



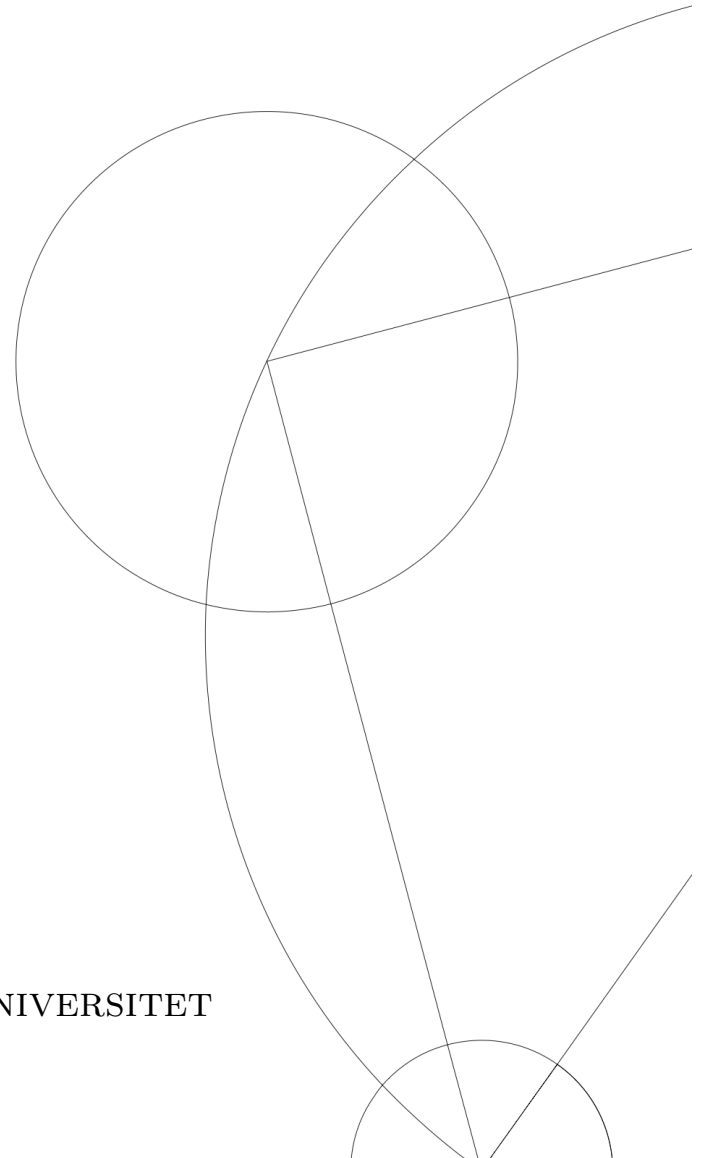
---

# THE ROLE OF WIND PARAMETERS IN ROGUE WAVE PROBABILITY PREDICTIONS

BSC. THESIS

Skrevet af *Yulii Bardenshtein*  
August 20, 2021

Under vejledning af  
Markus Jochum



KØBENHAVNS UNIVERSITET



UNIVERSITY OF  
COPENHAGEN

NAME OF INSTITUTE: University of Copenhagen

NAME OF DEPARTMENT: The Niels Bohr Institute

AUTHOR(S): Yulii Bardenshtein

EMAIL: jhw783@alumni.ku.dk

TITLE AND SUBTITLE: The role of wind parameters in Rogue Wave probability  
predictions  
-

SUPERVISOR(S): Markus Jochum

HANDED IN: 20.08.21

DEFENDED: 02.09.21

NAME Yulii Bardenshtein

SIGNATURE 

DATE 20-08-2021

# Contents

<b>1 Abstract</b>	<b>4</b>
<b>2 Introduction</b>	<b>5</b>
<b>3 Theory</b>	<b>5</b>
3.1 Rogue waves . . . . .	5
3.2 Wind waves . . . . .	6
<b>4 Methods</b>	<b>7</b>
4.1 Wind parameters . . . . .	7
4.2 Inclusion of wind data to FOWD . . . . .	8
4.3 Univariate analysis . . . . .	9
<b>5 Results</b>	<b>10</b>
5.1 Rogue Waves . . . . .	10
5.1.1 Stratification on crest-trough correlation . . . . .	11
5.1.2 Stratification on steepness . . . . .	12
5.2 Rogue Crests . . . . .	13
5.2.1 Stratification on steepness . . . . .	14
<b>6 Discussion</b>	<b>15</b>
<b>7 Conclusion</b>	<b>16</b>
<b>8 Additional information</b>	<b>17</b>
<b>9 Acknowledgments</b>	<b>17</b>
<b>Appendices</b>	<b>18</b>

# 1 Abstract

The study of rogue waves is essential to the safety and navigation of naval vessels [1]. Recent findings have discovered a new pathway in understanding the causes of rogue waves [2]; however, the effects of wind have yet to be established. This thesis investigates the impact of wind on rogue wave probabilities. A univariate Bayesian analysis is performed on a combined dataset of wind data from thirteen weather buoys and fifteen Free Ocean Water Dataset buoys. The results from the analysis indicate that wind parameters are poor predictors of rogue wave probabilities. However, the results suggest that the most critical wind parameter, namely inverse wave age, mainly influences the rogue wave probabilities due to its correlation with steepness. Another important finding from the analysis is the insignificant effect of the angular component of wind parameters, suggesting that the creation of rogue waves is independent of the increase or decrease of wave energy by wind.

The analysis of rogue crests reveals that wind parameters are good predictors of rogue crest probabilities. Inverse wave age is found to be the most critical wind parameter, with a lower bound predictive power of 0.56. The wind is found to cause rogue crests independently of steepness, revealing a direct pathway for wind-wave interactions. Wind parameters, especially inverse wave age, will need to be appreciated in further studies and forecasting algorithms regarding rogue waves.

