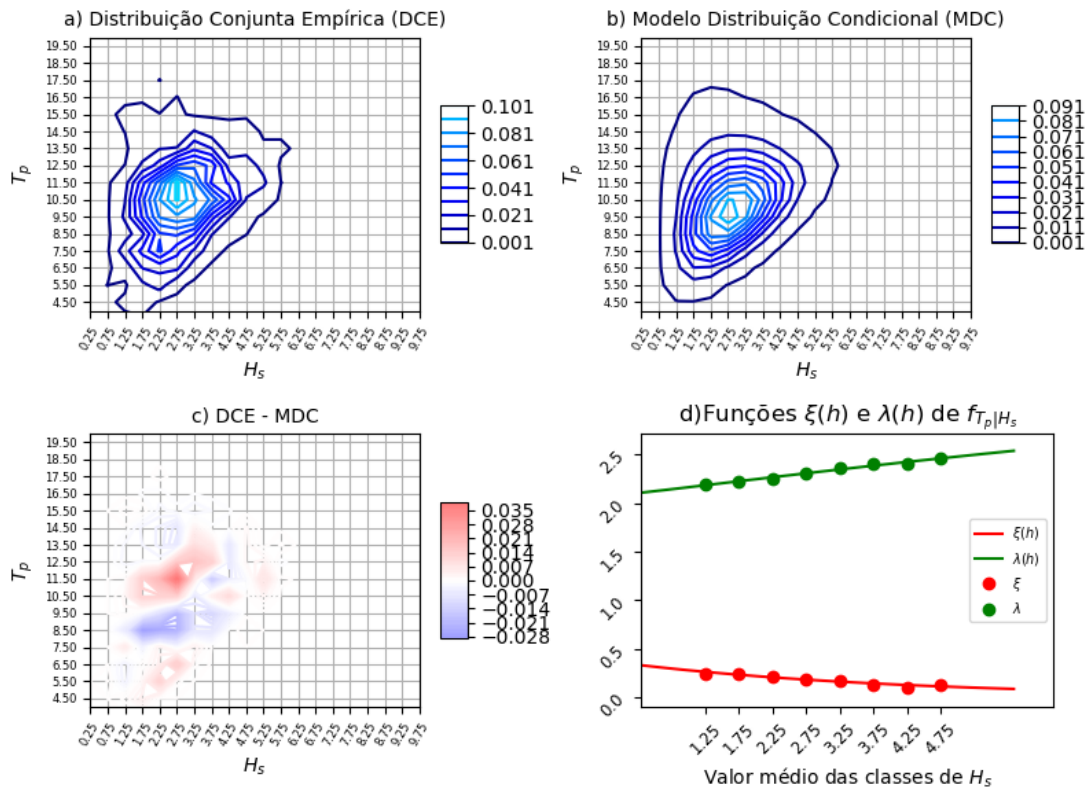


Figura 22- Distribuição conjunta pelo método da distribuição condicional – costa brasileira (11 anos de dados). Intervalo de H_s alterado.

A Figura 23 apresenta o resultado obtido utilizando-se a metodologia baseada no método de Nataf para obter a distribuição conjunta $H_s - T_p$. Aparentemente, a comparação entre a distribuição empírica da série bruta de dados e a distribuição obtida com esta metodologia não apresentam um aspecto tão semelhante nos limites do contorno. Porém, ao se fazer a comparação visual através do gráfico da diferença entre o histograma e o modelo é possível verificar que tanto este modelo quanto o MDC apresentam uma grande semelhança com a distribuição empírica.

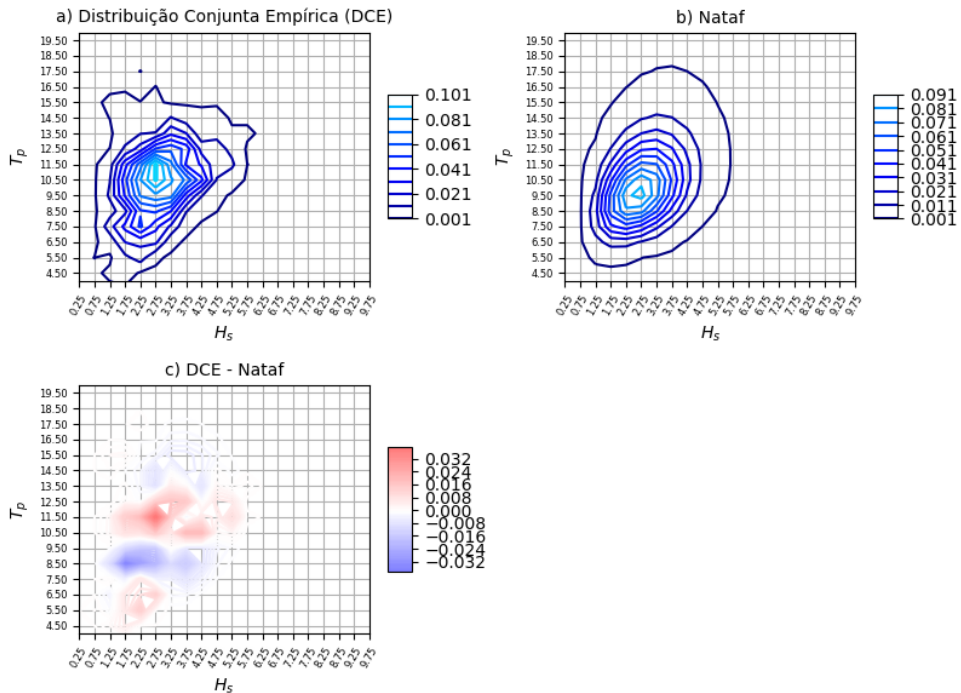


Figura 23- Distribuição conjunta pela metodologia baseada no modelo de Nataf – costa brasileira (11 anos de dados)

A Figura 24 ilustra o diagrama de Taylor para ambos os modelos investigados. Nesta figura observa-se um comportamento muito similar dos dois modelos, confirmando o visto através da análise das figuras, demonstrando que os valores característicos de desvio padrão, correlação e raiz do erro médio quadrático obtidos com os diferentes métodos são bem próximos, porém com uma ligeira vantagem para o MDC.

