A probabilistic prediction of rogue waves from a spectral WAVEWATCH III® wave model for the Northeast Pacific

by

Leah Cicon
B.Sc., University of Victoria, 2020

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the School of Earth and Ocean Sciences

A probabilistic prediction of rogue waves from a spectral WAVEWATCH III® wave model for the Northeast Pacific

by

Leah Cicon
B.Sc., University of Victoria, 2020

Supervisory Committee

Dr. Johannes Gemmrich, Co-Supervisor
(School of Earth and Ocean Sciences)

Dr. Jody M. Klymak, Co-Supervisor
(School of Earth and Ocean Sciences)

Dr. Brad Buckham, Outside Committee Member
(Department of Engineering)

ABSTRACT

Rogue waves are unexpected, individual ocean surface waves that are disproportionately large compared to the background sea state. They present considerable risk to mariners and offshore structures when encountered in large seas. Rogue waves have gone from seafarer's folktales to an actively researched and debated phenomenon. In this work an easily derived spectral parameter, as an indicator of rogue wave risk, is presented, and further evidence for the generation mechanism responsible for these abnormal waves is provided. With the additional goal of providing a practical rogue wave forecast, the ability of a standard wave model to predict the rogue wave probability is assessed. Current forecasts, like those at the European Centre for Medium-Range Weather Forecasts (ECMWF), rely on the Benjamin Feir Index (BFI) as a rogue wave predictor, which reflects the nonlinear process of modulation instability as the generation mechanism for rogue waves. However, this analysis finds BFI has little predictive power in the real ocean. From the analysis of long term sea surface elevation records in nearshore areas and hourly bulk statistics from open ocean and coastal buoys in the Northeast Pacific, crest-trough correlation shows the highest correlation with rogue wave probability. These results provide evidence in support of a probabilistic prediction of rogue waves based on random linear superposition and should replace forecasts based on modulation instability. Crest-trough correlation was then forecast by a regional WAVEWATCH III® wave model with moderate accuracy compared with the high performance of forecasting significant wave height. Results from a case study of a large fall storm October 21-22, 2021, are presented to show that the regional wave model produces accurate forecasts of significant wave height at high seas and presents a potential rogue wave probability forecast.

Table of Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I'm extremely grateful to my co-supervisor Dr. Johannes Gemmrich who presented me with this opportunity to join this project and has been hugely helpful, supportive and patient throughout. I am also thankful to my co-supervisor Dr. Jody M. Klymak and committee member Dr. Brad Buckham who with their expertise provided valuable feedback on my research along the way and on my final thesis. Special thanks to fellow graduate student Carmen Holmes-Smith for sharing this journey with me and for always being there with moral support and good humour. I would also like to extend my gratitude to the ECCC affiliates Benoit Pouliot who provided all the boundary conditions and wind input for the model and Dr. Natacha Bernier for her expertise in wave modelling. Finally, I gratefully acknowledge the student funding provided by the Search and Rescue New Initiatives Fund (SAR NIF).

1 Motivation

Between observations from buoys and open ocean platforms, numerical simulations based on theory and wave tank experiments, the underlying physics of rogue wave generation and the potential for predicting rogue waves is still up for debate. In this work, the aim is to further elucidate the underpinning physics of rogue wave occurrence and further identify key sea state parameters to be used for a practical rogue wave probability forecast from a standard forecast wave model in the Northeast Pacific (NEP).

Oceanic rogue waves are surface gravity waves that pose a great concern for mariners, offshore structures, and many users of the nearshore waters. Accounts of encounters with rogue waves include damage to marine structures and vessels and have resulted in injuries and the loss of lives. It is estimated that between 1969 and 1994 more than 22 supercarriers have been lost due to costly encounters with rogue waves (Kharif and Pelinovsky, 2003). Rogue waves however are not only an open seas phenomenon. A compiled list of extreme waves in 2005, included both offshore and nearshore events, where all of the reliably reported occurrences resulted in numerous injuries, damage to ships and coastal damage in the nearshore cases (Didenkulova et al., 2006). Unsurprisingly, rogue waves in high sea states receive the most attention. However, rogue waves occur in any height of seas, as they are waves with a wave height H or a crest height η that is disproportionately large compared to the background sea state represented by the significant wave height H_s. H is measured from the crest of the wave to the trough, η is measured from the crest to the mean sea level and H_s is the average of the third highest waves. Common criteria for rogue waves or rogue crests is $H > 2.2H_s$ and $\eta > 1.25H_s$, respectively. Seafarer's accounts of giant walls of water rising out of the seas were often dismissed as folktales, until the measurement of the Draupner wave ($\eta/H_s = 1.55$, H/H_s=2.15, H_s=11.9 m)

off an oil rig platform in the North Sea (Haver, 2004; Cavaleri et al., 2016). In the wake of the recording of the 25.6 m wave in 1996, research into these freak waves actively began. Since then there have been many other well-documented and-well researched rogue wave events in high sea states like the "Andrea" wave ($\eta/H_s = 1.62$, $H/H_s = 2.30$, $H_s = 9.2\,\mathrm{m}$) also recorded from an oil rig in the North Sea in 2007, and the "Killard" wave ($\eta/H_s = 1.57$, $H/H_s = 2.29$, $H_s = 11.4\,\mathrm{m}$) that struck a marine renewable energy site off the coast of Ireland (Donelan and Magnusson, 2017; Fedele et al., 2016). Ultimately it is these large waves that the research community wants to better understand and predict, as these are the waves that present the most risk. However, the same mechanism of generating these abnormally large waves is present in all sea states.

Research is ongoing to determine and understand the physics of rogue waves and the probability of their occurrence, and to determine in what sea state, if any, are they more likely to occur. The two main avenues of rogue wave generation in the absence of localized effects, like currents and seas with large directional spread, rely on either highly nonlinear effects or weakly nonlinear corrections to the Rayleigh wave height distribution obtained from linear wave theory. Current forecasts like those at the European Centre for Medium-Range Weather Forecasts (ECMWF) rely on the first mechanism, using the Benjamin Feir Index (BFI) as a rogue wave predictor, which is a reflection of the nonlinear process of modulation instability (Janssen, 2003; Janssen and Bidlot, 2009; Kirezci et al., 2021). These nonlinear interactions transfer energy within a wave group, leading to amplification of one or more waves within that group leading to the generation of a particularly large wave. This phenomenon requires a wave field that has narrow directional spread as well as a narrow peak in frequency. So, while there is plenty of experimental evidence for BFI in wave tanks and simulations, the relative importance of this process in the real ocean is unverified.

The other mechanism relies on simple linear superposition, where the oceanic waves are modelled by the random summation of elementary waves with nonlinear corrections. Second order linear wave theory agrees with observations for moderate non-dimensional wave height ($H/H_s < 2.1$) (Casas-Prat and Holthuijsen, 2010; Gemmrich and Garrett, 2011). It has been shown in multiple data sets that rogue waves ($H/H_s > 2.2$) occur more frequently than predicted by second order theory, where inclusion of higher order corrections up to fourth order show promising correlation to observations to account for the deviation (Gemmrich and Garrett, 2011; Gemmrich and Cicon, 2022).

Until recently, strong links between rogue wave probability and sea state parameters have remained relatively elusive. Hypothesized influential spectral parameters include measures of spectral bandwidth, wave steepness and as previously mentioned BFI, but no promising correlation has been presented with rogue wave probability. In the recent study of Häfner et al. (2021b) they report on the predictive power of easily derived spectral parameters on rogue wave occurrence. The front runner from the analysis of over a billion waves was crest-trough correlation r. The wide range applicability of this parameter and the forecast ability by current state of art spectral wave models is still unknown and is therefore needing to be explored. The proposed generation mechanisms of rogue waves and sea state parameters tested as rogue wave predictors are discussed more thoroughly in Chapter 3.

Overall, the generation mechanism of rogue waves is still contested and there has yet to be a proven effective rogue wave probability forecast for the real ocean. Therefore, in this work hypothesized influential parameters for rogue wave generation, which are discussed fully in Chapter 3, are evaluated as rogue wave predictors in the Northeast Pacific (NEP). Promising parameters are then assessed if they can be reliably computed from a wave model producing standard wave forecasts. The chapters in this thesis are arranged as follows. Chapter 2 describes the calibration and implementation of a WAVEWATCH III® (WW3) wave model for the NEP and serves as a record of the grid and physics included in the model. This regional model has been developed in partnership with Environment and Climate Change Canada (ECCC) and implemented into their wave forecast system. Sea state parameters identified as rogue wave probability predictors will be included into the model output. Chapter 3 describes the rogue wave analysis, where there is a further review of the statistical description of wave heights and the proposed rogue wave generation mechanisms, as well as the compilation of the rogue wave data set and the methods used to evaluate sea state parameters for rogue wave prediction. The rogue wave data set includes buoy observations from monitoring buoys from 2010-2021 and sea surface elevation records from research buoys from late 2019 to mid 2021. From this data set crest-trough correlation proved to be the most influential as a rogue wave predictor while BFI performed relatively poorly. A case study is then provided in Chapter 4 of a large storm on October 21-22, 2021, where the wave models storm forecast and the rogue wave risk during the event is evaluated. Finally, an overall summary, limitations and future proposed work are discussed in Chapter 5.

2 Wave Model

The Northeast Pacific (NEP) is known for severe storms and large waves produced by strong winter lows (Tillotson and Komar, 1997). It is also under resolved from a Canadian operational wave forecasting perspective. Previously, wave models for the west coast of Canada included the 1/4° ($\sim$ 25 km) Global Deterministic Wave Prediction System (GDWPS) and a 500 m regional wave model of the Strait of Georgia (SoG), while the east coast has an operational regional 5 km resolution wave model. The GDWPS is a global rectilinear model from 80°S to 86°N and is maintained by ECCC (`https://weather.gc.ca/model_forecast/wave_e.html`). The SoG model is also rectilinear and additionally includes current forcing. It is run by the University of British Columbia (UBC) (Gemmrich and Pawlowicz, 2020). This chapter describes the calibration of a high-resolution unstructured model for the NEP. It is foremost developed for ECCC for the purpose of improving the regional wave forecast and additionally for producing forecasts of controlling sea state parameters in rogue wave generation for this rogue wave research. The regional grid is nested in the GDWPS parent model and includes forecasts for coastal areas like the Salish Sea that were not covered by the coarser GDWPS, and introduces high resolution throughout the entire western Canadian coastline. The model required optimization with regards to grid resolution, solving scheme, computational time-steps and physics parameterization with additional tuning. New innovations that were implemented compared to other operational ECCC wave models is the triangular unstructured grid, utilizing the implicit solving scheme and the inclusion of shallow water physics such as triad interactions. The models source term parameterizations were selected based on comparison with point source wave buoy observations and the subsequent tuning was optimized using the Cyclops Cylc-based optimization suite (Gorman and Oliver, 2018). The optimized model was implemented into ECCC's Regional Deterministic

Wave Prediction System (RDWPS) version 4.0.0 December 1, 2021.

2.1 Introduction

Regional wave models are essential for forecasting and operational planning because in-situ observations are sparse and do not fully describe an active wave field. A substantial breakthrough in the wave modelling community was the description of the sea state as the superposition of sinusoidal waves, each specified in frequency, direction, and amplitude. This opened the door for third-generation spectral wave models such as WAM and WW3 (The Wamdi Group, 1988; The WAVEWATCH III ® Development Group (WW3DG), 2019). Third-generation wave models describe the evolution in time and space of a two-dimensional wave energy spectrum $F(k, \theta)$ without imposing any assumptions about the spectral shape, where k is the wavenumber and θ is the propagation direction. The model evolves the sea state according to a balance equation for the wave action density spectrum $N(k, \theta) = F(k, \theta)/\sigma$ referred to as the Wave Action Equation (WAE).

$$\frac{DN}{Dt} = \frac{S}{\sigma} \tag{2.1}$$

$\frac{D}{Dt}$ represents the total derivative (moving with a wave component), S represents the net effect of sources and sinks for the spectrum, and σ is the radian frequency associated with waves of wavenumber k by the dispersion relation.

$$\sigma^2 = gk \tanh(kd) \tag{2.2}$$

Where g is acceleration due to gravity and d is the water depth. The left hand side of Equation (2.1) represents spatial advection and the shift in wavenumber and direction due to varying water depth. The right-hand side represent all other processes that transfer energy between spectral components. The model evolves the wave action spectrum due to wave action being a conserved quantity with and without currents; however, the model output is the more traditional frequency-direction spectrum $F(f, \theta)$, where $f = \sigma/(2\pi)$.

A notable feature of third-generation wave models is the explicit parameterization of all the source terms in S. In deep water, the net source term S is generally considered to consist of three parts; an atmosphere-wave interaction term S_{in}, which is usually a positive energy input from the wind, but can also be negative in the case of swell, a nonlinear wave-wave interaction term S_{nl}, that describes the four-wave

interactions in deep water, and a wave-ocean interaction term S_{ds}, that is generally dominated by wave breaking and white-capping. To improve initial wave growth a linear input term S_{ln} can also be added. In intermediate water depth, additional processes become important, most notably bottom friction, S_{bot}. In extremely shallow water a depth induced breaking term, S_{db}, should be considered, if it is not already represented in S_{ds}. Additionally, in shallow water there are parameterizations for the three-wave nonlinear interactions called triads, S_{tr}. This results in a net source term of,

$$S = S_{nl} + S_{in} + S_{ds} + S_{ln} + S_{bot} + S_{db} + S_{tr} \tag{2.3}$$

Various research groups over the years have developed explicit formulations for these source terms, where the choice of parameterization and tuning of parameters is a key component of the model calibration process.

WW3 is known for its applicability for large scale rectilinear models, but can also accommodate high-resolution unstructured coastal models (The WAVEWATCH III® Development Group (WW3DG), 2019). Unlike structured grids, unstructured grids allow for variable resolution to better capture irregular shorelines and complicated bottom topography, which is desirable for high resolution regional models (Cavaleri et al., 2018). However, applying a high-resolution grid to resolve the more complicated wave dynamics near the coast causes an inflation of the computational expense. This can be addressed by the type of solving scheme employed by the model. The default explicit scheme in WW3 is based on simple Euler time-stepping and is solved by splitting the WAE using a fractional step method. The steps split off effects of temporal variations of the water depth, spatial propagation, intra-spectral propagation, and source terms, which then requires the definition of four time-steps. This splitting in multiple steps allows for efficient vectorization and parallelization in the model. However, the fractional step method can suffer from splitting errors which occur due to the uncoupling of the solutions of the different sub-problems and results in a time-step dependency (Roland, 2009). Nevertheless, explicit solving schemes are often employed by forecast centres for their robustness and stability. The principal fault of explicit solving schemes is that they can suffer from severe time-step constraints given by the Courant–Friedrichs–Lewy (CFL) condition. The Courant number C given by,

$$C \equiv \frac{c_g \Delta t}{\Delta x} \tag{2.4}$$

is a measure of how much of a signal travelling at the group speed c_g traverses a

computational grid cell Δx in a given time-step Δt. Stable solutions require $C < 1$. If $C > 1$ then information is propagating through more than one grid cell at each time step which results in an unstable solution and the model will blow up. Therefore, the time-step must decrease along with the smallest spatial grid size leading to an inflation of the computational expense for high resolution grids. An alternative to the explicit scheme is an implicit solving scheme, relatively newly implemented in WW3. This scheme integrates the WAE efficiently in time and solves all propagation dimensions without any splitting between the various dimensions. Therefore, only a single global time-step is applied in computation. The source terms are integrated directly using a block Gauss-Seidel solver and linearized based on Patankar rules avoiding splitting errors that arise in the fractional step method (Abdolali et al., 2020). This allows the model to resolve offshore and nearshore physics all at once with larger time steps. It can be shown analytically and by numerical experiments that if the time step for the explicit scheme maintains $C < 1$, the results will be consistent with the implicit scheme independent of C (Abdolali et al., 2020; Smith et al., 2018). Applying an unstructured grid in conjunction with an implicit solving scheme can be effective in managing the high resolution required for modelling nearshore dynamics. Of course, both numerical methods require proper optimization with respect to the global time-step and the spatial resolution to achieve a satisfactory level of accuracy for the given applications.