## 2 Introduction

This thesis adds a study of the role of wind parameters to the **Free Ocean Water Dataset** (FOWD) [3] analysis carried out in the recent work: *Real World Rogue Wave probabilities* by Häfner et al. (2021) [2]. The analysis from the paper investigated the importance of individual sea-state parameters for rogue wave probabilities. It concluded that most rogue waves are caused by linear effects, while second-order nonlinearities provide a minor correction. Furthermore, the study provided evidence that parameters such as steepness are one of the strongest predictors of rogue crest probabilities. Local Winds are hypothesised to play an indirect role in the analysis of rogue waves due to the correlation with lower wave periods and high-frequency seas. Steepness is observed to be more important in these sea-states. Furthermore, evidence suggests that wind can increase steepness [4], which offers a hypothetical pathway for the wind to play a role in the analysis of rogue waves.

It has also been argued that wind does not directly lead to the formation of rogue waves but mainly serves as an amplifier for larger waves to become rogue waves [1]. Thereby wind increases rogue wave activity. However, rogue wave analyses that take wind into account are limited by statistically stationary sea states and lack of actual ocean data. The FOWD sample used in the analysis allows the bypassing of the above-mentioned limitations.

The primary source of wind data comes from the **National Data Buoy Center** (NDBC) [5]. The combined dataset consists of fifteen buoys with wind sensors near FOWD buoys with aggregated versions of FOWD samples. The FOWD sample features buoys from different locations of the east and west coasts of the United States.

An essential aspect of the analysis is the incorporation of wind-wave parameters. These parameters allow for a detailed investigation, which can be tied to physical effects. These parameters include inverse wave age, saturation wave height and wind-wave misalignment.

## 3 Theory

### 3.1 Rogue waves

Rogue waves are characterized as waves with an abnormality index above 2. The abnormality index is defined as the fraction between the individual wave height and the significant wave height

$$AI = \frac{H_{wave}}{H_s}. \quad (1)$$

The significant wave height is defined as

$$H_s = 4\sqrt{m_0}, \quad (2)$$

where  $m_0$  is the zeroth moment of the spectral density as a function of the frequency

$$m_n = \int_0^{\infty} f^n S(f) df. \quad (3)$$

In deep water the probability distribution of the ratio between wave height and significant wave height follow a Rayleigh distribution [6]:

$$P(AI) = \exp\left(-\frac{AI^2}{0.5}\right). \quad (4)$$

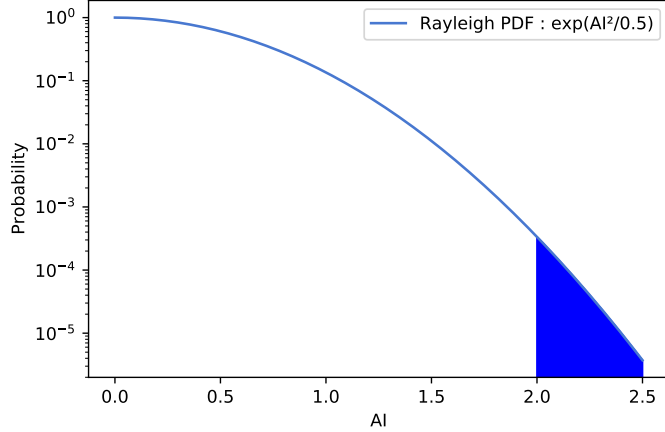


Figure 1: Rayleigh PDF for wave abnormality index - AI. The Rogue wave domain of the function ( $AI > 2$ ) is highlighted.

Linear superposition creates most rogue waves [2]. Crest-trough correlation, a linear parameter, was discovered as the highest predicting parameter for rogue waves. Narrowband seas have a higher crest-trough correlation due to a limited frequency range within the spectrum. Narrowband seas are associated with swells. Rogue crests are defined as wave crests with a crest-abnormality index above 1.2:

$$CAI = \frac{\eta}{H_s}, \quad (5)$$

where  $\eta$  is the crest-height. Second-order nonlinearities cause rogue crests. The best parameters for rogue crest probability prediction are surface elevation skewness, steepness and Ursell number. Surface elevation skewness is a measure of the vertical asymmetry between crests and troughs. Steepness is the ratio between wave height and wavelength. Ursell number is a measure of nonlinearity and is calculated from steepness and the relative depth [6].

### 3.2 Wind waves

Rogue waves have their origin as wind waves. Wind waves are created due to the presence of constant wind. The area in which the wind blows over the water is called fetch. The magnitude of the resulting waves is determined by the wind speed, duration, water depth and fetch length [6]. The Miles-Phillips mechanism describes the formation and growth of wind waves. The mechanism is a combination of the Miles and Phillips mechanisms. The Phillips mechanism describes the creation of wind waves, and the Miles mechanism explains their growth [7],[8]. Initially, wind creates pressure fluctuations on a flat water surface. The interaction causes irregularities on the flat surface, which in turn creates small capillary waves [8]. The restoring force of surface tension causes the waves to propagate. The initial energy transfer rate is linear. As the waves grow, the harmonic nature of the generated waves causes a disturbance of the airflow along the surface. This increases the number of interactions with the wind and induces exponential wave growth. The mechanism is a positive feedback mechanism [6]. The expression below is the source term for wave generation by constant wind and based on the formulation by Snyder et al. (1981) [6]

$$S(f, \theta) = \alpha + \epsilon_2 \frac{\rho_{air}}{\rho_{water}} \left[ 28 \frac{U}{c_p} \cos(\theta - \theta_{wind}) - 1 \right] E(f, \theta), \quad (6)$$

where  $U$  is the wind speed,  $f$  is the frequency,  $\rho_{water}$  and  $\rho_{wind}$  being the respective densities of water and wind.  $\theta$  and  $\theta_{wind}$  being the respective directions of wind and waves, respectively.  $\epsilon_2$  is a tunable coefficient. The constant  $\alpha$  is the initial linear growth derived from the Phillips mechanism.  $E(f, \theta)$  is energy density, which increases exponentially. Equation 6 implies that the wind will transfer energy to the waves as long as they are travelling in the same direction. If the winds are travelling in opposite directions the source term

is negative, the energy transfer from the winds to the waves is negative [6].

Nonlinear interactions result in energy exchange between waves, which creates waves with longer wavelengths that propagate faster than the wind. The wind will then reduce its energy transfer to the waves. These waves are called swells, and they will disperse in different directions away from the fetch [6]. Most rogue wave occurrences are present in swell conditions [2].