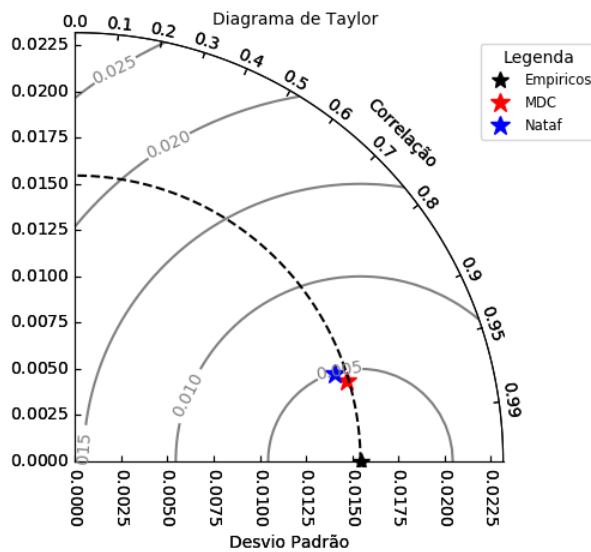


Figura 24- Diagrama de Taylor modelos condicional e Nataf - costa brasileira (11 anos de dados)

A Tabela IV apresenta uma comparação entre os principais parâmetros estatísticos das variáveis H_s e T_p . As células em verde indicam o método cujo resultado se aproxima mais dos dados.

Tabela IV - Comparação entre os parâmetros estatísticos: dados brutos e modelos condicional e Nataf (Costa brasileira – 11 anos de dados)

Parâmetros estatísticos	Modelo Condicional	Nataf	Empírica	Diferença Cond. - Emp.	Diferença Nataf - Emp.
Desvio padrão H_s (m)	0,8502	0,8490	0,8497	0,0005	0,0007
Desvio padrão T_p (s)	1,9581	2,0687	1,9329	0,0252	0,1358
Esperança de $H_s T_p$ (ms)	30,1179	30,1082	30,3198	0,2019	0,2116
Desvio padrão $H_s T_p$ (ms)	12,0030	12,4866	17,1540	5,1509	4,6674
$\rho_{H_s - T_p}$	0,3453	0,3681	0,3428	0,0025	0,0253

Observando todos os resultados descritos acima, é possível identificar que os dois modelos de distribuição conjunta fornecem valores similares entre si.

3.5.1.2 Ondas de Sudoeste – Série histórica de 2 anos de dados

A fim de comparar o desempenho dos modelos quando se há poucos dados disponíveis, tomou-se os valores observados num período de 2 anos, i.e., de 01/2015 a 12/2016. Nesse período, o total de estados de mar que apresentaram ondas na direção Sudoeste foi de 379. As distribuições usadas para esta base menor foram as mesmas do item anterior, variando-se apenas os parâmetros em função da base reduzida de dados.

A Figura 25 ilustra os resultados para a distribuição conjunta obtida pelo modelo condicional e a Figura 26 mostra os correspondentes resultados obtidos pela metodologia baseado no modelo de Nataf. Comparando-se estas duas figuras com os resultados anteriores (maior base de dados), verifica-se que o modelo Nataf é o que apresenta uma distribuição conjunta mais coerente com aqueles resultados. A distribuição conjunta obtida pelo método condicional é mais susceptível ao tamanho da base de dados do que aquela obtida com Nataf. Isto deve ser explicado pelo comportamento da função condicional $\lambda(h)$ que apresenta uma tendência aparentemente incorreta e incoerente quando comparada com a mesma função apresentada na Figura 22.

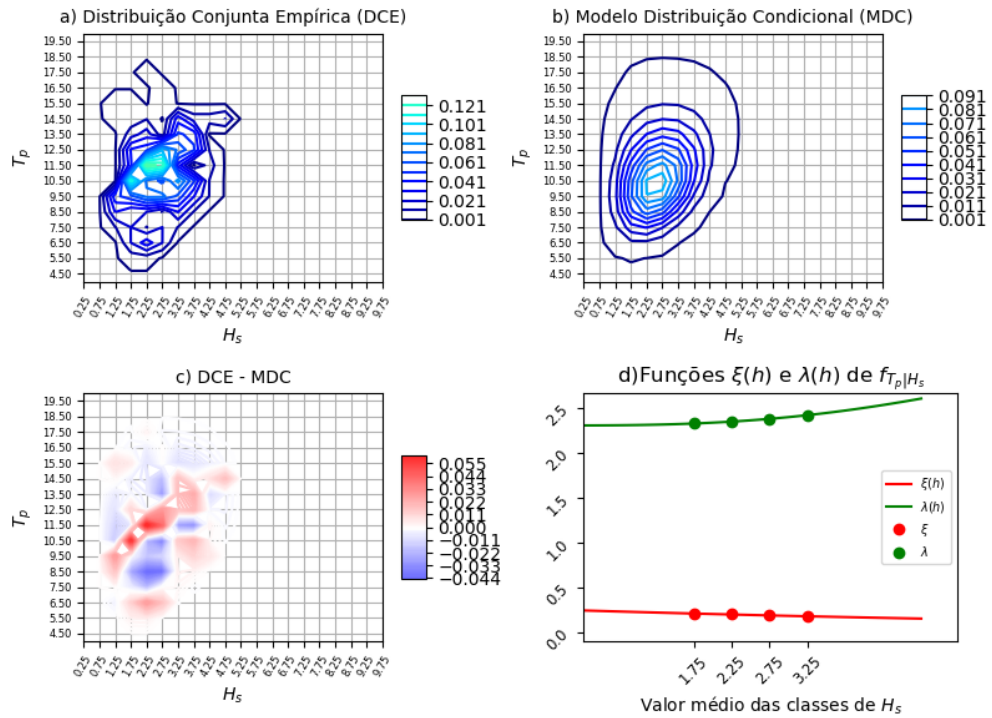


Figura 25- Distribuição conjunta pelo método da distribuição condicional – costa brasileira (2 anos de dados)