2.2 Methods

2.2.1 Wave Model

The WW3 regional wave model is calibrated for the NEP, where the domain extends from northwest and southwest corners located at approximately (60°N, 145°W) and (40°N, 134°W) to the coast (Figure 2.1a). The triangular unstructured mesh is generated with OceanMesh 2D (`https://github.com/CHLNDDEV/OceanMesh2D`) in Matlab. Boundary conditions are imposed every 5 km every 12 hours at the western and southern open boundary nodes from the GDWPS global model. The open ocean resolution is therefore set to 5 km to eliminate any filtering of the incoming swell. The final nearshore resolution of 1000 m was optimized based on computation cost and model performance compared to buoy observations. The coastline of the grid was obtained from the GSHHG coastline database provided by National Oceanic and

Atmospheric Administration's (NOAA) National Centers for Environmental Information (NCEI). The bathymetry data is a blend from the global relief ETOPO1 (NCEP, 2016) with NONNA 100 survey data from the Canadian Hydrographic Service (CHS) (`https://data.chs-shc.ca/map`). The wind forcing is provided every 6 hours at 2.5 km resolution by the High Resolution Deterministic Prediction System (HRDPS) from ECCC, which is linearly interpolated onto the grid nodes. The model resolves 36 frequency bins from 0.035 - 0.98 Hz corresponding to periods of approximately 1 s to 29 s. Directional resolution includes 36 direction bins spaced every 10°.

2.2.2 Buoy Observations

The model is calibrated against buoy observations throughout the model domain (Figure 2.1b). The buoy data are provided by the Marine Environment Data Section (MEDS) of the Department of Fisheries and Oceans (DFO), the Pacific Regional Institute for Marine Energy Discovery (PRIMED) at the University of Victoria (UVic), and the National Data Buoy Center (NDBC) under NOAA. The MEDS buoys include three open ocean 6 m NOMAD buoys (C46004, C46184 and C46036) and the remainder are 3 m discus buoys (Marine Environmental Data Section Archive, Ecosystem and Oceans Science, Department of Fisheries and Oceans Canada, 2021). The moorings record vertical acceleration at 1 Hz sampling rate for 34 minutes every hour and output bulk wave statistics and one dimensional frequency spectra. Due to outages in service and erroneous records some buoys and data sections were omitted for calibration, including the open ocean buoy C46036 due to insufficient observations and the SoG buoy C46146 due to questionable measurements and a 'gappy' record. The two buoys included from NDBC (46087 and 46088) are the same 3 m discus buoys however, only H_s records are available for calibration (NDBC, 2021). In addition to these monitoring buoys, data are utilized from two 1.1 m TRIAXYS buoys deployed close to the coast and operated by PRIMED, primarily used to inform Wave Energy Converter (WEC) design and assessment. The 'Nearshore buoy' was deployed 2 km off Wickaninnish Beach in 25 m depth from 5/10/2019 to 17/06/2021, and the 'Amphitrite Bank' buoy was deployed about 7 km offshore from Ucluelet, BC in 40 m depth from 13/12/2019 to 11/02/2021. These are the only buoys with surface elevation time series, recorded in 20 minute segments every 30 minutes at 1.3 Hz. The Nearshore buoy and Amphitrite Bank buoy were instrumental in calibrating the nearshore physics. To evaluate model performance in various sea states based on

location and orography, the buoys are separated into categories. The categories are 'open ocean' which are located far off shore, 'open coastal' which are coastal buoys that are still subject to the open ocean, 'sheltered' which are coastal buoys that are sheltered from the open ocean in some directions, and 'Strait of Georgia' which are buoys in the Strait of Georgia. See Table 2.1 for full buoy names, respective category and operating organization.

Category	**Name**	**Organization**
Open Ocean	C46184 - North Nomad C46004 - Middle Nomad C46036 - South Nomad (excluded)	MEDS - DFO
Open Coast	C46147 - South Moresby C46204 - West Sea Otter C46205 - West Dixon Entrance C46206 - La Perouse Bank C46207 - East Dellwood C46208 - West Moresby	MEDS - DFO
	N-S - Nearshore Buoy A-B - Amphitrite Bank	PRIMED - UVic
Sheltered	C46145 - Central Dixon Entrance C46185 - South Hecate Strait	MEDS - DFO
	46087 - Neah Bay	NDBC - NOAA
Strait of Georgia	C46131 - Sentry Shoal C46146 - Halibut Bank (excluded)	MEDS - DFO
	46088 - New Dungeness	NDBC - NOAA

Table 2.1: Buoys used for model calibration

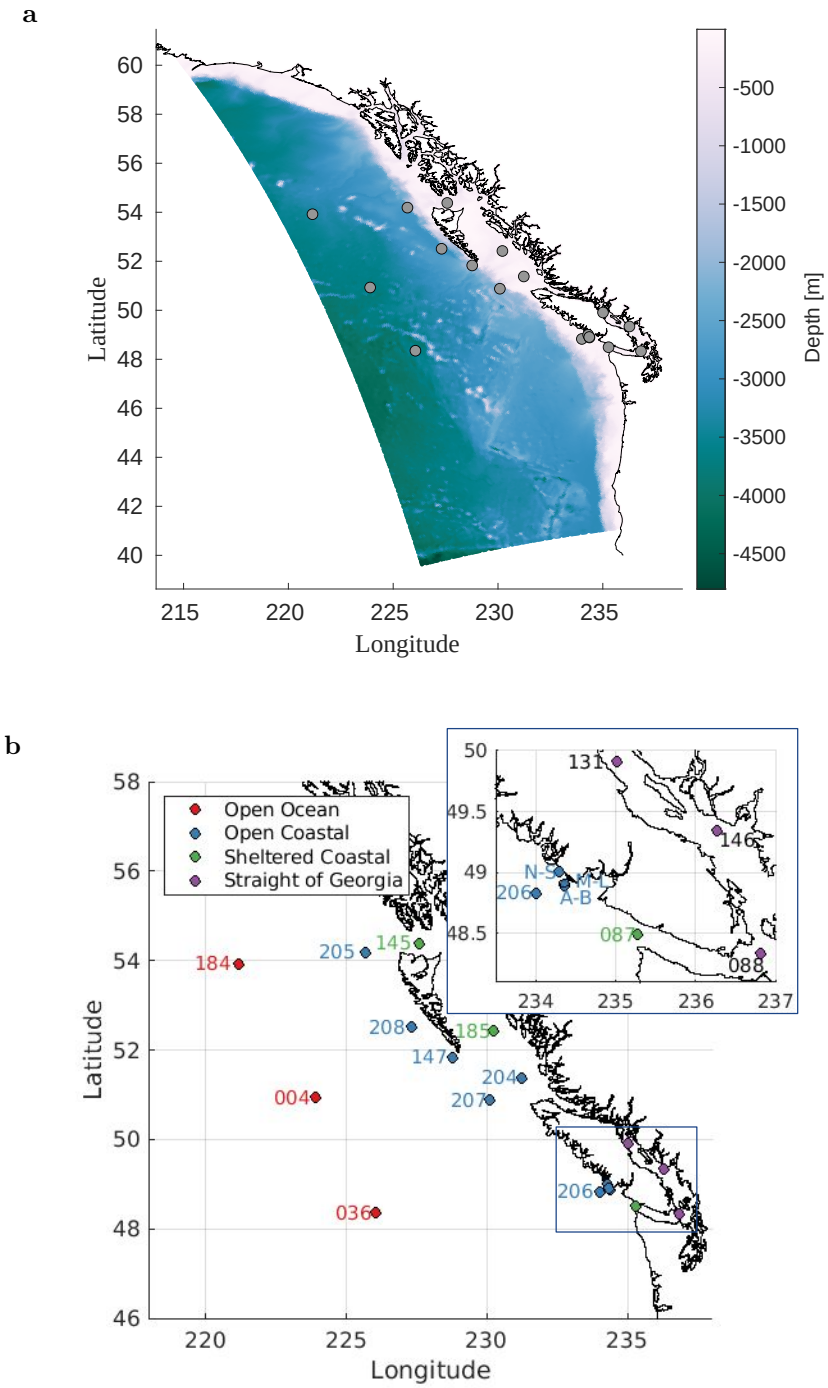

Figure 2.1: **a)** Depth field of the unstructured computational grid with buoy locations marked in grey. **b)** Buoy locations. Numbered buoys are missing the prefix 'C46' for the MEDS buoys and '46' for the NOAA buoys. Nearshore buoy is labelled N-S, Amphitrite Bank buoy is A-B and MarineLabs is M-L. The MarineLabs buoy was not available for the model calibration, but is utilized for the rogue wave analysis and is described in Chapter 3.

2.2.3 Calibration parameters and statistical scores

The accuracy of the model was determined by comparing wave forecast parameters such as wave height and period measurements with in-situ buoy observations. The level of agreement is quantified with statistical scores such as correlation coefficient, bias, and scatter index.

Significant wave height H_s is computed from the one-dimensional frequency spectrum $S(f)$ as

$$H_s = 4\sqrt{m_0}. \tag{2.5}$$

where $m_n = \int_0^\infty f^n S(f)df$ is the n^{th} spectral moment, and $S(f) = \int_0^{2\pi} F(f,\theta)d\theta$. For an intuitive understanding of H_s is it also approximately the average of the third highest waves. There are multiple definitions of wave period derived from the scaling of spectral moments. Two such quantities referred to as the mean spectral periods T_{01} and T_{02} were used for model calibration.

$$T_{01} = \frac{m_0}{m_1} \quad \text{and} \quad T_{02} = \sqrt{\frac{m_0}{m_2}}. \tag{2.6}$$

Where T_{01} is a mean period with larger emphasis on low frequencies and T_{02} has more weight on the high-frequency tail. T_{01} and T_{02} were used for calibration instead of the more commonly forecast peak period T_p to achieve a smoother evolution of the wave period. Although not used for wave model calibration, the model results for spectral bandwidth is also evaluated using a measure called narrowness σ_N.

$$\sigma_N = \sqrt{\frac{m_0 m_2}{m_1^2} - 1} \tag{2.7}$$

Bandwidth has been investigated as a relevant parameter for the performance assessment of wave energy converters (Saulnier et al., 2011; Hiles et al., 2015).

Statistical scores used to quantify the accuracy of model parameters are the correlation coefficient (R), the bias (b), and the scatter index (SI). The correlation coefficient is a measure of the strength of the linear relationship between model results and observations given by,

$$R = \frac{\sum(M_i - \bar{M})(O_i - \bar{O})}{\sqrt{\sum(M_i - \bar{M})^2}\sqrt{\sum(O_i - \bar{O})^2}}, \tag{2.8}$$

where M_i is the modelled data and O_i is the observations. The bias is a metric

showing the average difference between the model and observations.

$$b = \frac{1}{N}\sum(M_i - O_i) \tag{2.9}$$

Where N is the sample size. Bias is an indicator of overestimation or underestimate of a model. SI is a criterion indicating the scattering of points from the 1:1 line. It is obtained by normalizing the standard deviation of the differences between model results and observations by the average of all measurements, giving roughly equal weight to areas with different wave climates.

$$SI = \frac{\sqrt{\frac{1}{N}\sum(M_i - O_i)^2}}{\bar{O}} \text{ x } 100 \tag{2.10}$$

R value closest to 1, lower values of SI, and bias closest to 0 are indicate better model performance.

2.3 Model Calibration

2.3.1 Physics Parameterization

A considerable part of the calibration process is selecting the parameterizations for the source terms in Equation (2.3). The source terms and corresponding physics are implemented in the model by compiling the model with switches corresponding to certain parameterizations. The calibration of the model includes choosing which physics to incorporate and the tuning of parameters corresponding to the given parameterization. For some terms there is only one practical source term parameterization to be used in forecasting. This is the case for the nonlinear interactions between four wave components, the S_{nl} term, which for wave forecast purposes is computed with the Discrete Interaction Approximation (DIA) activated by the switch NL1 (Hasselmann et al., 1985). Although, other more complex parameterizations exist like the Webb-Resio-Tracy method and the Generalized Multiple DIA, they require a prohibitive amount of computational effort (The WAVEWATCH III ® Development Group (WW3DG), 2019). Although, the atmosphere-wave interaction term S_{in} and the wave-ocean interaction term S_{ds} represent separate processes, they are parameterized together, as the balance of these two source terms governs the growth characteristics of the wave energy. There are multiple source term packages that have

been developed for the $S_{in} + S_{ds}$ source term, however the ST4 package was selected as is it one of the more recent packages and employed in all other ECCC wave models (Ardhuin et al., 2010). The ST4 parameters and the proportionality constant for DIA were tuned using the Cyclops Cylc-based optimization method presented in Gorman and Oliver (2018). This allows a suite of parameters to be optimized efficiently instead of simply tuning with β_{max} (the coupling strength between wind and water in ST4). In general terms, the optimization algorithm explores the tunable parameter space and minimizes a computed error metric. This procedure was completed by the Meteorological Research and Development division of ECCC. The parameter defaults and the optimized values are listed in Table 2.2. For improved model initialization, and to provide more realistic initial wave growth from calm conditions, a linear input term S_{ln} was added by the LN1 switch.

Parameter name	**Default**	**Optimized**	**Description**
Betamax	1.43	1.50	Wind input
TAUWSHELTER	0.3	0.29	Wind sheltering
SWELLF	0.66	0.854	Swell propagation
SWELLF2	-0.018	-0.0188	
SWELLF3	0.022	0.022	
SWELLF4	1.50E+05	1.50E+05	
SWELLF5	1.20	1.20	
SWELLF7	3.60E+05	3.60E+05	
FXFM3	2.5	2.59	Cut-off frequency parameter tail
SDSCUM	-0.40344	-0.402	Wave breaking
SDSC1	0	0	
SDSC2	-2.2E-05	-2.2E-05	
SDSC6	0.3	0.203	
SDSBR	9.0E-04	9.1E-04	
SDSBCK	0	0	
NLPROP	2.50E+07	2.50E+07	Non-linear interactions

Table 2.2: Model physics parameters

In the nearshore region depth induced breaking, among other processes become important. Depth induced breaking is accounted for with the approach of Battjes and Janssen (1978) by the DB1 switch. Bottom friction is one of the physical mechanisms for the dissipation of wave energy in shallow water and its strength is very dependent on the bottom conditions. Consequently, the bottom friction is commonly part of the tuning procedure of a wave model and is specific to a given area. On the NEP coast the JONSWAP formulation (switch BT1) (Hasselmann et al., 1973) showed better agreement with the Nearshore buoy than the SHOWEX formulation (switch BT4) (Tolman, 1994). The shallow water process of triad interactions, an exchange of energy among three waves, was also included using the Lumped Triad Interaction model of Eldeberky et al. (1996) by the TR1 switch.

2.3.2 Solving scheme

The model was initially run with the default explicit scheme which requires global propagation, spatial propagation, intra-spectral propagation, and source term time steps. For a grid with minimum grid resolution of 1000 m, the explicit time steps of (120 s, 40 s, 60 s, 10 s) were selected, respectively. The explicit solving scheme is effective in long fetch areas like the open coast, but fails in sheltered regions like the Strait of Georgia and Hecate Strait by overestimating H_s significantly (Figure 2.2). Decreasing the minimum resolution to 500 m and therefore decreasing the time step did improve the results at the expense of computational cost, but not adequately. However, implementing the implicit numerical solver removes the overestimation in short fetch regions while maintaining agreement with the explicit scheme in unlimited fetch areas (Figure 2.2). This alludes to the overestimation being attributed to splitting errors that arise in the explicit solving scheme. A convergence analysis with respect to the time-step for the implicit solver was done and found 600 s maintained strong agreement with the buoys while being computationally efficient.