## 4 Methods

### 4.1 Wind parameters

Wind speeds and directions are not enough for the analysis since they do not quantify the interactions between the wind and its surrounding sea-state and cannot infer a physical mechanism behind the effects of the wind.

The analysis does not include wind gusts. Wind gust is a temporary increase in wind speed over several seconds to a minute. These periods are insignificant compared to those of constant wind speeds. Gust serves as a proxy for wind speed and would attribute the effects on rogue wave probabilities to wind speed.

Wind-wave parameters are essential to the analysis due to their physical meaning. The list of wind-wave parameters includes inverse wave age, saturation wave height, and wind-wave misalignment. The inclusion of inverse wave age is chosen due to its relation with the Miles-Phillips mechanism since the growth rate is dependent on the ratio between the wind speed and the peak phase velocity of the wave. The parameter was found to be related to the sea-state, where values below 0.15 indicate swell conditions and values above 0.83 indicate dominant wind-seas [9]. Regular wave age serves as an indication for the interaction time between wind and the water surface. Inverse wave age without the angular component is defined as

$$\beta = \frac{U_{10}}{c_p}, \quad (7)$$

Where  $U_{10}$  is the wind speed at a 10-meter height, and  $c_p$  is the peak phase velocity:

$$c_p = \sqrt{\frac{g}{k} \tanh(kh)}, \quad (8)$$

$$k = \frac{2\pi}{\lambda}, \quad (9)$$

where  $\lambda$  is the peak wavelength, and  $k$  is the wavenumber. Inverse wave age with the angular component is included in determining whether the parameter is tied to positive or negative energy transfer. The parameter serves as a comparison with the inverse wave age described above. Inverse wave age with the angular component is defined as

$$\beta_{ang} = \frac{U_{10} \cos(\theta_{wind} - \theta_{wave})}{c_p}. \quad (10)$$

Throughout this thesis, inverse wave age and the inverse wave age with the angular component will be referred to  $\beta$  and  $\beta_{ang}$ , respectively.

Saturation wave height serves as a proxy for the ratio between wind and wave energy. Saturation wave height and inverse wave age refer to the same mechanism; however, saturation wave height scales differently. Saturation wave height is defined as

$$\frac{U_{10}^2 \alpha}{H_s g}, \quad (11)$$

where  $g$  is the gravitational acceleration, and  $\alpha = 0.24$ . Saturation wave height with the angular component is included in the analysis and defined as

$$\frac{(U_{10} \cos(\theta_{wind} - \theta_{wave}))^2 \alpha}{H_s g}, \quad (12)$$

Wind-wave misalignment is included in determining if the relative angle between wind direction and wave direction is more important than the strength of the wind. Wind wave misalignment is defined as

$$\cos(\theta_{wind} - \theta_{wave}). \quad (13)$$

## 4.2 Inclusion of wind data to FOWD

The analysis includes weather buoys with anemometers from the **National Data Buoy Center** (NDBC), which are operated by the **National Oceanic and Atmospheric Administration** (NOAA). ERA5 winds [10] were an option initially, but due to the easy procedure of performing quality control on NDBC data and the easier data handling, the option settled on data from NDBC. Each weather buoy is directly accessible through their station map, which can help determining the geographical surroundings of the wind sensor. Weather buoys are chosen for the analysis since the conversion of wind speeds from different heights to a specific height requires specific knowledge of the surroundings.

Based on the maximum resolution from the ERA5 dataset, the distance between the wind sensor and its corresponding buoy has to be less than 30 km. The distance constraint is set to ensure a clear established connection between the local wind and its corresponding sea-state.

A combination of continuous winds and standard meteorological data is used for the analysis. The reason is that parameters of continuous winds are measured more frequent than standard meteorological data. However there are more datasets with standard meteorological data. The continuous wind datasets have a data acquisition period of 10 minutes, and standard meteorological data have an 8-minute acquisition window. The indicated values of wind speed and wind direction are average measurements within each acquisition window. The indicated timestamp in the NDBC datasets is the acquisition time of the data; therefore the datasets are merged on the FOWD equivalent of the aggregated datasets. An arbitrary tolerance limit of 5 minutes between the NDBC acquisition time and the FOWD acquisition time is set. The tolerance is set to ensure a decent accuracy between wind and FOWD measurement and include enough data to make the analysis possible.

The combined dataset includes data from thirteen weather buoys and fifteen FOWD buoys, that are located of the west and east coasts in the United States.

NDBC sensor id:	FOWD buoy id:	Distance [km]	Anemometer Height [m]
46099	036	28.4	4.5
46053	107	10.8	3.2
46022	128	18.0	3.8
46026	142	18.6	3.8
CHLV2	147	1.37	43.3
46092	156	7.2	4.0
46096	162	4.8	3.0
46022	168	21.8	3.8
44014	171	0.24	3.2
46029	179	13.5	3.2
46026	180	0.54	3.8
46042	185	10.6	3.8
41009	227	1.69	4.1
46025	234	7.05	4.1
46014	235	16.4	3.8

Table 1: Table of NDBC sensors and their corresponding FOWD buoys. Distance between anemometers and FOWD buoys in kilometres and the anemometer height above sea level are included

Emphasis is placed on consistency: all the anemometer heights are adjusted to a standard height of 10 meters using the logarithmic wind profile [11].

$$u(z_{10}) = u(z_{meas}) \frac{\ln((z_{10} - d)/z_0)}{\ln((z_{meas} - d)/z_0)}, \quad (14)$$

where  $z_{10}$  and  $z_{meas}$ , are the 10 meter height and the anemometer height, respectively.  $u(10)$  and  $u(z_{meas})$  are the wind speeds at their respective heights.  $z_0$  is the surface roughness length (meters), in this analysis



it is set to 0.0002, which is the value typically used for open waters. The zero-plane displacement  $d$  is set to zero as the sensors are located in the ocean.

### 4.3 Univariate analysis

The univariate analysis [2] is the primary method of determining the importance of individual wind parameters regarding rogue wave probabilities. The analysis is based on Bayesian inference and the univariate binning. Bayesian inference incorporates a posterior probability, which is derived from a probability distribution and its conjugate prior.