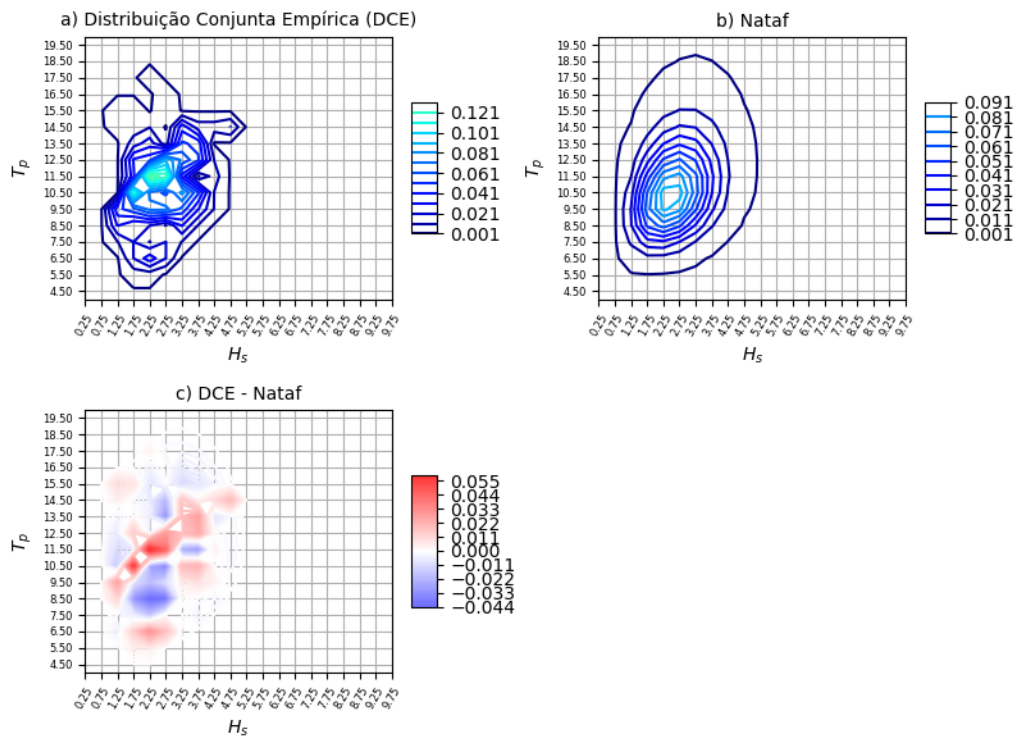


Figura 26- Distribuição conjunta pela metodologia baseada no modelo de Nataf – costa brasileira (2 anos de dados)

A Tabela V a seguir mostra os cálculos de alguns parâmetros estatísticos, comparando os resultados das distribuições conjuntas com o histograma dos dados brutos, e entre si. Através desses resultados, observa-se uma tendência do modelo de Nataf apresentar dados estatísticos

Tabela V- Comparação entre os parâmetros estatísticos: dados brutos e modelos condicional e Nataf (Costa brasileira – 02 anos de dados)

Parâmetros estatísticos	Modelo Condicional	Nataf	Empírica	Diferença Cond. - Emp.	Diferença Nataf - Emp.
Desvio padrão H_s (m)	0,7412	0,7412	0,7606	0,0194	0,0194
Desvio padrão T_p (s)	2,1466	2,1266	2,0692	0,0775	0,0575
Esperança de $H_s T_p$ (ms)	29,3708	29,1124	29,4891	0,1184	0,3767
Desvio padrão $H_s T_p$ (ms)	11,4671	12,2994	16,4432	4,9761	4,1438
$\rho H_s - T_p$	0,3971	0,3845	0,2834	0,1137	0,1011

A Figura 26 mostra o diagrama de Taylor das distribuições conjuntas comparando com os dados do histograma. Nesse diagrama ambos os modelos apresentam comportamento semelhante, porém, quando comprados aos correspondentes na Figura 24, ambos também são menos precisos que aqueles obtidos com a base de 11 anos de dados.

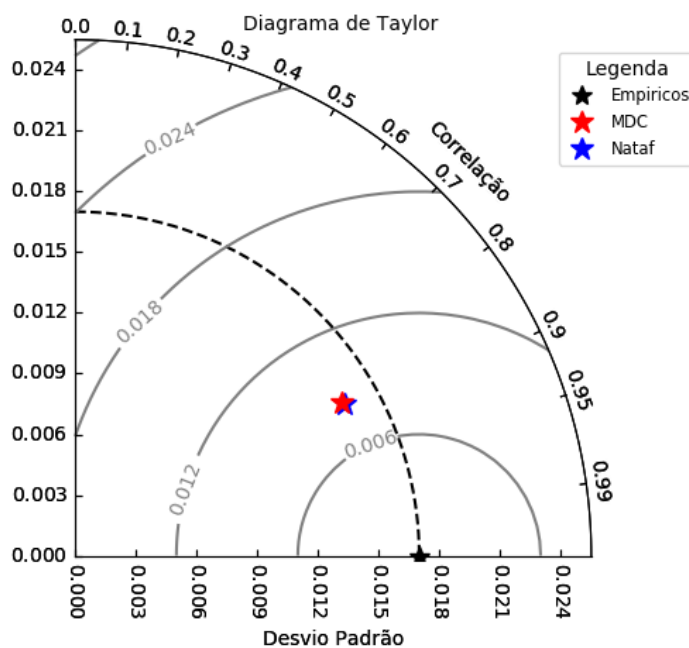


Figura 27- Diagrama de Taylor modelo condicional e Nataf - costa brasileira (2 anos de dados)

3.5.1.3 Ondas de Sudoeste – Resumo das análises

Para a amostra de dados de onda na direção Sudoeste na costa brasileira, contida na série histórica do banco de dados utilizado, verifica-se que a metodologia baseada no modelo de Nataf pareceu ser o modelo mais apropriado para representar a distribuição conjunta. Este modelo se mostrou pouco sensível ao tamanho da base de dados. O modelo baseado na distribuição condicional é bastante sensível com relação às funções utilizadas para representar os parâmetros da distribuição condicional.