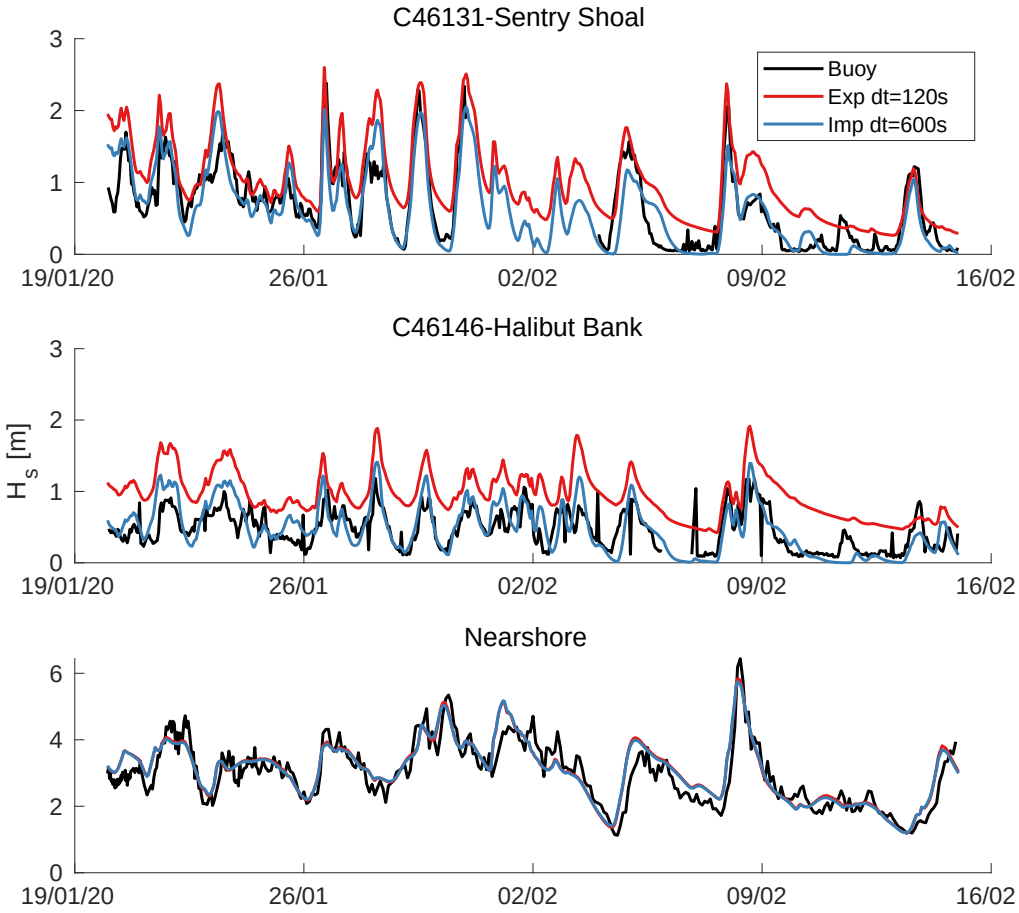

Figure 2.2: Sample time series from buoys in the Strait of Georgia (C46131 and C46146) and off the West Coast of Vancouver Island (Nearshore buoy, N-S) to illustrate the overestimation of H_s in sheltered areas using the explicit solving scheme (red) compared to the implicit scheme (blue) and their agreement on the open coast.

2.4 Results

The model was run with the settings described in Section 2.3 from 05/10/2019 to 11/07/2021 and model output of H_s, T_{01}, and T_{02} were compared with buoy observations available for that time period (Figure 2.3 and Figure 2.4). The results for T_{01} and T_{02} are comparable, but with T_{02} having longer periods than T_{01}, therefore only results for T_{01} are presented here.

In the open ocean a correlation coefficient R of 0.98, SI of 12.8%, and a bias of 0.13 m for H_s was achieved. The positive bias indicates an overall overestimation of the model compared to observations, but of a relatively small amount of 13 cm. When only considering the top 25% of H_s observations as seen in Figure 2.5, the bias

increases to 0.17 m. The scores at the open ocean buoys are quite dependent on the quality of the boundary input, as the waves seen at the open ocean buoy locations are strongly influenced by the incoming swell. Figure 2.6 shows the comparison of model runs with and without wind forcing, which highlights the influence of the input boundary in the open ocean compared to coastal regions. The collection of 'open ocean' buoys also reports the best correlation of the buoy categories for T_{01}, with an R of 0.94, SI of 7.2% and a minimal bias of 0.022 s.

Moving closer to the coast the overall scatter increases and R decreases. This is to be expected in the more dynamic coastal environment. At the 'open coastal' buoys, for H_s, R decreases slightly to 0.97, SI increases to 14.3%, and the bias nears zero at 0.023 m. So the bias is more favourably tuned to the coast then the open ocean for the bulk of the measurements. However, there is a negative bias in the top 25% of H_s measurements of 0.1 m. Meaning at larger wave heights the model tends to underestimate H_s in the coastal area. For T_{01}, there is an increase in scatter and a larger overestimation of 0.34 s.

At the 'sheltered' buoys where there are more intermediate sea heights, SI increases to 19.3%, but maintains a reasonable R of 0.94 and low bias of 0.045 m. T_{01} is overestimated at large periods, however on average underestimated by 0.14 s.

While the bias is low at the 'Strait of Georgia' buoys at 0.032 m, SI is large at 41.9%. The results for T_{01} in the SoG are poor with large underestimation of 0.89 s on average, however the sample size is smaller due to only C46131 having available and reliable results for mean period in Figure 2.4.

In lower sea state regions including the SoG and to a lesser extent at the sheltered buoys, SI is observed to be comparatively high. Despite SI being normalized by the average of the observations and therefore should be scaled appropriately with the height of seas, it is observed that SI increases when only considering low H_s observations. This trend is observed at other buoys locations as well when taking only the lower third of observations into account.

Overall the model tends to overestimate large H_s in the open ocean but, underestimates large H_s in coastal regions. These results emphasize the challenges of tuning a wave model for a variety of sea conditions and having to balance the bias throughout the domain.

Spectral bandwidth σ_N is included in this analysis in Figure 2.7 for energy resource assessment considerations. In the model, bandwidth is generally overestimated at the 'open ocean', 'open coastal' and 'sheltered' buoys with SI in the range of 30-45%. This

could be due to the difference in frequency resolution between the model and wave buoys as all the spectral parameters are computed over the full range of frequency. The MEDS buoys resolve 29 frequency bins from 0.05-0.45 Hz and the Nearshore buoy resolves 127 frequency bins from 0.005-0.62 Hz, while the model resolves 36 frequency bins from 0.035-0.98 Hz. The bandwidth results are not accurate in the SoG based on the observations from the C46131 buoy, which show negative correlation in Figure 2.7, where the ranges of the observed σ_N compared to the model are also not consistent, with the model output ranging from 0.2-0.5 while the observations range from 0.2-0.75.

To evaluate the overall performance of the model as a 48 h forecast for H_s, the bias and SI was calculated for lead times of 0-47 hours. For this analysis November 2020 was chosen as there is higher variability in the winds in the winter than in summer, therefore it provides an lower bound of the forecast capacity of the model. 48 hour forecasts were run every 6 h with the wind and boundary updating every 6 h and 12 h, respectively. At each hour within the 48 hours forecast the bias and SI was calculated against the buoys observations then grouped by location into the same 'open ocean', 'open coast', 'sheltered', and 'Strait of Georgia' categories. Of course there are not necessarily buoy observations every hour on the hour, therefore the buoy time series were interpolated but only if there were records within a 4 h window surrounding the model output time. Figure 2.8 shows that the bias and SI generally increase with increasing lead time. This is to be expected for SI as the performance will decrease in accuracy as your forecast projects in time from when the input field was calculated. In the case of bias, an increase from an initial negative bias means that the bias actually decreases with lead time and the model improves. This is case for the 'open coast' buoys and all the buoys grouped together for this time period. Overall the bias and SI increase is relatively low and the model performs quite well over a 48 hour forecast. As previous discussed the SI in the SoG is much higher than the other regions and this is reflected in Figure 2.8.

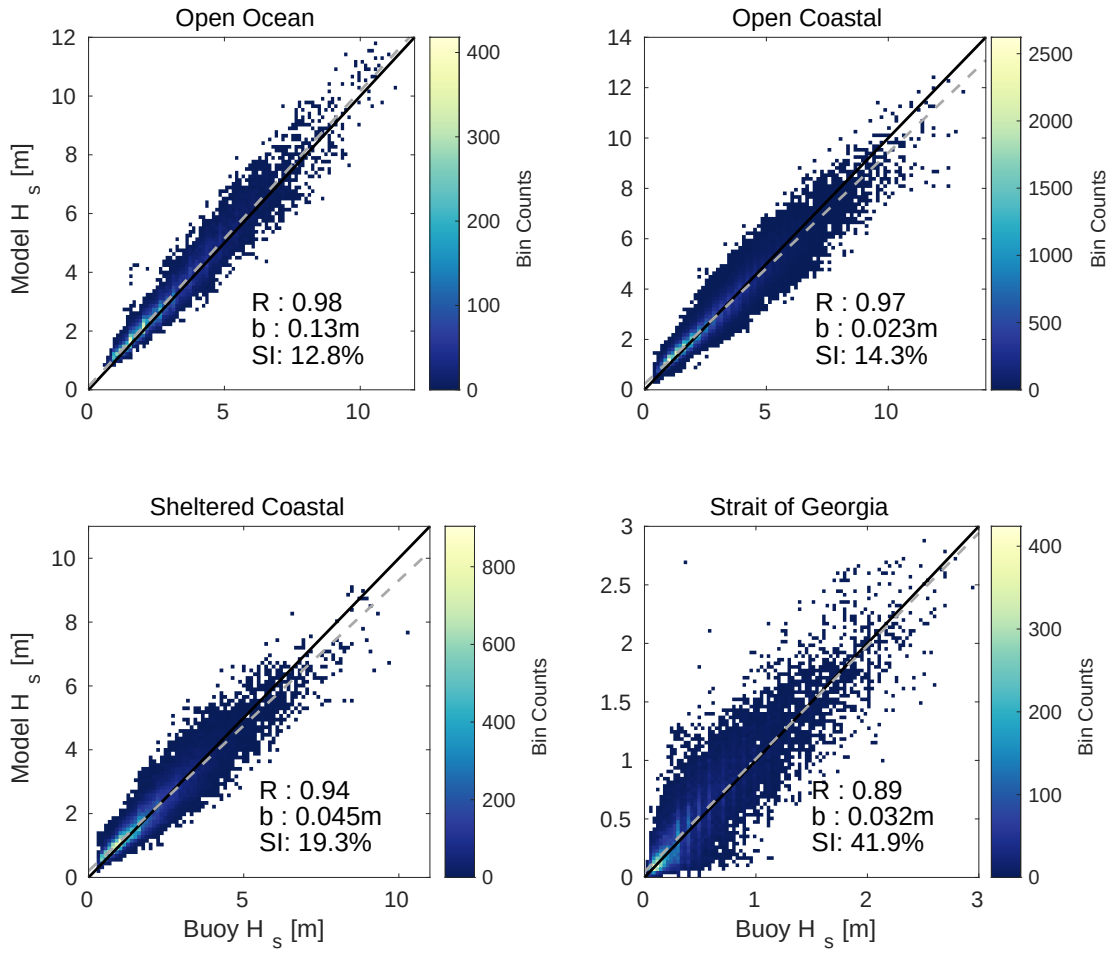

Figure 2.3: Scatter plots of observations of H_s versus model H_s. Buoys grouped into open ocean (C46004 and C46184), coastal (C46147, C46204, C46205, C46206, C46207, C46208, Nearshore and Amphitrite Bank), sheltered (C46145, C46185 and 46087) and Strait of Georgia (C46131 and 46088). Solid black line is the 1:1 line and dashed grey line is the linear fit. Statistical scores for correlation coefficient R, bias b and scatter index SI found on plot.

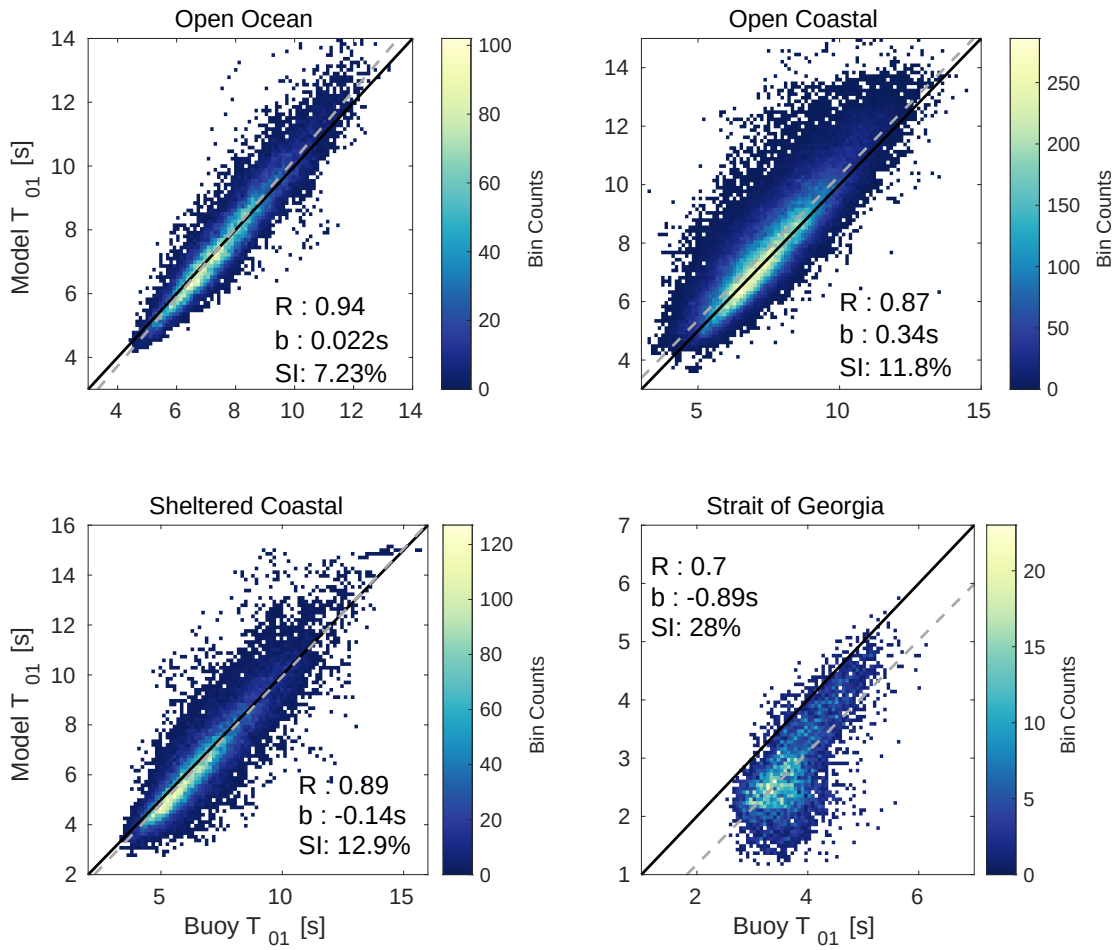

Figure 2.4: Scatter plots of observations of T_{01} versus model T_{01}. Buoys grouped into open ocean (C46004 and C46184), coastal (C46147, C46204, C46205, C46206, C46207, C46208 and Nearshore), sheltered (C46145 and C46185) and Strait of Georgia (C46131). Solid black line is the 1:1 line and dashed grey line is the linear fit. Statistical scores for correlation coefficient R, bias b and scatter index SI found on plot.

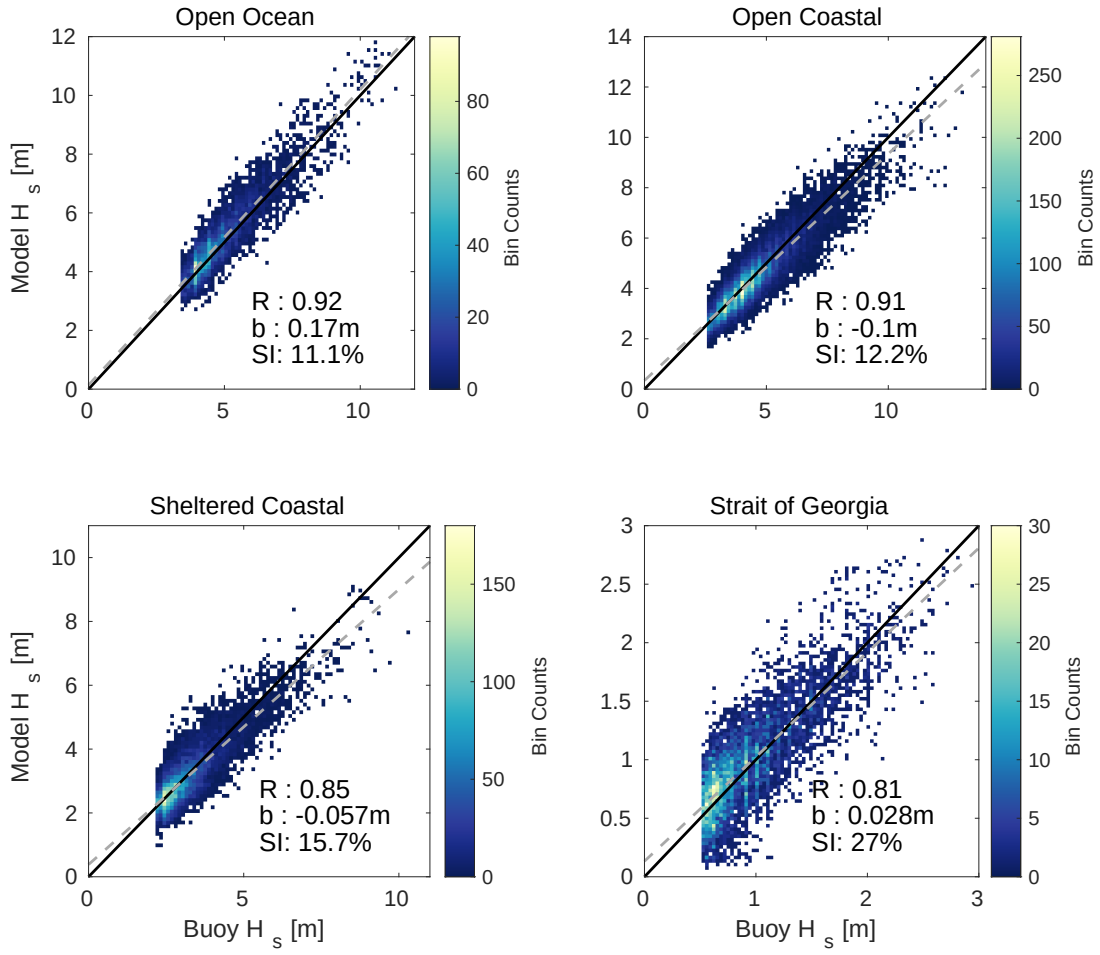

Figure 2.5: Same as Figure 2.3 except considering only the highest 25% of H_s observations.