Rogue wave occurrences are given by the binomial distribution, with the probability  $p$  [3]

$$n^+ = \text{Binom}(n^+, n^-, p), \quad (15)$$

where  $n^+$  and  $n^-$  being rogue wave observations and non-rogue wave observation, respectively. Bayes theorem is applied to the binomial distribution and the conjugate prior of the binomial distribution, which is a Beta distribution

$$p_{\text{prior}} = \text{Beta}(\alpha_0, \beta_0) \quad (16)$$

The resulting predictive posterior distribution is a Beta distribution

$$P(p|n^+, n^-) = \text{Beta}(n^+ + \alpha_0, n^- + \beta_0) \quad (17)$$

The probability of encountering a rogue wave for every aggregated wave chunk of 100 waves is corrected

$$p = 1 - (1 - p_{100})^{\frac{1}{100}} \quad (18)$$

The univariate binning involves separating all relative wave height observations into equal-sized bins for each parameter. Bins with less than ten rogue wave observations are removed, and the probability  $p$  is assumed to be an identically and independently distributed random variable within each bin. The essential idea behind this analysis is converting binary labels, e.g. non-rogue waves and rogue waves, into probabilities. The two classes are inseparable in parameter values as rogue waves appear in sea states containing regular waves. Therefore, the analysis hinges on determining the importance of a parameter with its relation to rogue wave probabilities. Since rogue waves are scarce, a constraint is set by incorporating a prior beta distribution with the parameters  $\alpha_0$  and  $\beta_0$ . They represent the total number of waves for which either one rogue wave or rogue crest occurs.

$\alpha_0$	$\beta_0$	
1	$10^4$	AI >2.0
1	$10^4$	CAI >1.2
1	$10^6$	AI >2.4
1	$10^6$	CAI >1.5

Table 2: The table shows the values of  $\alpha_0$  and  $\beta_0$ , with their respective abnormality index.

The univariate predictive power determines the ability of an individual parameter to predict rogue waves. The univariate predictive power is defined as

$$P_x = \log_{10} \left( \frac{p_{i_{max}}}{p_{i_{min}}} \right) \quad (19)$$

$p_{i_{max}}$  is the highest lower bound probability and  $p_{i_{min}}$  is the lowest upper bound probability, both for  $i$ 'th bin respectively.

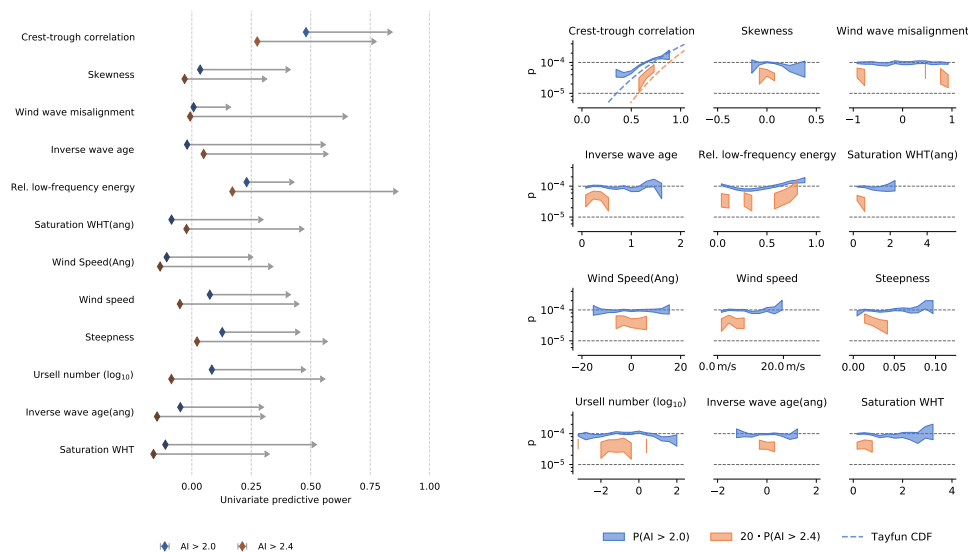
Plots of scaled rogue wave probabilities for individual parameters are included in the analysis. The parameter values are within a 95% credible interval. They are included since they provide an observable effect of each parameter on rogue wave probabilities.

A conditional independence analysis is performed via stratification on crest-trough correlation and steepness. This is to exclude causal paths related to sea-state parameters and observe the direct effect of wind parameters on rogue wave probabilities.

Stratification on Ursell number is not included in the analysis. Ursell Number is governed by steepness and the relative depth, where relative depth is related to the topology [2]. Stratification on Ursell Number would therefore not make much sense since the direct effects of wind cannot be isolated.

## 5 Results

### 5.1 Rogue Waves



The 2.5 percentile lower bound predictive power for each parameter. Arrows indicate the 97.5 percentile.

Scaling of Rogue Wave probability for each parameter within a 95% credible interval. The probability for waves with  $AI > 2.4$  is scaled by 20.

Figure 2

The predictive power and probability scaling of crest-trough correlation match its counterpart from the primary FOWD analysis. Thus, indicating that this FOWD sample is representative of the primary dataset. The predictive power in figure 2 reveals that all wind parameters are poor predictors of Rogue wave probabilities for  $AI > 2.0$  and  $AI > 2.4$ . Furthermore, all of them have predictive powers below 0.20.

The probability scalings shown in figure 2 demonstrate that most of the wind parameters have no changes in probability. Where most parameters have constant probabilities around  $10^{-4}$ . A small statistically insignificant slope appears for wind speed  $17m/s$  and  $j\beta_{ang}j = 1$ .  $\beta$  and  $\beta_{ang}$  are directly associated with sea state conditions. Lower values of  $\beta_{ang}$  indicate sea states dominated by swell conditions, which have a naturally larger rogue wave probability due to their bandwidth effects [2].  $j\beta_{ang}j = 1$  represents the regime where wind interacts strongly with waves via the Miles-Phillips mechanism. Dominant wind-wave interactions could potentially lead to a slight excess of rogue waves. The probability of wind-wave misalignment does not change throughout the entire range of parameter values, suggesting that wind-wave interactions' strength dominates over the relative direction between wind and waves.