3.5.2 Mar da Noruega

Dois estudos de caso foram feitos com os dados da costa da Noruega, e ambos considerando a mesma direção de ondas, mudando apenas a quantidade de dados disponíveis para cada estudo. A direção escolhida foi a direção Noroeste (NW) por apresentar a maior correlação entre H_s e T_p entre todas as direções disponíveis e por também apresentar altos valores destas duas variáveis (veja Tabela III), parâmetros que podem influenciar os modelos de distribuição condicional.

Tal como no caso anterior, foram selecionados dois períodos de dados: uma série histórica de 11 anos (01/2006 a 12/2016) e uma série menor de 02 anos (01/2015 a 12/2016). A primeira série contém 4239 estados de mar na direção selecionada e a segunda 621. Esta segunda série foi selecionada para comparar o desempenho dos modelos quando se dispõe de poucos dados disponíveis.

3.5.2.1 Ondas de Noroeste: Série histórica de 11 anos de dados

Para essa série de dados, foram testadas duas distribuições de probabilidade para H_s , uma Lognormal e uma Weibull. Os parâmetros destas distribuições para esta amostra foram obtidos através do método da verossimilhança e os ajustes obtidos são mostrados na Figura 28. Ao contrário da costa brasileira, agora fica claro que a distribuição de Weibull é a que melhor se ajusta à distribuição dos dados empíricos.

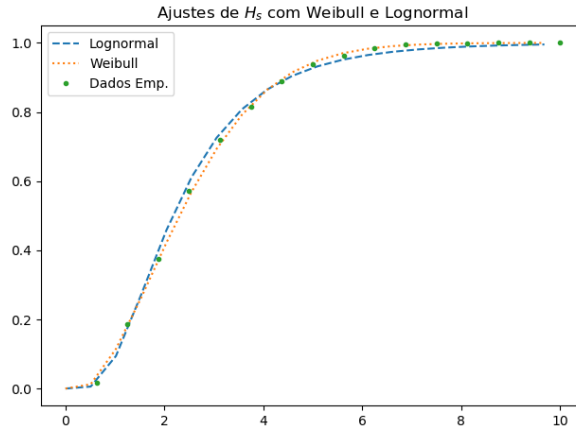


Figura 28- Ajuste de CDFs Lognormal e Weibull para H_s . Costa da Noruega.

A Figura 29 mostra os modelos de Weibull e Lognormal ajustados para a amostra de T_p , onde identifica-se que a ambas as funções Weibull e Lognormal se ajustam adequadamente aos dados. Assim como no caso da região brasileira, a distribuição Lognormal foi a escolhida.

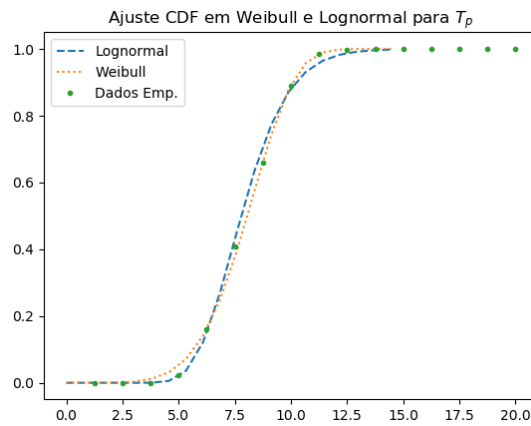


Figura 29- Ajuste de CDFs Lognormal e Weibull para T_p . Costa da Noruega.

Os modelos de Weibull e Lognormal também foram usados na modelagem da distribuição de T_p condicional a valores de H_s . Como ilustra a Figura 30, a distribuição Lognormal foi a escolhida para representar a distribuição de dados de T_p .

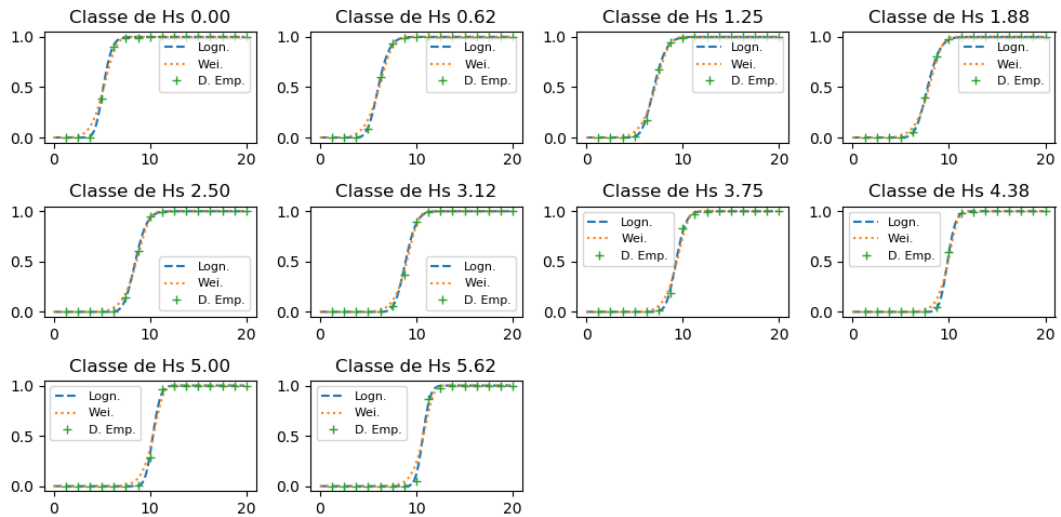


Figura 30 - Ajuste da CDF condicional de T_p : teste com Weibull e Lognormal. Costa da Noruega.

A Figura 31 mostra os resultados obtidos para a distribuição conjunta de H_s e T_p através do modelo baseado na distribuição condicional. Este modelo apresenta uma grande semelhança com os dados empíricos. Além disto, os ajustes das funções de $\xi(h)$ e $\lambda(h)$ também demonstram um ótimo ajuste aos dados pontuais calculados, mantendo uma tendência coerente para extrapolações além da base de dados observados.