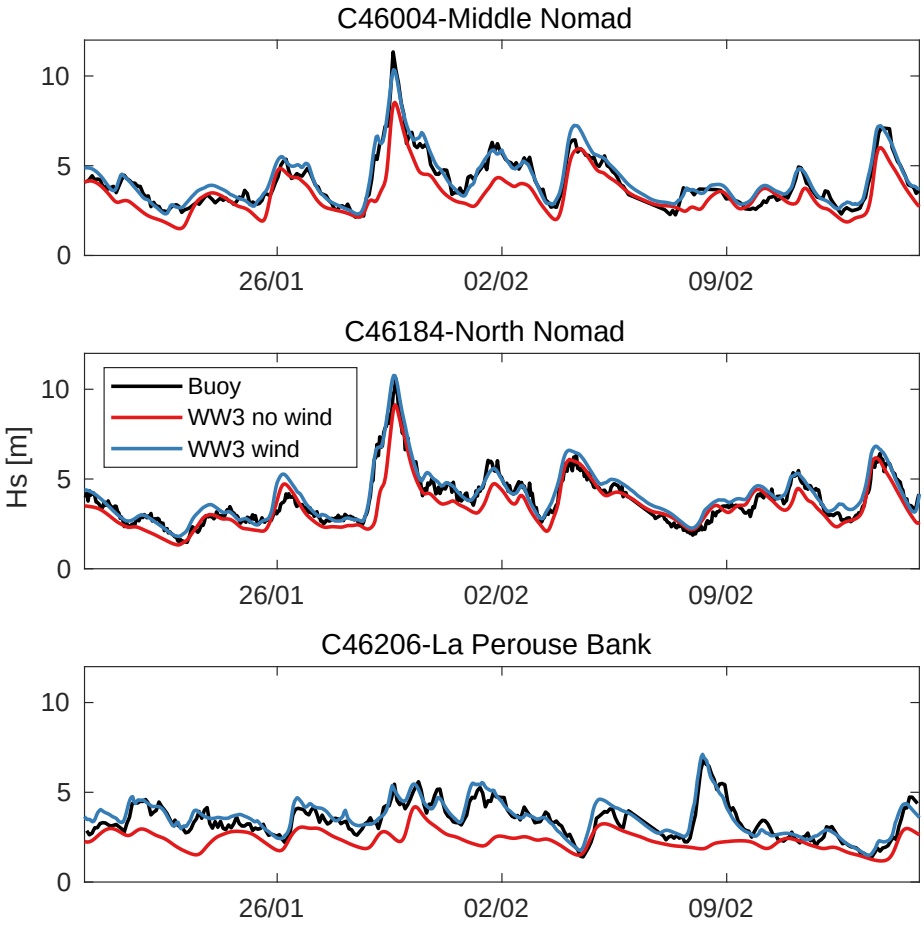

Figure 2.6: Time series of WW3 output for models runs with and without wind forcing to showcase the dominant influence of the boundary conditions in the open ocean at C46004 and C46184 compared to the coast at buoy C46206.

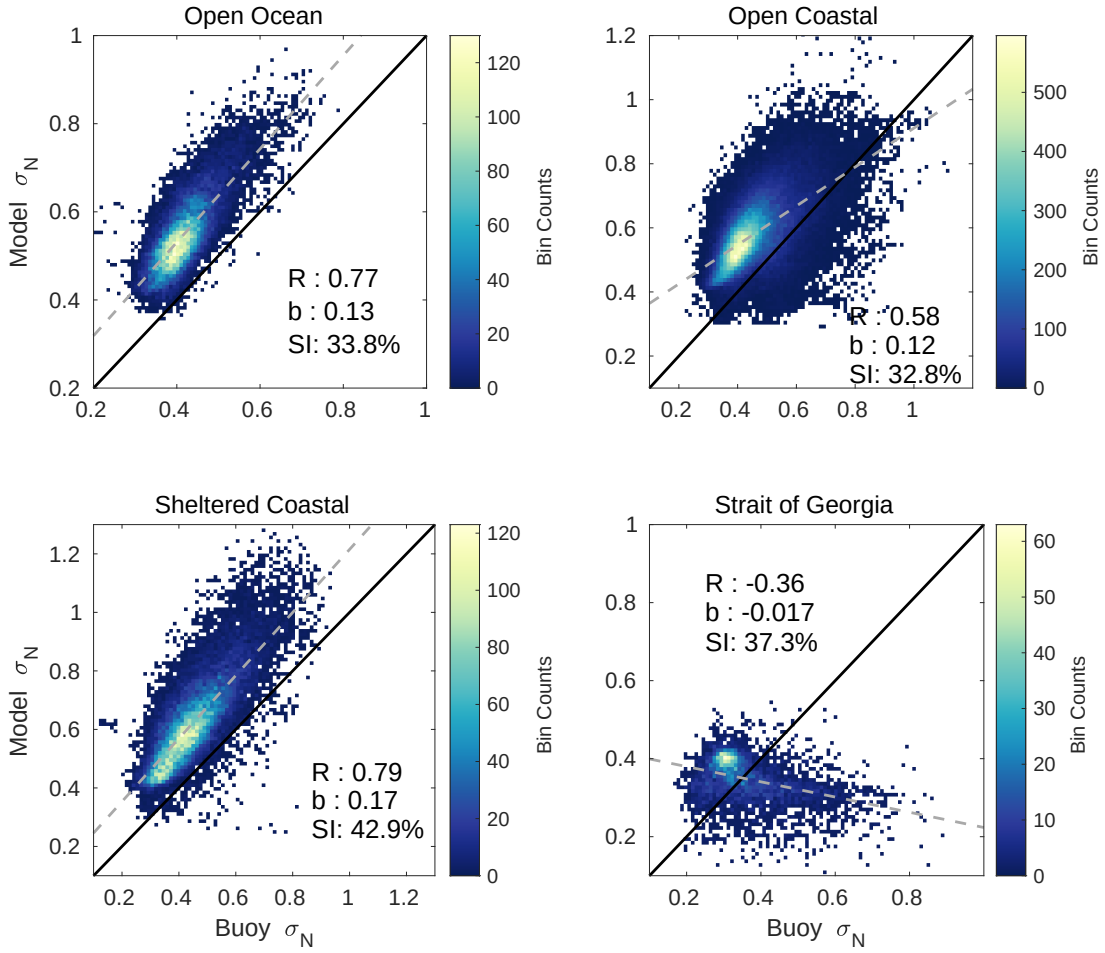

Figure 2.7: Same as Figure 2.4 except considering observations and modelled output of the bandwidth parameter σ_N.

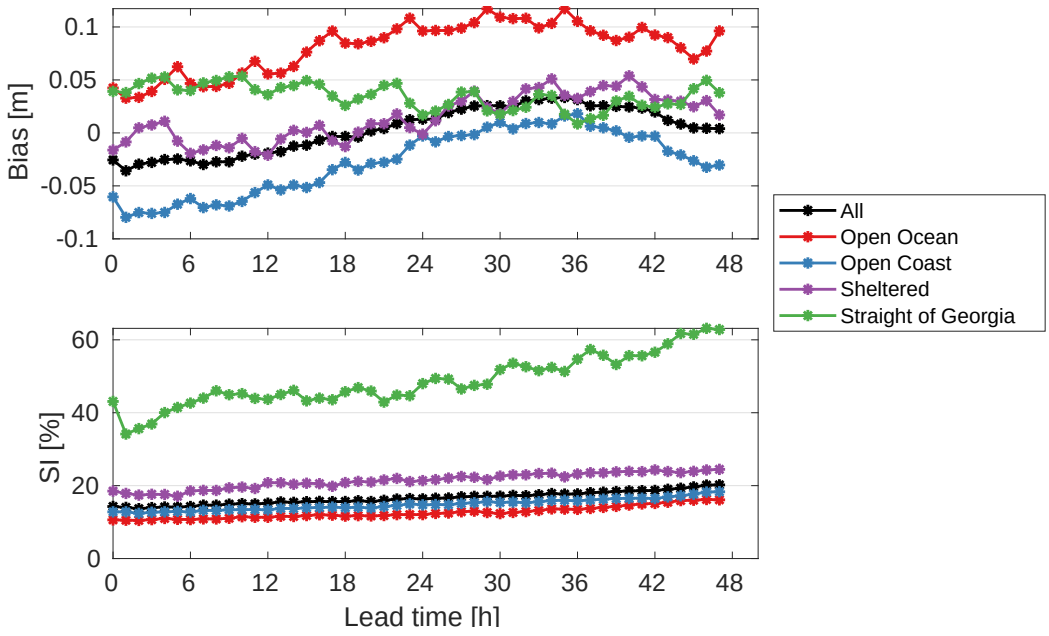

Figure 2.8: Bias and scatter index (SI) for H_s with forecast lead time for November 2020.

2.5 Discussion

Global wave models have now achieved a high degree of reliability with a scatter index of 10% and a bias on the order a few centimetres (Cavaleri et al., 2020). Nevertheless, operational regional wave models are required to achieve local high resolution forecasts, which is the operative goal of the regional NEP WW3 model described in this work. The regional model is nested within the ECCC global GDWPS model to increase the resolution to 1000 m in the nearshore. Compared to global modelling, nearshore wave forecasting is more challenging due to the variation in bottom topography, the complicated coastlines and the introduction of shallow water physics (Cavaleri et al., 2018). This is reflected in the statistical scores of the regional model results, however good agreement with buoy observations is maintained for the bulk of the data. An intercomparison of different wave buoy sensors revealed an average bias of $O(0.1\,\mathrm{m})$ (Jensen et al., 2021), which can be regarded as an estimate of the uncertainty of the buoy observations. A bias of under 10 cm was achieved for the NEP model with the exception a bias of 13 cm at the open ocean buoys. Further tuning of the model for a bias of under 10 cm is potentially unproductive due to the uncertainty of the observations. A sub 10 cm bias is also largely maintained when evaluating the model bias with forecast leadtime (Figure 2.8). The larger bias at the 'open ocean' buoys is mainly due to the influence of the boundary input. The open ocean region in the model is largely dominated by the incoming swell from the global model, so there is little to be done to correct the open ocean scores without negatively impacting the coastal results (Figure 2.6). With nesting model grids, the nested grid, to some extent is at the mercy of the calibration of the parent model.

Spectral bandwidth was included in this analysis due to wave energy assessment considerations. Historically, estimation of long-term wave energy production at a given deployment site has commonly been limited to H_s and mean period. However, studies have shown that bandwidth is a valuable characteristic to complete the set of wave parameters describing the sea state for the purpose of estimating wave energy production (Saulnier et al., 2011; Hiles et al., 2015). Unfortunately, relative to H_s the model shows poor estimation of σ_N with a domain wide overestimation of bandwidth and unreliable results in the SoG. This could be due to the difference in frequency resolution between the model and buoys and bandwidth being a more complex parameter to predict compared to H_s, as H_s only depends on the zeroth moment of the spectrum. The scores σ_N could also potentially be improved with specific model

tuning.

For the model and calibration process there are a few things to note. Firstly, the nearshore region and therefore shallow water physics were largely calibrated with the available N-S and A-B buoys. The buoys are in close proximity and relatively similar conditions and therefore do not represent the large range of coastline within the domain. The SoG is also poorly resolved with respect to observations. The current MEDS buoys in the Strait are affected by mooring dynamics and often report 'gappy' records (Gemmrich and Pawlowicz, 2020) and the 46088 buoy from NDBC does not provide sufficient raw data to compute parameters like T_{01} or T_{02}. Further validation of the model for the Salish Sea would benefit from more reliable sea state observations. Lastly, this model does not include currents, which are considered in the 500 m SoG regional model. Hence, the nonlinear interaction between waves and varying current that result in the modulation of wavelength and amplitude are not represented (Longuet-Higgins and Stewart, 1961). This distortion could become important in some areas within the model domain with considerable currents like the Juan de Fuca Strait, Puget Sound and some areas of the Strait of Georgia, which may contribute to the low scores. Incorporation of current forcing is a possibility for the NEP model and could be considered for further development. Perhaps current forcing could be limited to the Salish Sea where the affects are most prevalent and where a current model is already available.

In general wave models are tuned to the bulk of the data, which does not include the rare extreme events and therefore they have a tendency to underestimate the largest wave heights, and in particular the peaks of the storm (Cavaleri, 2009). While overall the model performs well as a wave forecast, a case study of a large storm is performed in Chapter 4 to evaluate the model capacity to capture the high seas of a large storm.

3 Extreme Waves

Following the calibration of the wave model, the next step in the development of the rogue wave probability forecast is identifying parameters correlating to rogue wave occurrence and subsequently assessing the wave model's ability to forecast these parameters. Parameters under investigation in rogue wave formation are wave steepness, spectral bandwidth, Benjamin Feir Index and crest-trough correlation. The sea state parameters observations were compiled from long term sea surface elevation records and hourly bulk statistics from buoys in the NEP. Uncertainties on the probability of encountering a rogue wave are calculated following Häfner et al. (2021a) to ensure that the changes in rogue wave probability with a given parameter are significant. This analysis finds consistent results with Häfner et al. (2021b), that crest-trough correlation r has the strongest predictive power for rogue wave probability, while Benjamin Feir Index BFI performs relatively poorly. Furthermore, it is demonstrated that r can be forecast by the regional wave model with moderate accuracy, which forms the basis of a rogue wave probability forecast. This chapter begins with a review of the statistical distributions of wave height, a description of the generation mechanisms of rogue waves, and a partial summary of the studies thus far aiming to correlate sea state parameters to rogue wave occurrence. This chapter in part has been published in Gemmrich and Cicon (2022) along with other analyses, however here additional material is presented.

3.1 Introduction

3.1.1 Rogue wave generation mechanism

The distribution of wave heights in the real ocean remains a dynamic area of research due to the availability of increasingly more and better data. Some examples being the use of laser altimeters to investigate nonlinear effects that may not be represented in surface buoy records (Casas-Prat and Holthuijsen, 2010) and the compilation of wave data sets on the order of a billion waves (Cattrell et al., 2018; Häfner et al., 2021a). However, it was Longuet-Higgins (1952) who first proposed that wave heights followed the Rayleigh distribution based on linear wave theory. In linear wave theory, the sea is approximated by a random superposition of linear (small amplitude) waves distributed by a narrow frequency spectrum. The wave height H then follows the Rayleigh distribution (Longuet-Higgins, 1952).

$$P(H) = \frac{4H}{H_s^2} \exp\left(-\frac{2H^2}{H_s^2}\right) \tag{3.1}$$

Where the probability of a wave or crest exceeding the significant wave height H_s by a multiple z is,

$$P(H/H_s > z) = \exp(-2z^2) \text{ and } \; P(\eta/H_s > z) = \exp(-8z^2) \tag{3.2}$$

for wave heights H and crest heights η, respectively. Rogue waves populate the tail of this distribution where a common criterion is waves exceeding $2.2H_s$ ($z = 2.2$) and wave crests exceeding $1.25H_s$ ($z = 1.25$). This results in an occurrence rate of rogue waves to be 1 in 16000 waves or since the dominant wave period in the ocean is $\mathcal{O}(10\,s)$ a rogue wave is expected approximately once every two days. The assumptions made to develop linear wave theory are the wave amplitude is small compared to the water depth and wavelength referred to as small amplitude waves and the frequency spectrum is narrow. These assumptions are rarely representative of the real ocean. To allow for nonlinear waves, Stokes theory was developed, where a Stokes wave is a nonlinear (finite amplitude) and periodic surface wave. Stokes theory uses a perturbation series approach, known as the Stokes expansion with respect to wave steepness to obtain approximate solutions of the wave equation for nonlinear wave motion (Kinsman, 1965). Unlike small amplitude waves, finite amplitude waves with 2nd order Stokes corrections have tall, peaked crests and broad shallow troughs,

which is a manifestation of the non-linearity. These second order corrections do not change the wave height distribution, since both the crest and the trough are elevated from the mean water level by the same amount on average. In most conditions the distribution of wave and crest heights are well described by nonlinear Stokes theory with second order corrections. However, it has been observed that a deviation begins from the 2nd order distribution for $H/Hs \gg 2$, where large non-dimensional wave height occur with higher probability (Dysthe et al., 2008; Gemmrich and Thomson, 2017). This deviation can be accounted for by including up to 4th order Stokes corrections where the wave heights are amplified as a function of wave steepness (Gemmrich and Garrett, 2011; Gemmrich and Cicon, 2022). This mechanism explains rogue wave generation as linear superposition of multiple steep waves. Additionally, real sea states are not necessarily narrow in frequency. Finite bandwidth also affects the distribution of wave height and serves to reduce the wave heights compared with the narrow-band approximation (Longuet-Higgins, 1980). Naess (1985) demonstrated that for a finite, but still narrow bandwidth the pdf of wave heights is still given by a Rayleigh distribution, but is modulated by an amount related to the bandwidth.