### 5.1.1 Stratification on crest-trough correlation

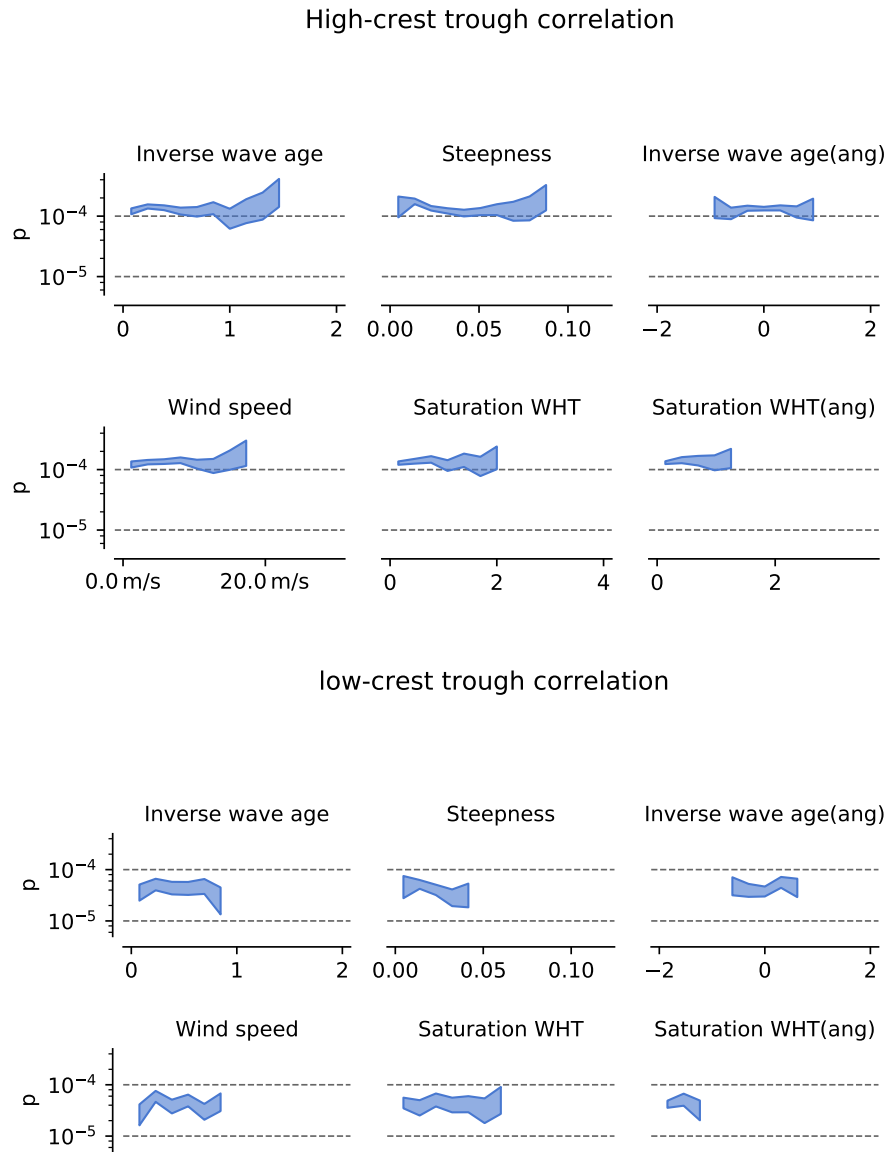
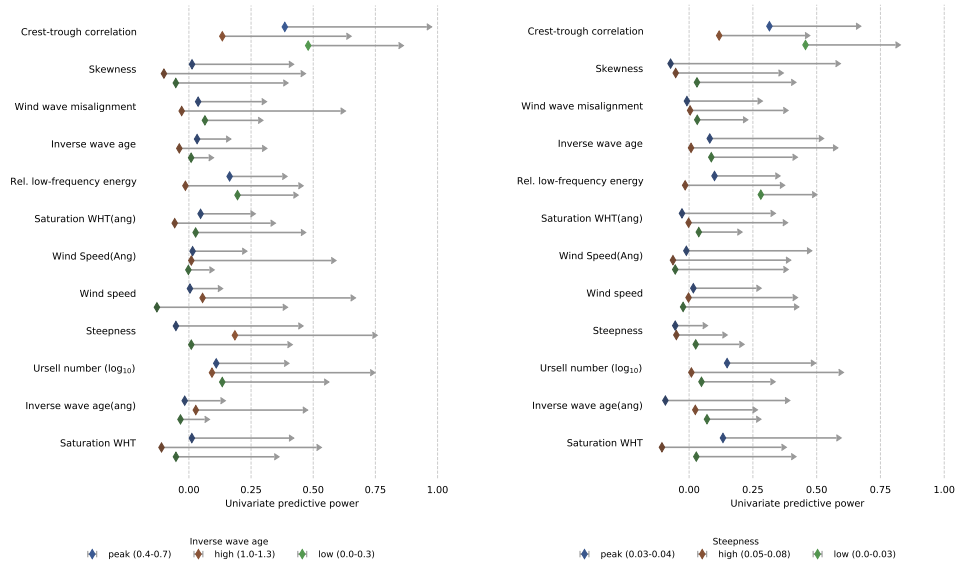


Figure 3: Scaling of rogue wave probability with regards to sample stratification on crest-trough correlation

Sample stratification on crest-trough correlation shown in figure 3 reveals that for high crest-trough correlation, there is a small and statistically insignificant rogue wave probability increase for  $\beta = 1$ . However, the large uncertainty for  $\beta = 1$  indicates the lack of data to make a conclusive statement.