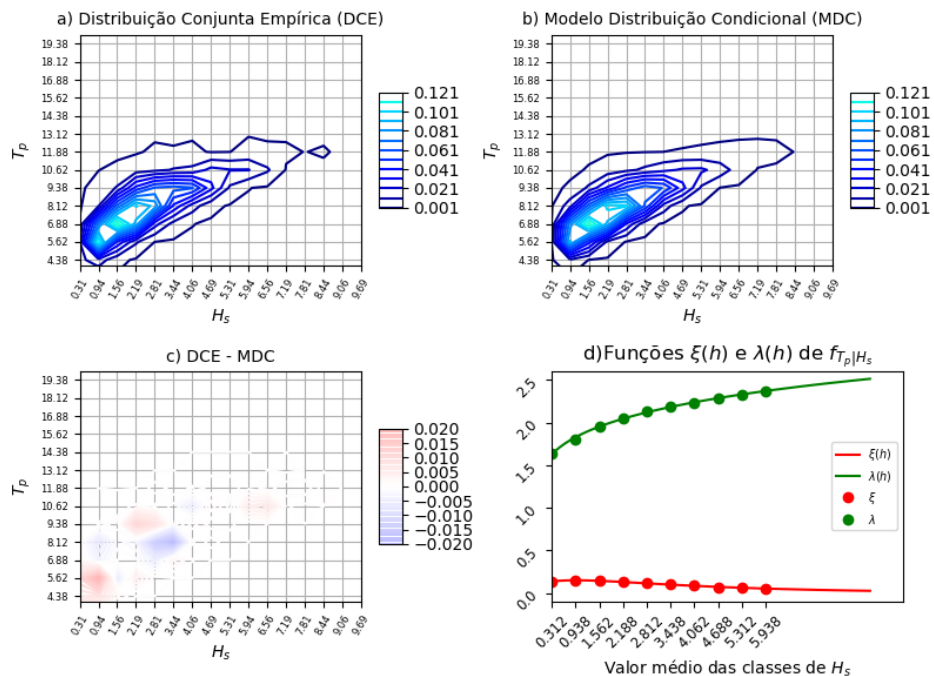


Figura 31- Distribuição conjunta pelo método da distribuição condicional – costa da Noruega (11 anos de dados).

A Figura 32 apresenta os resultados obtidos para a distribuição conjunta de H_s e T_p obtidos através da modelagem baseada no modelo de Nataf. No geral este modelo também apresenta uma boa representação do modelo empírico, porém, quando se analisa as ondas com maiores valores de H_s , nota-se um descolamento das curvas de nível onde o modelo associa valores de T_p maiores daqueles observados no histograma dos dados brutos.

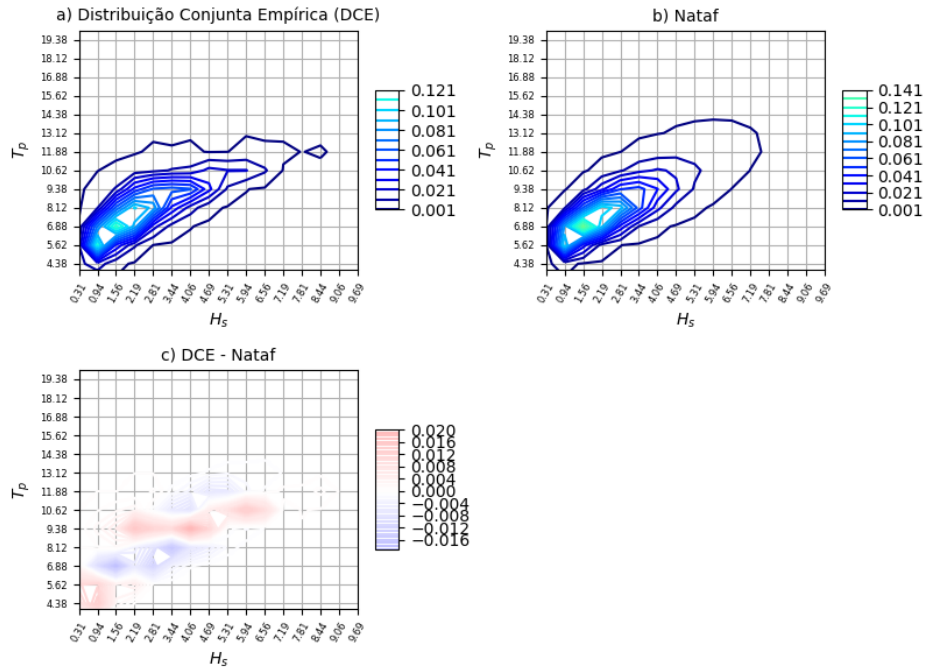


Figura 32- Distribuição conjunta pelo método Nataf – costa da Noruega (11 anos de dados).

A Tabela VI mostra que as estatísticas calculadas para a distribuição de Nataf levam uma ligeira vantagem sobre aquelas obtidas pela metodologia da distribuição condicional. Por outro o diagrama de Taylor, representado na Figura 33, apresenta uma leve vantagem para esta última metodologia.