While the random instantaneous summation of steep waves has been shown to explain the occurrence of rogue waves, there exists another school of thought in which rogue waves are caused by a highly nonlinear effect called the modulation instability (Janssen, 2003). By modulation instability, rogue waves are generated by third order nonlinear interactions between four wave components. Essentially, in a weakly nonlinear, uniform wave train, small sideband frequencies exchange energy over time and reinforce each other which results in instability and transient events of abnormally large waves. Theoretically, these rogue waves occur when the waves are sufficiently steep, because nonlinear focusing may then overcome the spreading of energy by linear dispersion. Therefore, spectral bandwidth in combination with steepness further instigates or limits the modulation instability condition (Janssen, 2003). This initiated the introduction of the Benjamin Feir Index BFI, which is the ratio of steepness to bandwidth as a proxy for rogue wave probability. The Zakharov equation is a deterministic nonlinear evolution equation for surface gravity waves in deep water, in which the modulation instability is well described (Zakharov, 1968). Monte Carlo simulations of the Zakharov equation for periodic waves of finite amplitude show that modulation instability occurs for larger BFI (Janssen, 2003). BFI is used in routine forecasts as a rogue wave predictor at the European Centre for Medium-Range Weather Forecasts (ECMWF) (Janssen and Bidlot, 2009). However, modulation in-

stability is most effective in directionally limited and narrow banded seas, which is rarely the case in the real ocean. The discourse in the physics of rogue waves endures due to the difficulty of assessing the generation mechanism from a buoys point measurement. One cannot determine whether a rogue wave is one steep wave generated over time by the modulation instability, or by a superposition of many moderately steep waves from a buoys time series.

3.1.2 Rogue wave prediction parameters

Rogue waves cannot be predicted from forecasts in any deterministic fashion as phase averaged spectral models such as WW3 treat the wave field as a stochastic phenomenon and do not reproduce the sea surface explicitly. Therefore, a probabilistic approach is taken with the aim of uncovering spectral sea state characteristics that relate to rogue wave probability. There have been several such studies that aim to determine in what seas, if any, are rogue waves more likely to occur with varying degrees of success (Adcock and Taylor, 2014). This section gives an overview of spectral sea state parameters that have been hypothesized to be important as rogue wave indicators and will be re-examined in this work.

As previously mentioned BFI has been shown to be indicative of the modulation instability. It is calculated as the ratio of wave steepness ε to spectral bandwidth ν (Thomson et al., 2019).

$$\mathrm{BFI} = \frac{\varepsilon}{\nu} \tag{3.3}$$

Modulational instability is expected for values of BFI > 1. However, this threshold is somewhat dependent on the definition of the BFI, but it does not change the overall interpretation of trends or relative changes (Serio et al., 2005). Despite being proven theoretically and in simulations BFI has yet to be shown as a rogue wave predictor for real ocean data (Janssen and Bidlot, 2009; Kharif and Pelinovsky, 2003; Adcock and Taylor, 2014).

Rogue wave generation is influenced by spectral width by way of group dynamics (Gemmrich and Thomson, 2017). There are several definitions of spectral bandwidth ν of which narrowness σ_N (Longuet-Higgins, 1984) and peakedness σ_P (Serio et al., 2005) are considered here.

$$\sigma_N = \sqrt{\frac{m_0 m_2}{m_1^2} - 1} \tag{3.4}$$

$$\sigma_P = \frac{m_0^2}{2\sqrt{\pi}} \left(\int_0^\infty f S(f)^2 df \right)^{-1} \tag{3.5}$$

For narrow bandwidths, σ_N and σ_P approach zero, and the wave energy is concentrated near the peak frequency, as individual waves will have similar frequency with differing amplitudes modulated by the wave envelope. Cattrell et al. (2018) find that spectral bandwidth parameters display different probability distributions between rogue and normal seas, however no predictors were identified from 1.1 billion individual wave profiles, as the subset of seas containing rogue waves sits within the set of seas without (Cattrell et al., 2018).

The 4th order Stokes correction is greater for greater steepness. This implies that rogue waves are a linear superposition of steep waves. Modulation instability is also expected for waves with greater steepness. The characteristic wave steepness ε can be calculated from the peak wave number k_p and the characteristic amplitude ($a_0 = H_s/(2\sqrt{2}) = \sqrt{2m_0}$) (Serio et al., 2005).

$$\varepsilon = a_0 k_p = k_p \sqrt{2m_0} \tag{3.6}$$

Where k_p is the peak wave number associated with the spectral peak period T_p calculated from the weighted spectrum (Young, 1995) and related to wave number by the dispersion relation in Equation (2.2).

$$T_p = \frac{\int_0^\infty S(f)^4 df}{\int_0^\infty f S(f)^4 df} \tag{3.7}$$

To determine the wave number for a given T_p, the approximate inverse of Equation (2.2) is used given in Fenton (1988). No clear link between steepness and rogue wave probability has been established, as steep waves are not necessarily rogue waves (Gemmrich and Garrett, 2011).

Crest-trough correlation r was first presented as a parameter in relation to wave height distributions by (Tayfun, 1990), where r is the lagged auto-correlation of the sea surface elevation at half the mean wave period. It can be estimated from the spectral density, $S(f)$ using the Wiener-Khinchin theorem. Following Tayfun and

Fedele (2007) r is computed as,

$$
\begin{aligned}
r &= \frac{1}{m_0}\sqrt{\rho^2 + \lambda^2} \\
\rho &= \int_0^\infty S(f)\cos(2\pi f\tau)df \\
\lambda &= \int_0^\infty S(f)\sin(2\pi f\tau)df
\end{aligned}
\tag{3.8}
$$

Where $\tau = \frac{T_{01}}{2}$ is the lag time at half the spectral mean period T_{01} (Equation (2.6)). Since r is the auto-correlation of the sea surface elevation at half the wave period it is essentially an estimate of the correlation between the crest heights and the trough depths (Tayfun, 1990; Tayfun and Fedele, 2007). Therefore, r is closely related to spectral bandwidth, as in narrow banded seas the crests and successive troughs are approximately the same size (high r) and by broadening the spectrum the sea state becomes increasingly chaotic and uncorrelated (low r). By extension, r is also a measure for the tendency of the sea state to form wave groups. A recent study found that in wave buoy observations the probability of rogue waves varied by one order of magnitude with r (Häfner et al., 2021b). However, the observations in this study were concentrated in coastal areas and therefore the results may not be reflected in open ocean conditions.

3.2 Methods

3.2.1 Buoy data records

The initial foundational part of this analysis required surface elevation records which were available from the Nearshore buoy and an additional buoy provided by MarineLabs. The MarineLabs buoy is a 0.9 m CoastScout wave buoy that was deployed from 13/08/2020 to 21/05/2021 at Amphitrite Bank about 7 km offshore in 45 m water depth (labeled as M-L in Figure 2.1b). The buoy recorded surface elevation at 5 Hz from a 3D inertial measurement unit (IMU), and from differential GPS. Data from both buoys were recorded in 20 minute segments every 30 minutes. The Nearshore and Marinelabs buoy records were combined into a single data set as they are representative of the same environment with similar depth and distance to the coast. The combined record results in 28 months of sea surface elevation data.

To extend this work to offshore areas the long term time series of H_{max}, H_s and frequency spectra available from the MEDS buoys from DFO are utilized (Marine Environmental Data Section Archive, Ecosystem and Oceans Science, Department of Fisheries and Oceans Canada, 2021). This work included data from 01/01/2010 to 15/08/2021 which had received a quality control check. Any additional suspicious data sections were removed. Due to outages in service and erroneous records, the time series from the MEDS buoys are variable in length and do not necessarily span the full 11 years. Furthermore, the buoys in the SoG were excluded due to their poor reliability and low rogue wave threat in sheltered seas.

3.2.2 Estimation of Rogue Wave Probabilities

To determine the dependence of rogue wave occurrence rates on sea state parameter x, the univariate rogue wave probability p is computed while varying x. A similar method as Häfner et al. (2021a) is followed, which includes evaluating uncertainties on the estimates of p to establish if correlations are significant. p is defined as the probability that the next wave height will exceed the threshold $2H_s$ or $2.2H_s$. To evaluate how p varies with wave parameter x, x and H/H_s are split into N bins with approximately equal number of observations in each bin to achieve good statistics. It is assumed that the number of rogue waves n^+ and the number of non-rogue waves n^- in each bin are identically and independently distributed (iid) according to a binomial distribution with probability p.

$$P(n^+, n^-|p) = \text{Binom}(n^+ + n^-, p) \tag{3.9}$$

It is also assumed the prior distribution of p is a Beta distribution with parameters $\alpha_0 = 1$ and $\beta_0 = 3000$ for $P(H/H_s > 2)$ and $\beta_0 = 16000$ for $P(H/H_s > 2.2)$ given by,

$$P(p) = \text{Beta}(\alpha_0, \beta_0) \tag{3.10}$$

Where the values α_0 and β_0 are calculated from the Rayleigh distribution. The purpose of the Beta prior is only to constrain p to a reasonable order of magnitude, therefore the exact choice of α_0 and β_0 does not influence the final results. Applying Bayes theorem the posterior probability for p becomes another Beta distribution, due to the Beta prior for p being conjugate to the binomial distribution of n^+.

$$P(p|n^+, n^-) = \frac{P(n^+, n^-|p)P(p)}{P(n^+, n^-)} = \text{Beta}(n^+ + \alpha_0, n^- + \beta_0) \tag{3.11}$$

The estimate of p with uncertainties is calculated from the median and the 95% credible interval of Equation (3.11) (based on 2.5^{th} and 97.5^{th} percentiles). The benefit of this analysis is to be able to generate uncertainties for p to determine whether the variation with x is meaningful. This analysis hinges on the assumptions that p is iid distributed in each bin, which is not the case if p depends on more than one parameter. Therefore, these uncertainties indicate the level of confidence in the best estimate of p if only one parameter is considered at a time.

For buoys where surface elevation is unavailable, the normalized wave is given by H_{max}/H_s, where H_{max} is the maximum wave in a record and the total number of waves N_0 is estimated using the mean period T_{01} and the record length RL where, $N_0 = \frac{RL}{T_{01}}$. The limitation with this simplification is that the number of flagged rogue waves in a record is limited to one.

The predictive power is used to evaluate the degree of variability of p with parameter x defined as, $\mathbb{P}_x = \log_{10}\left(\frac{p_{imax}}{p_{imin}}\right)$, where i_{max} is the bin index where x is highest and i_{min} is bin index where x is lowest. This is a measure of how much p varies with x if only considering this one parameter. The uncertainty in $\mathbb{P}_x$ is quantified through Monte Carlo sampling, based on the known distributions of p_{imax} and p_{imin} given in Equation (3.11) and the 95% confidence interval calculated from the 2.5^{th} and 97.5^{th} percentiles of the distribution of $\mathbb{P}_x$.

3.3 Results

3.3.1 Rogue wave occurrence rates

In the combined record of sea surface elevation from the Nearshore and MarineLabs buoys there are 160 waves where $H/H_s > 2$ and 22 waves where $H/H_s > 2.2$ out of a total of 5.7 million individual waves (Figure 3.1). Figure 3.1b shows an extreme example of a rogue wave event on 17/11/2020 that is specifically analyzed in Gemmrich and Cicon (2022), where $H/H_s = 2.9$ in high seas of $H_s = 6.05\,m$. Since large non-dimensional wave height correspond to small values of exceedance probability a convenient representation is to plot $\ln(-\ln P)$ against H/H_s on a logarithmic scale. The Rayleigh distribution then becomes a straight line, $\ln(\ln P(H/H_s > z)) = 2\ln z + \ln 2$,

where deviation from linear theory are more easily seen. Figure 3.2 shows the distribution of wave heights and corresponding exceedance probabilities from the combined surface elevation records. As commonly observed, the probabilities of wave heights is overall lower than predicted by the Rayleigh distribution due to the spectrum being insufficiently narrow (Longuet-Higgins, 1980). The observations follow the linear model up to $H/H_s < 2.1$, but beyond that show a dramatic increase of probabilities of rogue waves, which has been observed in other data sets (Dysthe et al., 2008; Gemmrich and Thomson, 2017). Through simulations of random superposition of linear waves, Gemmrich and Cicon (2022) showed that allowing for 4th order Stokes correction and intermediate water depth resulted in good agreement with the observations from the MarineLabs buoy.

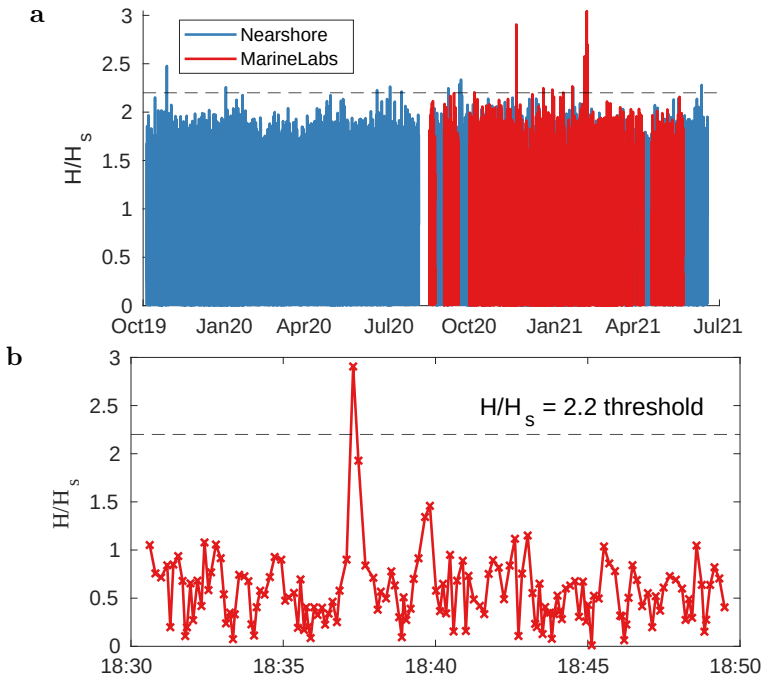

Figure 3.1: **a)** Entire time series of non-dimensional wave height H/H_s at the Nearshore and Marine Labs buoys. **b)** An example of a extreme rogue wave event at the Marine Labs buoy on 17/11/2020, where $H/H_s = 2.9$ in high seas of $H_s = 6.05\,\mathrm{m}$.

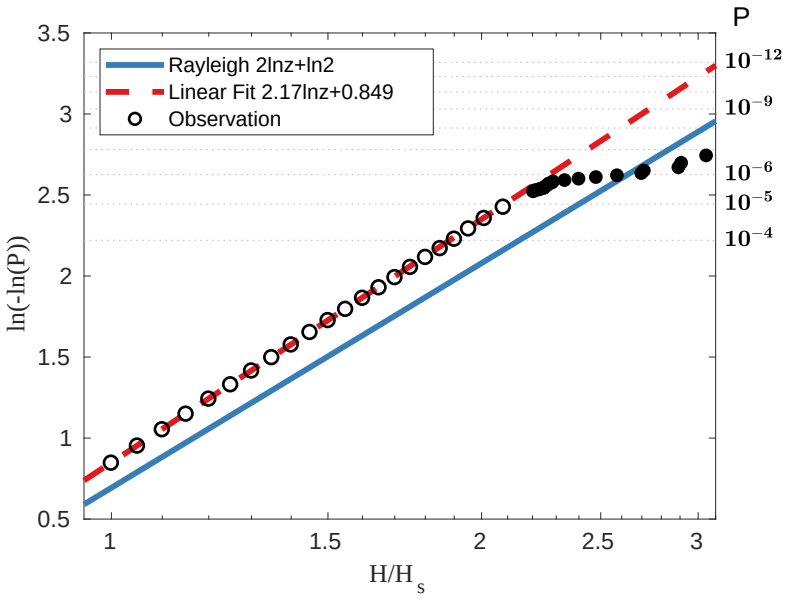

Figure 3.2: Exceedance probability P of non-dimensional wave height H/H_s of combined data set of Nearshore and Marine Labs buoys. Filled symbols show observed rogue waves where $H/H_s > 2.2$. The observations were down sampled for plotting purposes.