### 5.1.2 Stratification on steepness



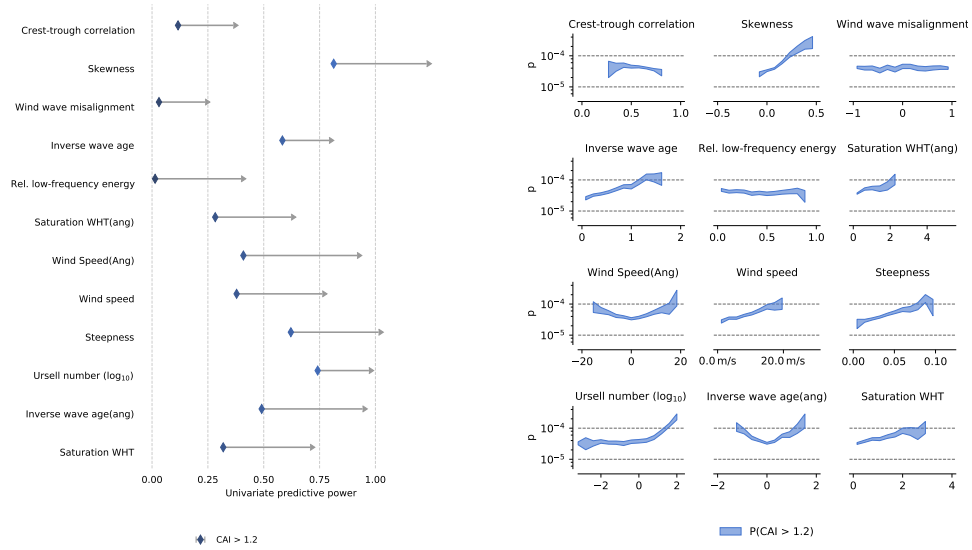
The 2.5 percentile lower bound predictive power for values stratified on  $\beta$  for rogue waves with  $AI > 2.0$ .

The 2.5 percentile lower bound predictive power for values stratified on steepness for rogue waves with  $AI > 2.0$ .

Figure 4: Steepness becomes slightly more informative when stratified on  $\beta$

Figure 4, reveals that the wind could potentially act through steepness for rogue waves. However, the results only apply for  $\beta$  between 1.0 and 1.3, which tells that the effect is statistically insignificant. More data is needed to disentangle the interaction between inverse wave age and steepness properly. Due to the correlation between inverse wave age and steepness in figure 9 and the results from figure 4, the relation between inverse wave age and steepness serves as an essential starting point for further investigation.

## 5.2 Rogue Crests



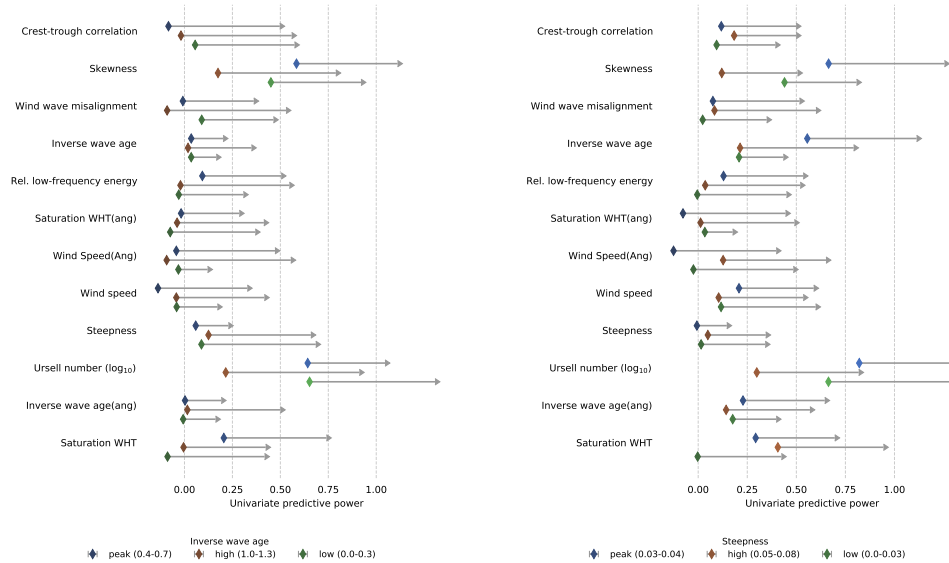
The 2.5 percentile lower bound predictive power for each parameter. Arrows indicate the 97.5 percentile.

Scaling of Rogue Crest probability for each parameter within a 95% credible interval.

Figure 5

Predictive powers from figure 5 reveal that all wind parameters perform better on rogue crests, where inverse wave age( $\beta$ ) is the most critical wind parameter, with a lower bound predictive power of 0.56. The poor performance of wind-wave misalignment further establishes the effect of the wind strength rather than the relative angle between wind and wave directions. The more substantial predictive power of  $\beta$  compared to  $\beta_{ang}$  and the symmetry around 0 for the scaling of  $\beta_{ang}$ , reveals that the angular component is not essential for rogue crests. Opposing winds take energy from waves, but the energy of the waves is related to the significant wave height, which does not influence the creation of outliers [2].

### 5.2.1 Stratification on steepness



The 2.5 percentile lower bound predictive power for values stratified on  $\beta$  for rogue crests with  $CAI > 1.2$ .

The 2.5 percentile lower bound predictive power for values stratified on steepness for rogue crests with  $CAI > 1.2$ .

Figure 6: The results suggest that for crests  $\beta$  is still relevant when stratified on steepness.

The plots from figure 6 reveal that the predictive powers of steepness are below 0.20 when stratified on  $\beta$ , but not vice versa. Conditional stratification on steepness removes the causal path of steepness, whereby direct wind-wave interactions are investigated. The results from figure 4 reveal that the wind can cause rogue crests independently of steepness. The supposed mechanism behind this interaction has yet to be revealed and serves as a starting point for future studies. Results from figure 7 show that there is a statistically insignificant increase in rogue crest probability for negative values of  $\beta_{ang}$ , which further establishes that rogue wave occurrences are not related to the increase of wave energy.