Tabela VI- Comparação entre os parâmetros estatísticos: dados brutos e modelos condicional e Nataf (Costa da Noruega – 11 anos de dados)

Parâmetros estatísticos	Modelo Condicional	Nataf	Empírica	Diferença Cond. - Emp.	Diferença Nataf - Emp.
Desvio padrão H_s (m)	1,3636	1,3476	1,3849	0,0212	0,0373
Desvio padrão T_p (s)	1,5759	1,6628	1,6219	0,0461	0,0408
Esperança de $H_s T_p$ (ms)	21,8805	21,7657	21,9499	0,0694	0,1842
Desvio padrão $H_s T_p$ (ms)	15,5100	15,7653	17,9467	2,4367	2,1814
$\rho_{H_s - T_p}$	0,8004	0,8190	0,8103	0,0099	0,0087

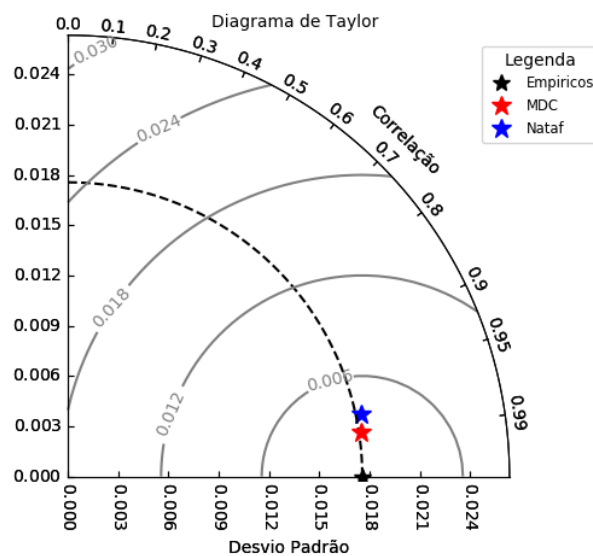


Figura 33- Diagrama de Taylor modelo condicional e Nataf - costa da Noruega (11 anos de dados)

Desta forma, nesta locação, ambos os modelos podem ser considerados razoáveis para representar a distribuição conjunta de H_s e T_p .

3.5.2.2 Ondas de Noroeste – Série histórica de 2 anos de dados

Os resultados obtidos para a distribuição conjunta de H_s e T_p para a direção Noroeste da costa da Noruega considerando uma série histórica reduzida de dados são mostrados na Figura 34 e na Figura 35, respectivamente, para os modelos baseado na distribuição condicional e o de Nataf. A Tabela VII apresenta uma comparação entre os principais parâmetros estatísticos das variáveis obtidas com os modelos e os valores correspondentes calculados através da amostra de dados. A Figura 36 ilustra o diagrama de Taylor que compara a performance dos dois modelos. Mais uma vez a performance dos dois modelos é bastante similar. Entretanto, é interessante observar que, em contraste com a performance para a locação na costa brasileira, o modelo condicional não apresentou nenhuma anomalia com o conjunto reduzido de dados.

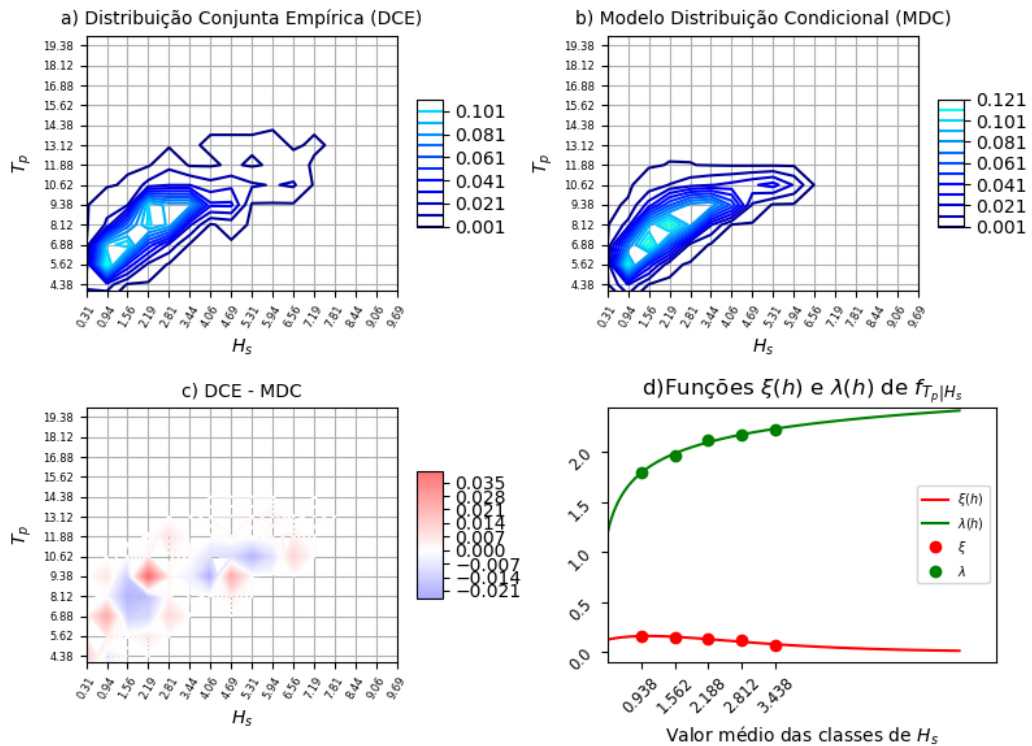


Figura 34- Distribuição conjunta pelo método da distribuição condicional – costa da Noruega (2 anos de dados).