3.3.2 Rogue wave risk prediction

To develop the rogue wave risk prediction system, the rogue wave probability p was computed while varying sea state parameters. The parameters evaluated are crest-trough correlation r, measures of spectral bandwidth narrowness σ_N and peakedness σ_P, steepness ε, and Benjamin Feir Index given by $\text{BFI}_N = \frac{\varepsilon}{\sigma_N}$ and $\text{BFI}_P = \frac{\varepsilon}{\sigma_P}$.

An initial analysis was performed using the surface elevation measurements from the combined Nearshore and MarineLabs buoys. As described in Section 3.2.2, to evaluate p's dependence on parameter x, p and the 95% confidence interval are calculated by quantifying the number of rogue events in each x bin for $H/H_s > 2$ and $H/H_s > 2.2$. Figure 3.3 shows a strong positive correlation with r for both rogue wave limits and a strong negative correlation with σ_N for $p = P(H/H_s > 2.2)$. With r the occurrence rate of a wave exceeding $2H_s$ varies from 1 in 110,000 to 1 in 16,000 and the occurrence rate of a wave exceeding $2.2H_s$ varies from 1 in 1,400,000 to 1 in 140,000. For σ_N the occurrence rate of a wave exceeding $2H_s$ varies from 1 in 44,000 to 1 in 24,000 and the occurrence rate of a wave exceeding $2.2H_s$ varies from 1 in 1,400,000 to 1 in 110,000. When the maximum non-dimensional wave height is plotted against r and σ_N from the buoy observations and the wave model output the

same trend is observed as with the rogue wave probability (Figure 3.4).

To quantify the variation of p with x the 'predictive power' is defined, which is the logarithmic ratio of p between the bins where x is highest and lowest. For example if $\mathbb{P}_x = 2$, it means that p changes by 2 orders of magnitude as x is varied. Based on this observational data set for $P(H/H_s > 2)$, r has by far the strongest predictive power with $\mathbb{P}_x = 0.85$ with lower bound of 0.56. For $P(H/H_s > 2.2)$ r and σ_N have comparable predictive powers of 0.99 and 1.1, respectively (Figure 3.5). All other parameters perform poorly as a rogue wave indicators including the measures of BFI where the predictive power lower confidence bound scarcely falls above 0 for $p = P(H/H_s > 2)$ and includes zero for $p = P(H/H_s > 2.2)$.

r and σ_N are closely related and are correlated when considering mean trends, which explains their joint correlation with non-dimensional wave heights and rogue wave probability. However, Figure 3.6 shows that low bandwidth σ_N correspond to high r, but high r does occur at any bandwidth. This is also reflected in the distribution of rogue events. Therefore, using low σ_N as rogue wave indicator would not include rogue events that would be considered under the umbrella of high r.

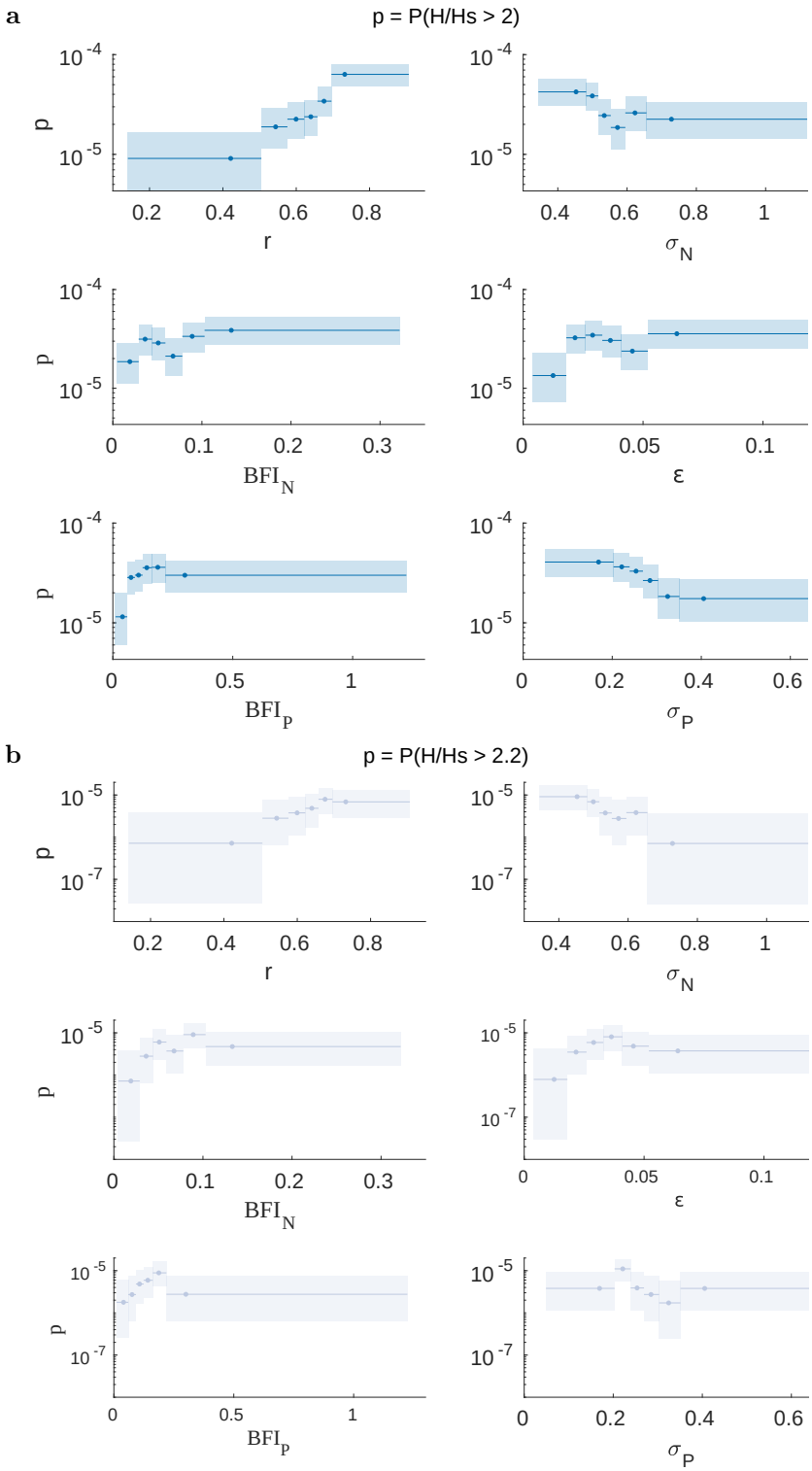

Figure 3.3: Variation in rogue wave probability p with seas state parameters, where in each bin there are equal number of observations. Shaded area is the 95% confidence interval for p, where **a)** $p = P(H/H_s > 2)$ and **b)** $p = P(H/H_s > 2.2)$. The point markers are plotted at the mean sea state parameter in the bin.

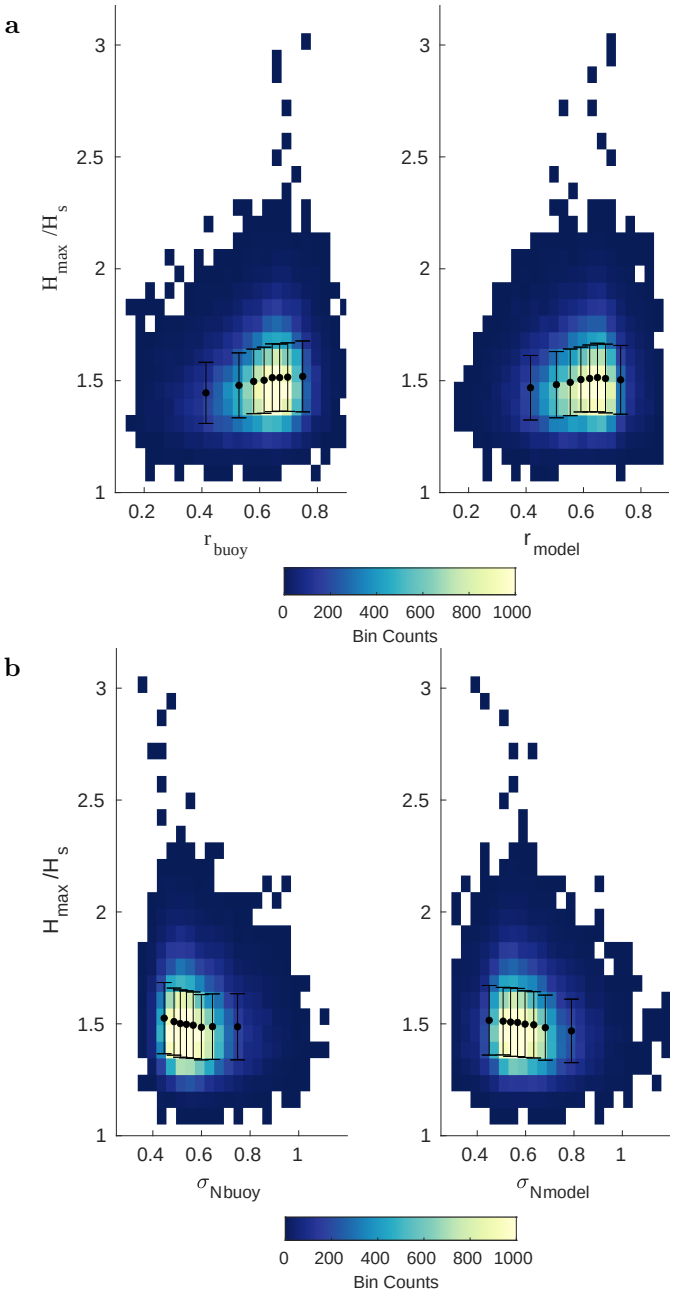

Figure 3.4: Normalized maximum non-dimensional wave height H_{max}/Hs within 20 minute record, as function of **a)** r and **b)** σ_N. With buoy observations on the left and model output on the right. Individual data are shown as scatter plot, with data density given by colour, and constant-size bin averages and $\pm\sigma$ are shown in black.

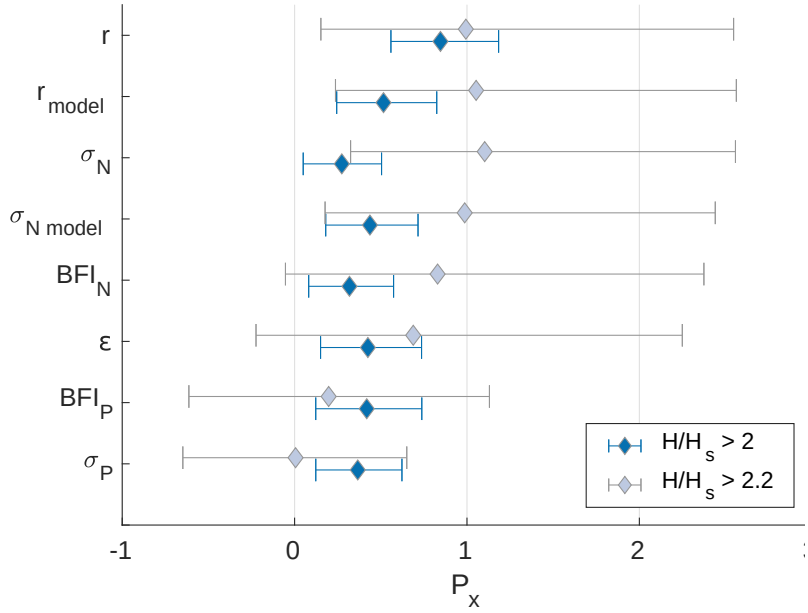

Figure 3.5: Predictive power of sea state parameters. Error bars represent 95% confidence bounds. Where the correlation between p and parameter x is negative (negative $\mathbb{P}_x$) the absolute value is taken for comparison. For example in the case of narrowness σ_N (Figure 3.3).

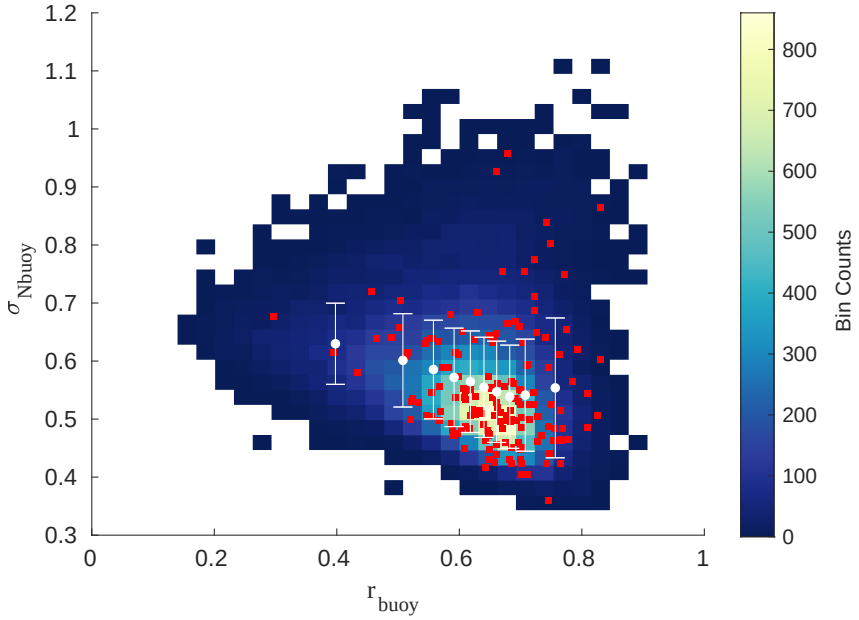

Figure 3.6: Observations of r versus σ_N. Individual data are shown as scatter plot, with data density given by the colour bar, and constant-size bin averages and $\pm\sigma$ are shown in white. Red markers represent individual rogue waves events where $H > 2H_s$.

An influential factor to consider in developing the rogue wave probability forecast is the model's capacity of predicting the risk parameters. Despite the regional WW3 model predicting H_s with high reliability at the Nearshore and MarineLabs buoy locations ($R = 0.96$), the model's prediction of r and σ_N are less reliable with similar statistical scores with correlation coefficient of 0.61 and 0.59, respectively (Figure 3.7). However, when r and σ_N model output was examined at an offshore location like the C46004 open ocean buoy, the prediction of r is greatly improved, while σ_N shows high bias and scatter (panels b,c,e and f in Figure 3.7). There is a similar trend at the other MEDS buoy locations where r is far better predicted by the model and σ_N is greatly overestimated. This is demonstrated in the scatter plots for r and σ_N (Figure 3.8), which include observations for the entire 21.5 month model time series from 05/10/2019 to 11/07/2021.

The correlation between rogue wave probability and r and σ_N output from the model inevitably varies due to the ability of the model to predict those parameters, which has been shown to be more challenging in near shore waters. This deficiency is translated into changes in the correlation and predictive power between p and r and σ_N computed from the model (Figure 3.5 and Figure 3.9). Most notably, for r_{model} and $p = P(H/H_s > 2)$, $\mathbb{P}_x$ decreases to 0.49 from 0.85. For σ_N and $p = P(H/H_s > 2)$ the absolute value of $\mathbb{P}_x$ actually increases from 0.27 to 0.38. $\mathbb{P}_x$ for $H/H_s > 2.2$ stays relatively consistent for both r and σ_N between the observational and modelled parameters.