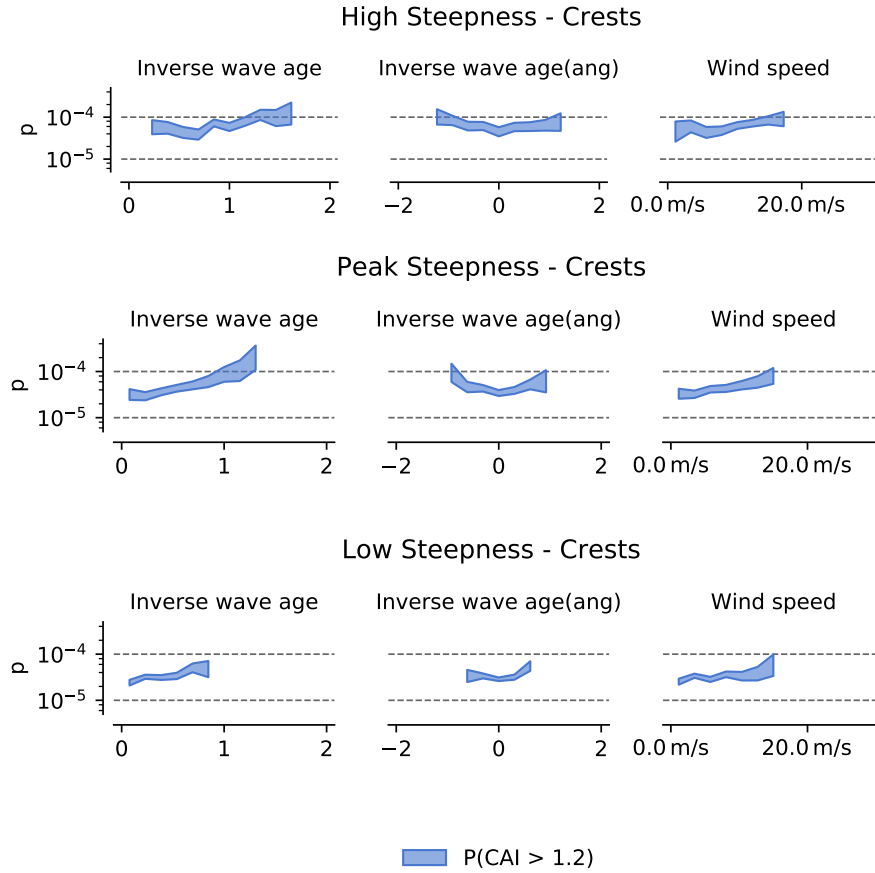


Figure 7: Scaling of rogue crest probability with regards to sample stratification on steepness.

## 6 Discussion

The analysis hinges on a well-combined dataset between meteorological and FOWD data. Therefore, the distance between NDBC buoys and FOWD buoys may seem like a biasing factor. The correlation between wind speed and significant wave height, were plotted against the distance between FOWD buoys and their corresponding weather buoys. Figure 9 reveals that there is an overall decrease in the correlation between wind speed and significant wave height with increasing distance. This could have impacted the results and therefore serves as a potential flaw within the analysis. Therefore, future studies are encouraged to anemometers within a range of 15 kilometres from the nearest FOWD buoy.

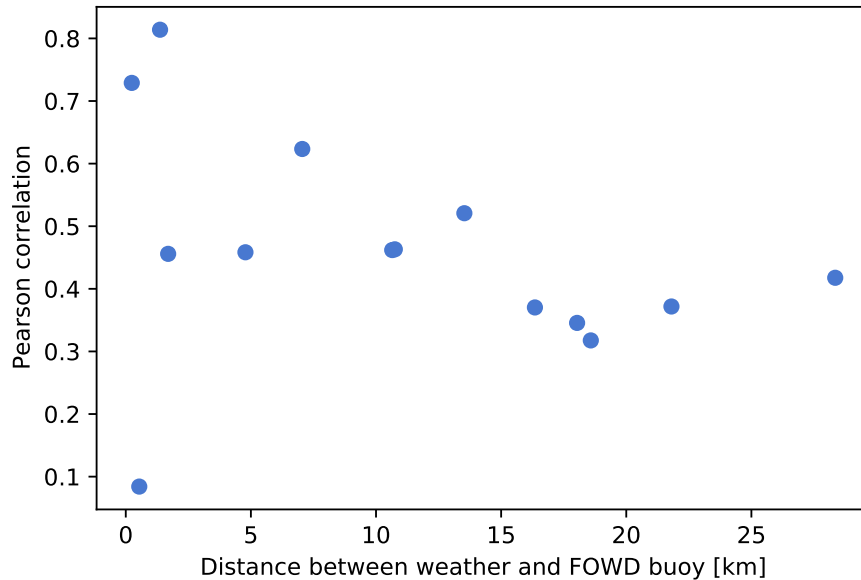


Figure 8: Scatter plot of the Pearson correlation coefficient between wind speed and significant wave height from the individual datasets, plotted against the distance between FOWD and weather buoys.

Inland sensors were not used, which excludes a large amount of data from the analysis. Data from inland sensors is gathered every two minutes, which is more accurate than the eight or ten-minute gathering from the weather buoys. The height environment between in-land sensors varied, and it was difficult to deduce the proper zero-plane displacement when converting the wind speed to a 10-meter height. However, if the right conditions around the anemometers are known regarding the zero-plane displacement and the roughness length. In that case, further studies are encouraged to incorporate inland sensors as they provide more data for analyses.

Since this analysis was performed on a small FOWD sample, the findings might come across as unreliable. The sample consists of 6384 rogue waves, potentially indicating a significant sampling error which casts some doubt on the statistical significance of the results.

An issue that was brought by Häfner et al. (2021) is the classification of rogue waves based on an arbitrary threshold. This definition is inherently flawed since most rogue waves are found in calm swell conditions where most do not pose any significant danger to their surroundings. Wind parameters are poor predictors of these rogue waves. However, excluding wind from future analysis based on this definition of rogue waves might be detrimental. Wind can play an essential role in predicting rogue crests. The inverse wave age parameter should therefore be included in forecasting and future analyses of similar types.

Lastly, the analysis conducted in this thesis mainly focused on the univariate analysis. The multivariate algorithm from the primary FOWD analysis was excluded due to a lack of data to train the algorithm. Machine learning methods are helpful and can provide several advantages; however, it is harder to infer meaningful causal relations. Nevertheless, the univariate analysis delivered a good platform for conditional independence analysis, which helped untangle the effects of wind parameters.

## 7 Conclusion

Wind data from thirteen NDBC weather buoys were combined with aggregated data from a FOWD sample of fifteen buoys. The univariate analysis revealed that winds are inadequate predictors of rogue wave probabilities. This is mainly because most rogue waves are found in calm swell conditions, where strong wind-seas do not occur. Wind-wave misalignment and saturation wave height were unimportant for the analysis due to their inadequate predictive powers.



The wind parameter inverse wave could however play a role by working through steepness. Based on lack of data, this cannot be concluded upon. This result serves as a potential starting point in further investigation.