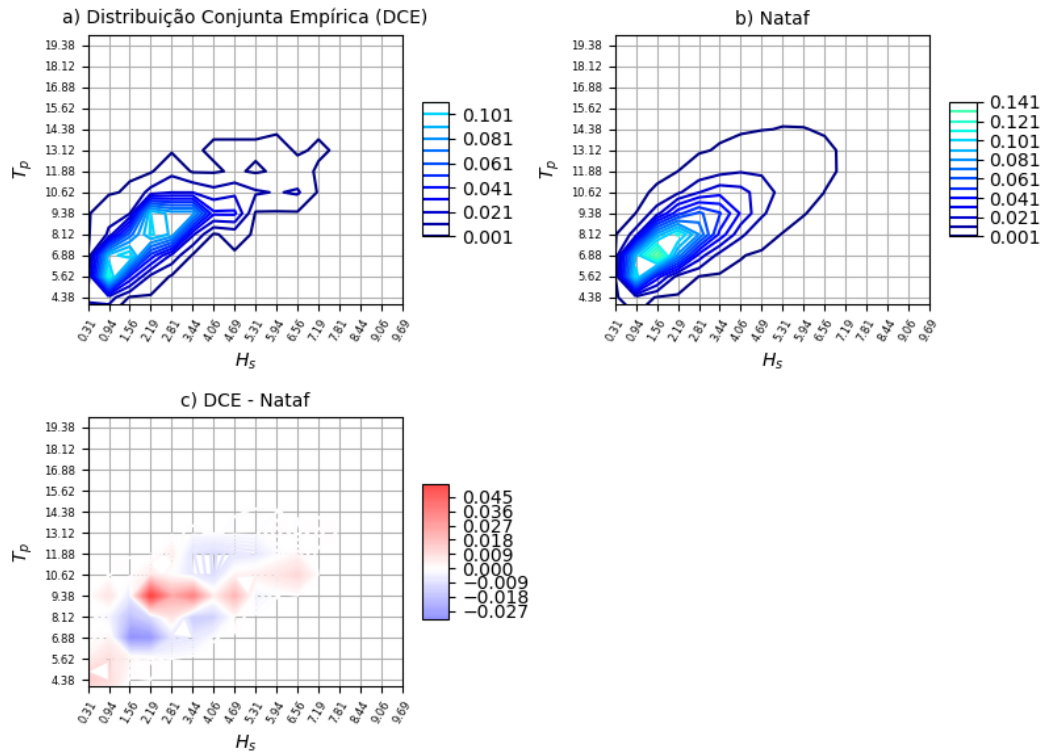


Figura 35- Distribuição conjunta pelo método Nataf – costa da Noruega (2 anos de dados).

Tabela VII- Comparação entre os parâmetros estatísticos: dados brutos e modelos condicional e Nataf (Costa da Noruega – 2 anos de dados)

Parâmetros estatísticos	Modelo Condicional	Nataf	Empírica	Diferença Cond. - Emp.	Diferença Nataf - Emp.
Desvio padrão H_s (m)	1,2117	1,1949	1,2453	0,0336	0,0504
Desvio padrão T_p (s)	1,6524	1,7560	1,7134	0,0610	0,0426
Esperança de $H_s T_p$ (ms)	20,6843	20,5173	21,0280	0,3437	0,5107
Desvio padrão $H_s T_p$ (ms)	13,8035	14,3807	16,8079	3,0044	2,4272
$\rho_{H_s - T_p}$	0,8104	0,8227	0,7905	0,0199	0,0322

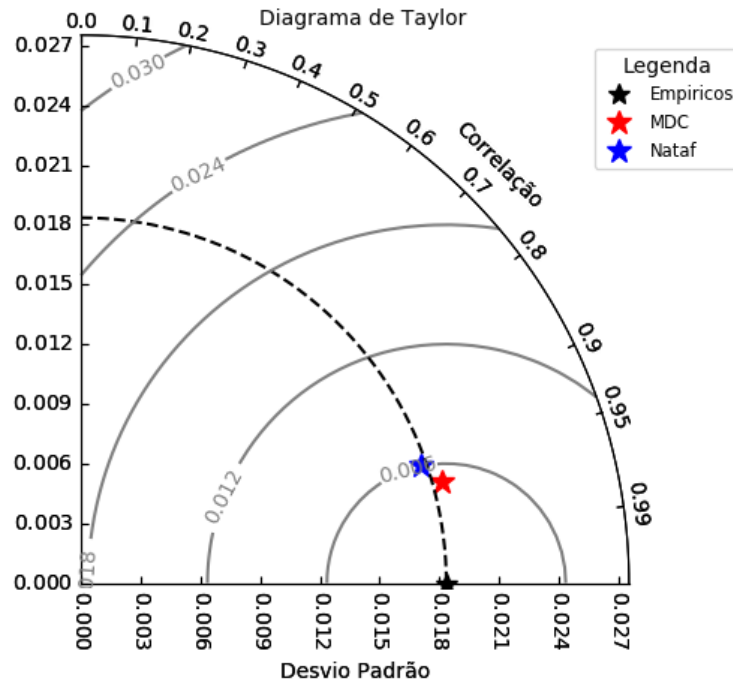


Figura 36- Diagrama de Taylor modelo condicional e Nataf - costa da Noruega (2 anos de dados)

3.5.2.3 Ondas de Noroeste – Resumo das análises

Para a amostra de dados de onda na direção Noroeste na costa da Noruega, contida na série histórica do banco de dados utilizado, verifica-se que ambas as metodologias de modelagem de distribuição conjunta investigadas neste trabalho apresentam um desempenho similar. Além disto, neste caso o modelo condicional se mostrou pouco sensível ao tamanho da base de dados. Este desempenho do modelo baseado na distribuição condicional pode ser atribuído ao uso de funções paramétricas para representação dos parâmetros da distribuição condicional já desenvolvidos para a região do mar do Norte por MATHISEN e BITNER-GREGERSEN (1990).

4 Conclusões e Sugestões para Trabalhos Futuros

Neste trabalho foram investigados dois modelos para obtenção da distribuição conjunta dos parâmetros de onda: altura significativa H_s e período de pico T_p . Esta distribuição é muito importante para análise e projeto de estruturas marítimas.

Os dois modelos investigados foram: a) um modelo tradicional baseado no uso de distribuições condicionais de probabilidade e b) um modelo que usa uma transformação probabilística conhecida como transformada de Nataf. Conceitualmente, o primeiro modelo seria o modelo ideal, porém, ele apresenta uma grande dificuldade para ser usado na prática porque requer o uso de funções contínuas para representar os parâmetros da distribuição condicional em função dos valores da variável independente. Neste trabalho foram utilizadas as equações paramétricas sugeridas em MATHISEN e BITNER-GREGERSEN (1990). Já o segundo modelo é de bastante apelo prático, pois o mesmo requer apenas as distribuições marginais das variáveis envolvidas e o valor do coeficiente de correlação entre elas. Porém, conceitualmente, a correlação pode algumas vezes não representar apropriadamente a dependência estatística entre os dados.