Long term surface elevation records, which are the basis of the analysis presented thus far, are few and are mainly restricted to coastal areas. However, time series of hourly H_{max}, H_s and frequency spectra from the MEDS buoys are more readily available, span decades and include offshore areas. There are 15000 records in which $H_{max}/H_s > 2$ and 3596 records in which $H_{max}/H_s > 2.2$. However, the analysis must be modified slightly where the number of waves in a record N_0 must be estimated by the mean period and the non-dimensional wave height is given by H_{max}/H_s. To evaluate if this approach is viable these assumptions were included on the Nearshore and MarineLabs data set to compare with previous results using the individual wave heights. Using H_{max}/H_s excludes only three rogue events in the entire record. Those events correspond to segments where there were two waves with height exceeding $2H_s$. Overall, using H_{max} and T_{01} to get N_0 does not greatly affect the results for the Nearshore and MarineLabs buoys (Figure 3.10). However, the expansive data set of Häfner et al. (2021a) contains a relatively high number of multiple rogue waves in

rapid succession (about 2500 waves with $H/H_s > 2$ within 30 s of each other), which corresponds to about 3% of all rogue events. Therefore, by limiting the number of rogue events per record to one, the rogue wave occurrences are under represented.

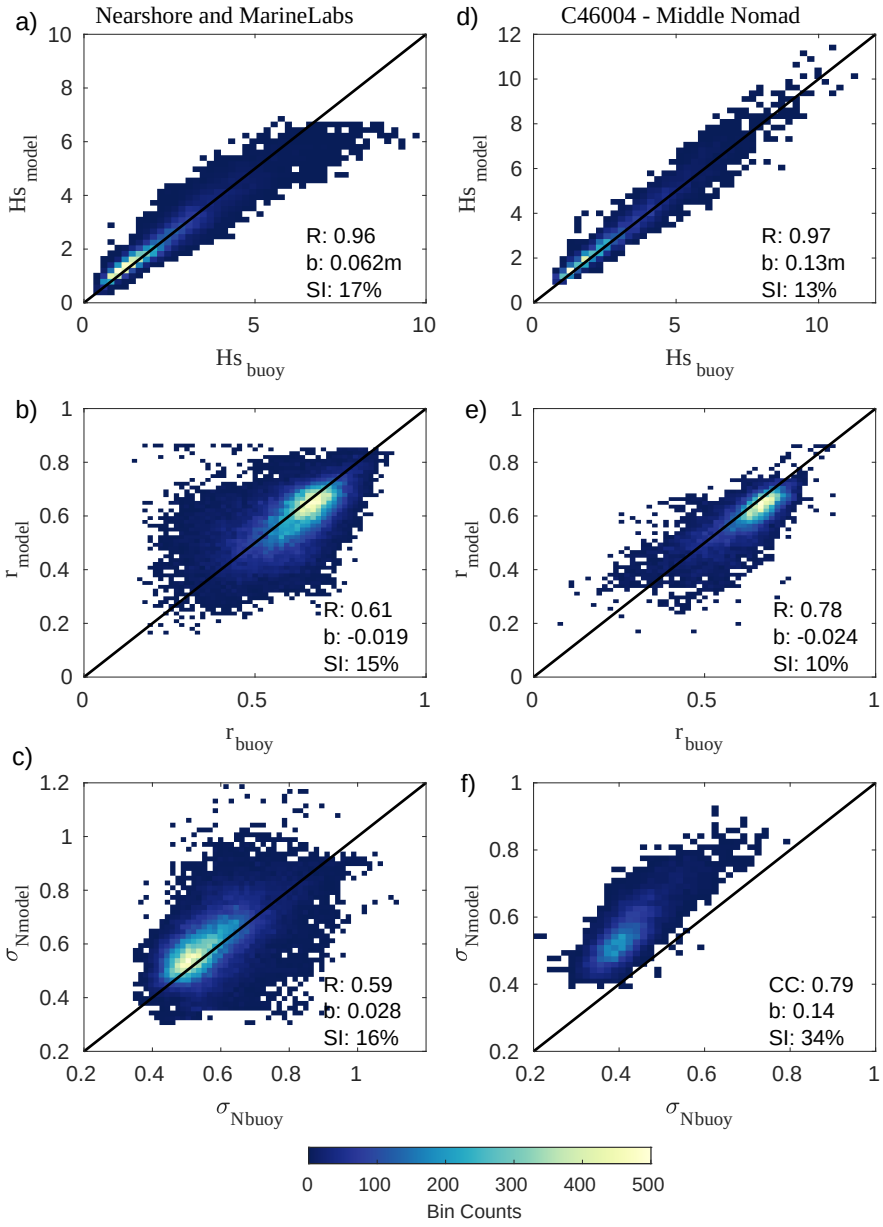

Figure 3.7: Scatter plots of combined 28 months of observations from Marine Labs and Nearshore buoys versus WW3 model output for **a)** significant wave height H_s, **b)** crest-trough correlation r and **c)** narrowness σ_N. **d-f)** Same plots for C46004 - Middle Nomad buoy over the same time period. Correlation coefficient R, bias b and scatter index SI reported on the figure. Solid black line is the 1:1 line.

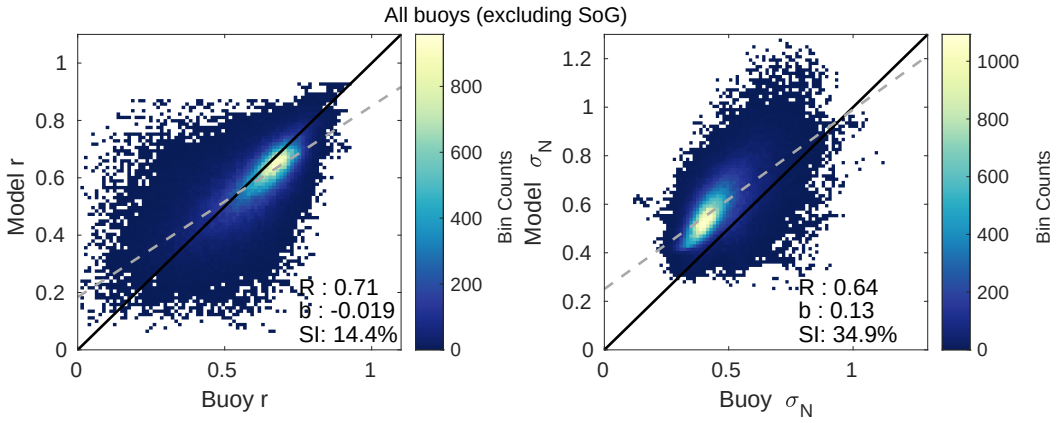

Figure 3.8: Scatter plots of all the buoys, excluding the SoG buoys, for the entire calibration time period 05/10/2019 to 11/07/2021 for crest-trough correlation r and narrowness σ_N. Correlation coefficient R, bias b and scatter index SI reported on the figure. Solid black line is the 1:1 line.

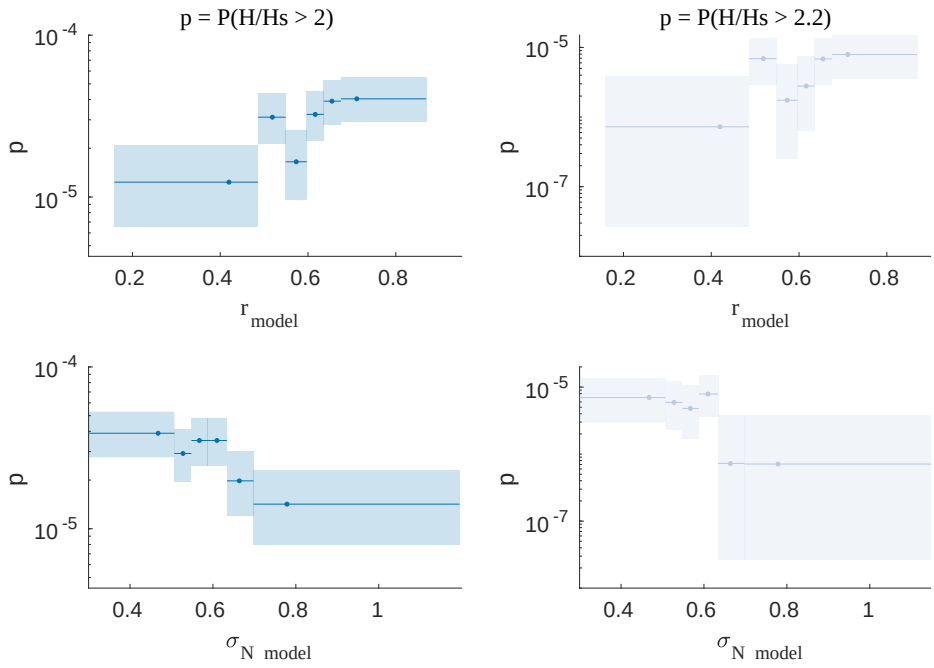

Figure 3.9: Rogue wave probability p as a function of sea state parameter crest-trough correlation r and spectral narrowness σ_N from the WW3 model. Shaded area is the 95% confidence interval for p and the point markers are plotted at the mean sea state parameter in the bin.

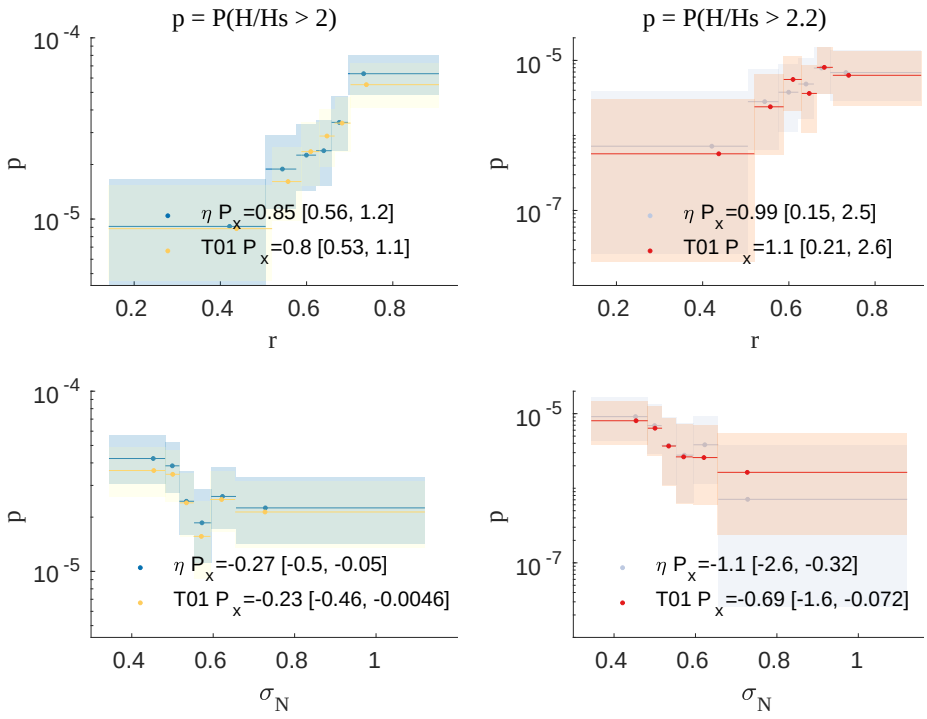

Figure 3.10: Variation in rogue wave probability p with seas state parameters r and σ_N with either using the sea surface elevation η to get the number of rogue waves n^+ and total number of waves N_0 or using H_{max}/H_s to get n^+ and T_{01} to get N_0. Each bin has an equal number of observations. Shaded area is the 95% confidence interval for p and the point markers are plotted at the mean sea state parameter in the bin.

Extending this analysis with the required assumptions to the offshore MEDS buoys, a similar strong trend of high correlation between p and r is present (Figure 3.11). At all MEDS locations the median $\mathbb{P}_x$ for $p = P(H_{max}/H_s > 2)$ and $p = P(H_{max}/H_s > 2.2)$ increases and the lower bound either stay roughly the same or increases. Figure 3.12 demonstrates that there is still a negative correlation present with σ_N at most of the buoy locations, however to a lesser degree than r with lower predictive power and a uninspiring level of correlation at the open ocean buoy locations C46004 and C46184. Figure 3.13 shows the comparison of predictive power between r and σ_N at the MEDS buoy locations, which is greatly in favour of r. Furthermore, the range of σ_N greatly decreased when going from the nearshore buoys to offshore, whereas the range and distribution of r is relatively constant. This is illustrated in Figure 3.14 using the distribution of r and σ_N from the combined Nearshore and Marine Labs buoys versus the distributions from C46004. With the extension of the analysis, r is proven to be an effective parameter for the full domain of the operational wave forecast.

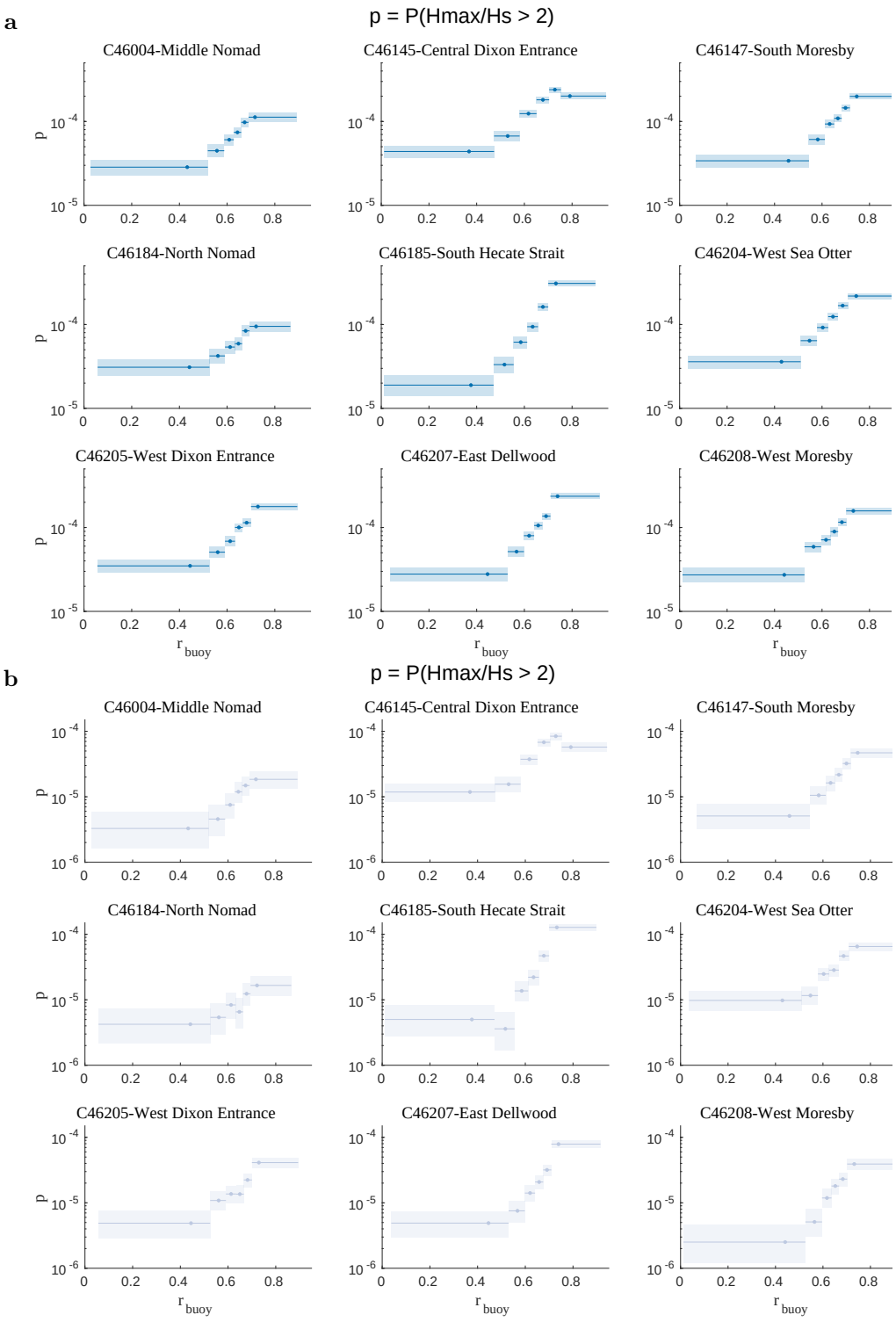

Figure 3.11: Variation in rogue wave probability p with crest-trough correlation r with shaded area as 95% confidence intervals for MEDS buoys where **a)** $p = P(H/H_s > 2)$ and **b)** $p = P(H/H_s > 2.2)$. Each bin has an equal number of observations. The scatter points are plotted at the mean parameter value in the bin.