All wind parameters had better predictive powers for rogue crests, except for wind-wave misalignment. Specifically, inverse wave age was found to be the best predictor of rogue crest probabilities, with a predicting power of 0.56. Conditional stratification on steepness revealed that wind could independently cause rogue crests without interacting with steepness. Winds can come from multiple directions and impact waves, which suggests that the interaction is independent of the energy transfer from winds to waves. The findings from the rogue crest analysis serves as a stepping stone in understanding the effects of wind on rogue crests. Future analyses and forecasting algorithms for rogue waves are encouraged to factor in the wind due to its direct causal path for rogue crests.

## 8 Additional information

The data and the code used for the analysis can be accessed via. the github repository:

<https://github.com/YuliiBardenshtein/Inclusion-of-wind-to-the-FOWD-analysis>

Note: A github account is required to access the repository.

## 9 Acknowledgments

I would like to express my gratitude to Dion Häfner and Markus Jochum for the great supervision. The project has been a great learning experience, within the fields of statistics and ocean physics. An additional thank you is owed to Team Ocean at the Niels Bohr Institute and Johannes Gemmrich from the University of Victoria.

## References

- [1] A. A. Adcock & P. H. Taylor, *The physics of anomalous ('rogue') ocean waves*, doi : 10. 1088/0034-4885/77/10/105901, 14-10-2014
- [2] D. Häfner, J. Gemmrich & M. Jochum, *Real-world rogue wave probabilities*, <https://doi.org/10.1038/s41598-021-89359-1>, 12-05-2021
- [3] D. Häfner, J. Gemmrich & M. Jochum, *FOWD: A Free Ocean Wave Dataset for Data Mining and Machine Learning*, <https://arxiv.org/abs/2011.12071>, last revised 28-04-2021
- [4] J. H. Lee J. P. Monty, *On the Interaction between Wind Stress and Waves: Wave Growth and Statistical Properties of Large Waves*, <https://doi.org/10.1175/JPO-D-19-0112.1>, 01-02-2020
- [5] *National Data Buoy Center*, <https://www.ndbc.noaa.gov/>
- [6] L. H. Holthuijsen, *Waves in Oceanic and Coastal Waters*, <https://doi.org/10.1017/CB09780511618536>, 9780511618536, 2007
- [7] J. H. Miles, *On the generation of surface waves by shear flows*, <https://doi.org/10.1017/S0022112057000567>, 27-05-1957
- [8] O.M. Phillips, *On the generation of waves by turbulent wind*, <https://doi.org/10.1017/S0022112057000233>, 18-02-1957
- [9] K. Zheng, J. Sun, C. Guan & W. Shao, *Analysis of the Global Swell and Wind Sea Energy Distribution Using WAVE-WATCH III*, <https://doi.org/10.1155/2016/8419580>, 29-12-2015

- [10] ERA5 documentation,  
<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>
- [11] Log wind profile,  
[https://en.wikipedia.org/wiki/Log\\_wind\\_profile](https://en.wikipedia.org/wiki/Log_wind_profile)

## Appendices

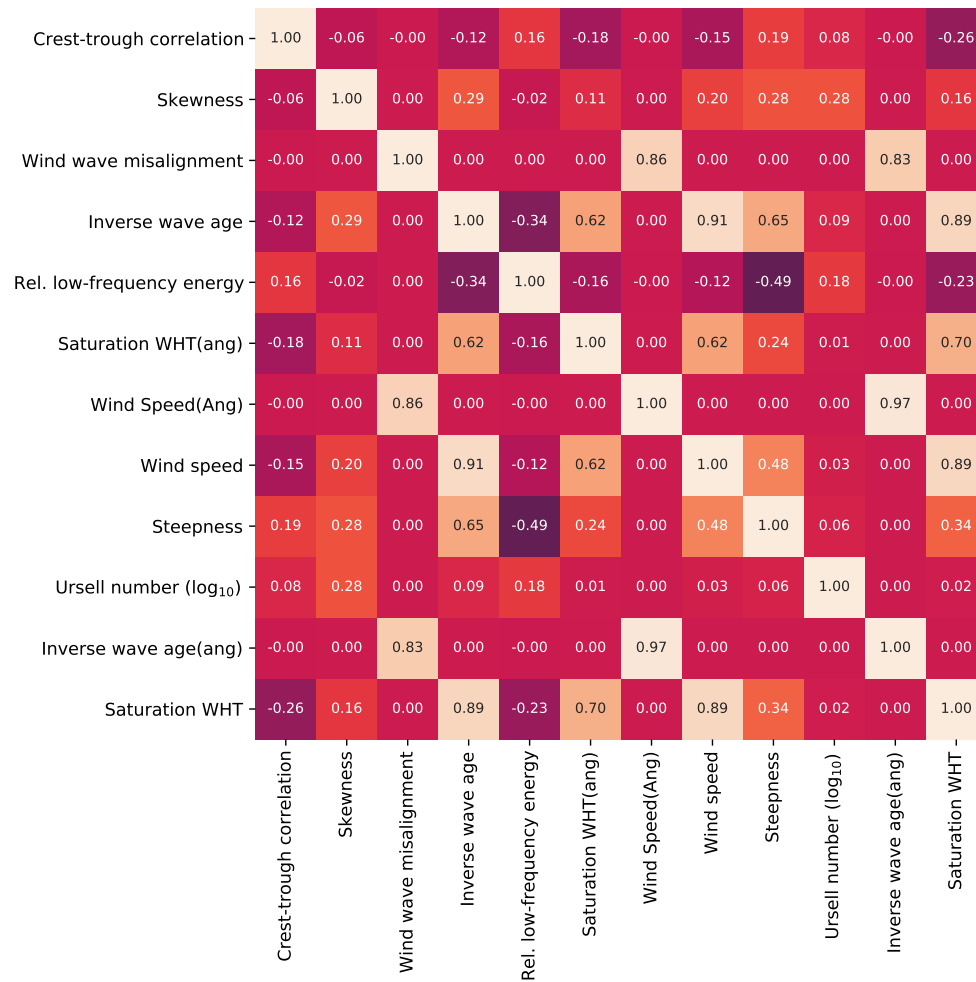


Figure 9: Matrix of Pearson correlations between parameters