Utilizando-se um conjunto de 11 anos de dados históricos simultâneos de H_s e T_p gerados através do modelo de ondas WAVEWATCH III, os dois modelos descritos anteriormente foram utilizados para gerar correspondentes distribuições conjuntas em duas localidades: costa brasileira (Bacia de Santos) e costa da Noruega (Tromsøflaket). Para as duas localidades verificou-se que a distribuição de Weibull foi a mais apropriada para modelar a distribuição marginal de H_s e a distribuição Lognormal foi a mais apropriada para modelar tanto a distribuição marginal quanto a condicional de T_p . Baseando-se em observações visuais, comparação de parâmetros estatísticos e no diagrama de Taylor foi possível observar o seguinte:

- o modelo de Nataf apresentou uma boa performance para ambas localidades mesmo quando uma amostra de tamanho inferior (2 anos de dados) foi usada na modelagem;
- o modelo baseado na distribuição condicional apresentou certas dificuldades em seu ajuste quando usado na modelagem dos dados da costa brasileira, porém,

a sua performance foi boa quando foram tratados os dados da costa da Noruega (independentemente do tamanho da amostra).

Pela facilidade de modelagem e pelos resultados até aqui obtidos seria aconselhável o uso do modelo de Nataf para representação da distribuição conjunta de H_s e T_p . Entretanto, como foram poucos os casos testados, sugere-se que este estudo continue através da consideração de outras localidades para obter-se uma conclusão definitiva sobre este tópico. Além de outras localidades seria interessante que em estudos futuros fossem também abordados outros tópicos:

1. Busca de outras equações paramétricas para representação dos parâmetros da distribuição condicional;
2. Investigação sobre o efeito do uso dos dois modelos de distribuição conjunta na análise de fadiga e de valores extremos de estruturas marítimas;
3. Comparações entre os contornos ambientais de projeto (WINTERSTEIN *et al.*, 1993);
4. Incluir nos modelos as incertezas nos parâmetros estatísticos em função do tamanho da amostra de dados.

5 Referências Bibliográficas

- ANG, A.H.S., TANG, W.H., 1975, *Probability Concepts in Engineering Planning and Design. Volume I, Basic Principles*. New York, John Wiley & Sons.
- ANG, A.H.S., TANG, W.H., 1984, *Probability concepts in engineering planning and design. Volume II – decision, risk and reliability*. New York, John Wiley & Sons.
- ATLASUL, 2018, *Site do ATLASUL – Sistema de previsão de ondas e monitoramento costeiro*. Disponível em: <<http://www.lamma.ufrj.br/sites/spo/atlasul/scripts/taylor/diagrama.php>>. Acesso em: 23 de outubro de 2018.
- BITNER-GREGERSEN, E.M., 2015, “Joint met-ocean description for design and operations of marine structures”, **Applied Ocean Research**, v. 51, n. 1, pp. 279-292.
- BITNER-GREGERSEN, E.M., HAVER, S., 1991, “Joint environmental model for reliability calculations”. In: **Proceedings of the First International Offshore and Polar Engineering Conference**, pp. 246-253, Edinburgo, Ago. 1991.
- DER KIUREGHIAN, A., LIU, P.L., 1986, “Structural reliability under incomplete probability Information”, **Journal of Engineering Mechanics**, v. 112, n. 1, pp. 85-104.
- DITLEVSEN, O., MADSEN, H., 1996, *Structural Reliability Methods*. Chichester, Wiley.
- HOLTHUIJSEN, L., 2007, *Waves in oceanic and coastal waters*. Cambridge, Cambridge University Press.
- MATHISEN, J., BITNER-GREGERSEN, E., 1990, “Joint distributions for significant wave height and wave zero-up-crossing period”, **Applied Ocean Research**, v. 12, n. 2, pp. 93-103.
- MONSALVE-GIRALDO, J.S., DANTAS, C.M.S., SAGRILO, L.V.S., 2016, “Probabilistic fatigue analysis of marine structures using the univariate dimension-reduction method”, **Marine Structures**, v. 50, n. 1, pp. 189-204.

- NATAF, A., 1962, "Détermination des distribution dont les marges sont données", **Comptes rendus hebdomadaires des séances de l'Académie des sciences**, v. 255, n. 1, pp. 42-43.
- PETROBRAS, 2018, *Site da Petrobras*. Disponível em: <<http://www.petrobras.com.br/pt/nossas-atividades/principaisoperacoes/bacias/bacia-desantos.htm>>. Acesso em: 8 de agosto de 2018.
- SAGRILO, L.V., PRATES DE LIMA, E., PAPAEO, A., 2011, "A Joint probability model for environmental parameters", **Journal of Offshore Mechanics and Arctic Engineering**, v. 133, n. 3, pp. 1-7.
- SILVA-GONZÁLEZ, F., HEREDIA-ZAVONI, E., MONTES-ITURRIZAGA, R., 2013, "Development of environmental contours using Nataf distribution model", **Ocean Engineering**, v. 58, n. 1, pp. 27-34.
- TAYLOR, K.E., 2001, "Summarizing multiple aspects of model performance in a single diagram", **Journal of Geophysical Research**, v. 106, n. D7, pp. 7183-7192.
- TOLMAN, H.L., 1989, *The numerical model WAVEWATCH: a third generation model for the hindcasting of wind waves on tides in shelf seas*. In: Report nº. 89-2, Communications on Hydraulic and Geotechnical Engineering, Delft University of Technology, 72 pp.
- TOLMAN, H.L., 1991, "A third-generation model for wind waves on slowly varying, unsteady and inhomogeneous depths and currents", **Journal of Physical Oceanography**, v. 21, n. 1, p. 782-797.
- TOLMAN, H.L., 1992, "Effects of numerics on the physics in a third-generation wind-wave model", **Journal of Physical Oceanography**, v. 22, n. 1, p. 1095-1111.
- WINTERSTEIN, S.R., UDE, T.C., CORNELL, C.A., BJERAGER, P., HAVER, S., 1993, "Environmental parameters for extreme response: inverse form with omission factors". In: **Proceedings of ICOSSAR 93**, pp. 1-8, Innsbruck, Ago. 1993.