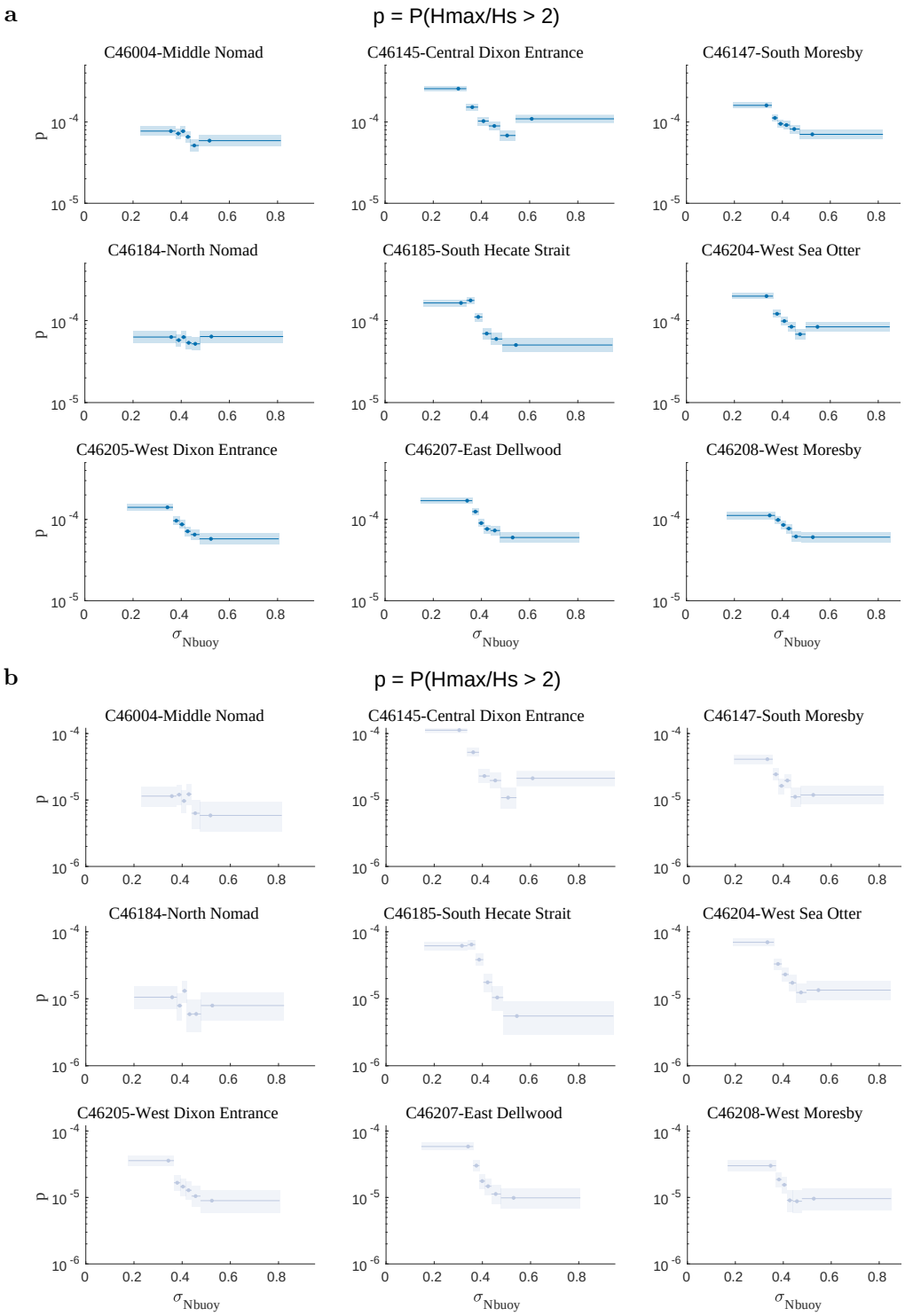

Figure 3.12: Variation in rogue wave probability p with spectral bandwidth narrowness σ_N with shaded area as 95% confidence intervals for MEDS buoys where **a)** $p = P(H/H_s > 2)$ and **b)** $p = P(H/H_s > 2.2)$. The scatter points are plotted at the mean parameter value in the bin and each bin has an equal number of observations.

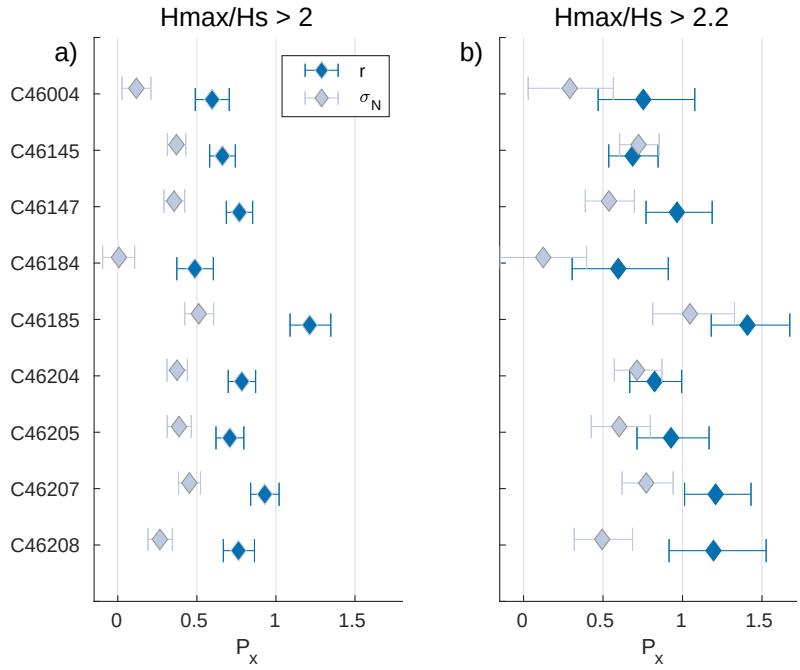

Figure 3.13: Predictive power of r and σ_N for MEDS buoy locations. Error bars represent 95% confidence bounds. Where the correlation between p and parameter σ_N is negative the absolute value is taken for comparison.

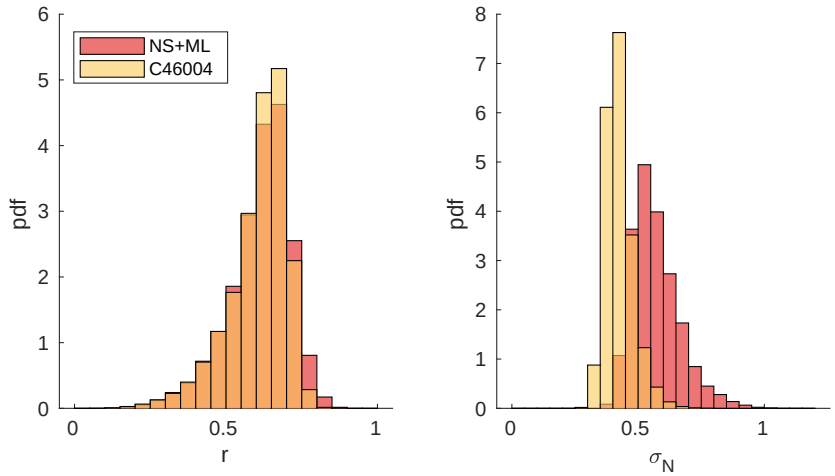

Figure 3.14: Distribution of r and σ_N from the combined Nearshore and MarineLabs buoy (NS+ML) and offshore buoy C46004.

3.3.3 Risk forecast

Overall crest-trough correlation r is a more robust prediction parameter over narrowness σ_N based of the following,

- Stronger correlation with p when considering all buoy locations.
- Models predictions of r is better than σ_N.
- Increased risk of rogue events at low bandwidth is included in high r sea states (low bandwidth correspond to high r.)
- Range of σ_N decreases considerably from nearshore locations to offshore while the range of r is maintained.

To generate a risk assessment semi-logarithmic fits were applied at all buoy locations, where the average r value in each bin is used for fitting (Figure 3.15). The slopes at each buoy are relatively similar and there is no strong trend between slopes when grouped by depth, region or distance from the coast. The mean fits for $P(H/H_s > 2)$ and $P(H/H_s > 2.2)$ are as follows,

$$\begin{aligned} P\left(H/H_s > 2\right) &= 10^{(2.54\pm0.48)r-(5.70\pm0.29)} \\ P\left(H/H_s > 2.2\right) &= 10^{(3.31\pm0.86)r-(6.90\pm0.53)} \end{aligned} \tag{3.12}$$

While the sampling range of r is poorly resolved above 0.8 and below 0.4, the distribution of p with r shows no indication of deviating from the log linear relationship. Therefore, it is believed to be appropriate to extend the fit to r values outside of the densely populated sampling range. Using this relationship in Equation (3.12) the probability of a wave exceeding $2H_s$ or $2.2H_s$ can be calculated with the spatial and temporal fields of r output from the regional wave model. An example of such output of H_s, r and p is shown in Figure 3.16, where the frequency spectra were output every $1/4^\circ$ to compute r. Combined output of H_s and r provides a complete risk assessment, with overlapping areas of high r and large H_s being the most dangerous.

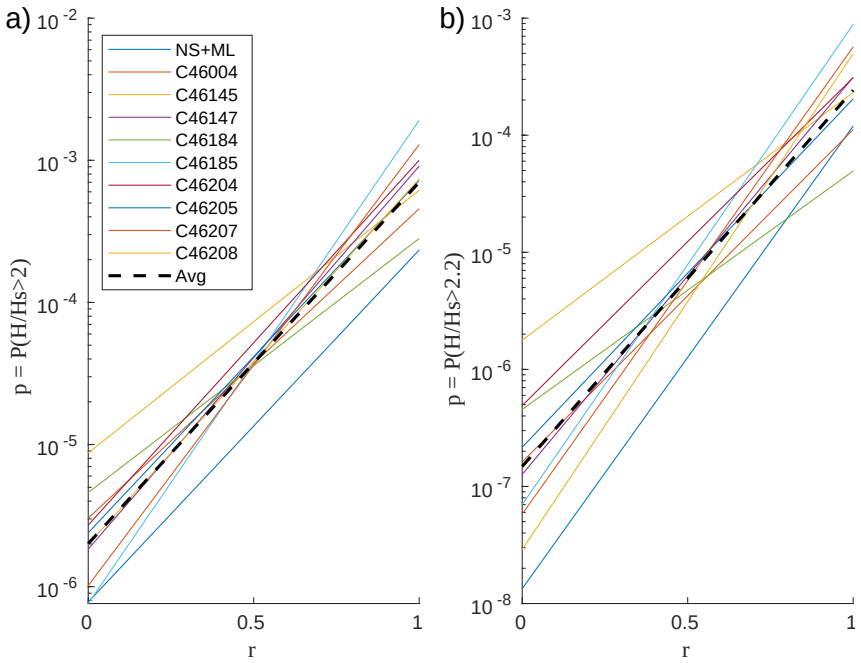

Figure 3.15: Semi logarithmic fit of p versus r for **a)** $p = P(H/H_s > 2)$ and **b)** $p = P(H/H_s > 2.2)$. Where the dashed black line in the average over all buoys.

3.4 Discussion

This study analyses multiple sea state parameters for suitability as rogue wave indicators and has found additional evidence in favour of developing a rogue wave risk forecast based on crest-trough correlation r. A strong correlation between rogue wave probability p and r is demonstrated (Figure 3.3 and Figure 3.11). Computing uncertainties for the rogue wave probability ensures that the variation of p is significant while varying the sea state parameters. This work also expands r's applicability from coastal areas to offshore locations with some simplifications by utilizing time series of hourly wave height maxima H_{max} and significant wave height H_s, which are more readily available. This allowed us to extend r as a prediction parameter for the full domain of the NEP regional wave model. Although, the analysis using H_{max} limits the number of rogue waves per record to one so rogue wave occurrences are under represented, the total number of rogue waves increases from 160 waves where $H/H_s > 2$ to 15000 and from 22 waves where $H/H_s > 2.2$ to 3600. This greatly increases the sample size of rogue events and thus the confidence in the relationship with r in the present domain. r is also shown to be easily computed by a routine wave forecast and

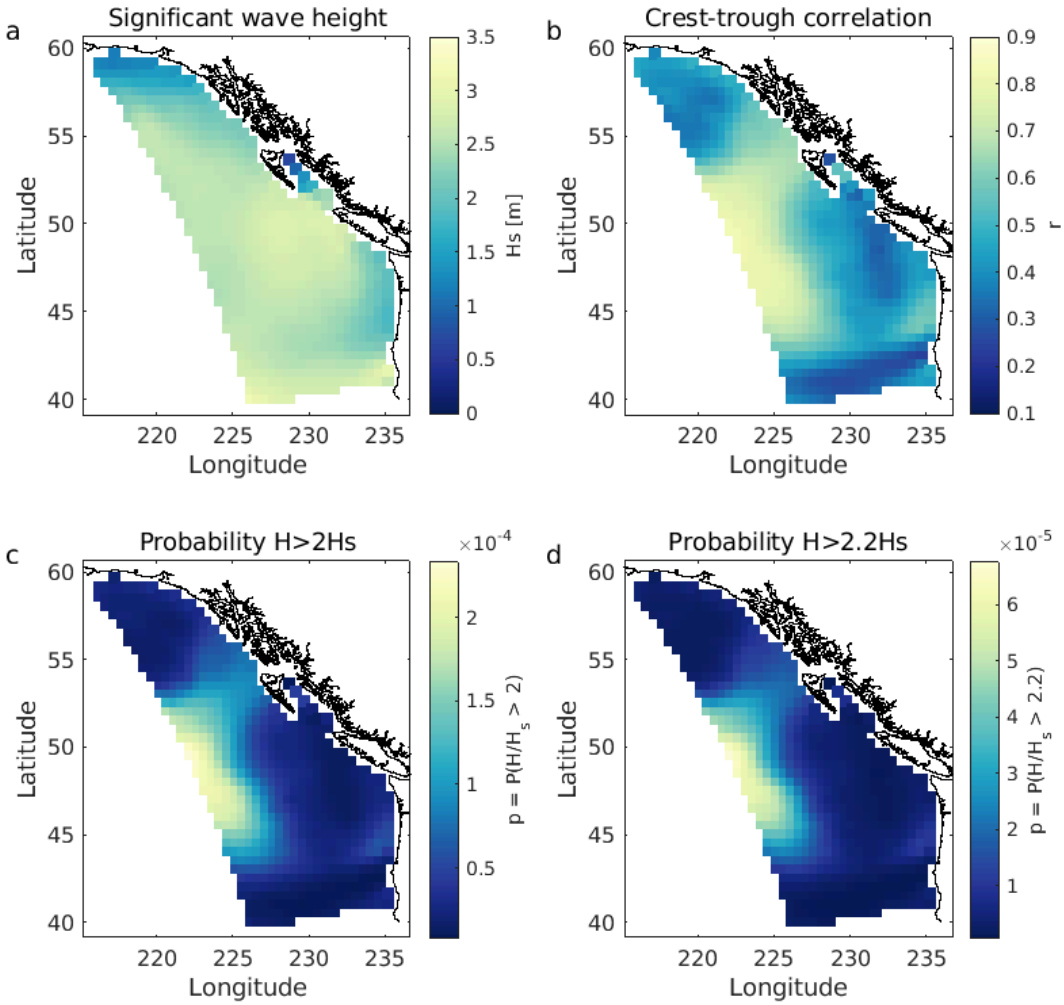

Figure 3.16: Example forecast output for March 22, 2020 00h UTC of **a)** significant wave height H_s,**b)** crest-trough correlation r, **c)** probability of a wave exceeding $2H_s$ and **d)** probability of a wave exceeding $2.2H_s$.

along side a forecast for H_s provides a complete risk assessment, where overlapping areas of high H_s and r pose the greatest risk (Figure 3.16).

It is reasonable that r is linked with rogue wave probability as it is an estimate of the auto-correlation between the crest heights and the trough depths. Therefore, in narrow banded seas the crests and successive troughs are approximately the same size resulting in high r values. Narrow banded seas also maintain wave groups and rogue waves are more likely to form in 'groupy' sea states (Gemmrich and Thomson, 2017). Therefore, it is also unsurprising a correlation between p and σ_N was found, where p increases with high r and low σ_N. However, r is an overall more robust prediction parameter based on the predictive power and forecast ability. This work focused on extreme wave heights and not extreme crest heights due to the records available, as there is no maximum wave crest height recorded by the MEDS monitoring buoys. It should be noted that by definition wave crests are unaffected by crest-trough correlation. This was verified in the surface elevation data set and evidenced in Figure 3.17, which demonstrates the correlation between r and the probability of rogue crests is low, as well as the predictive power. Some concern was raised about the binning of the wave parameters being biased using approximately equal number observations in each bin. Therefore, the correlation between p and r was tested using an adaptive bin pruning algorithm on the sea surface elevation records, which allows one to determine whether two bins are significantly different (Häfner, 2021). One starts with a high number of bins and merges non-significant bins until the significant optimal binning is left. Applying the adaptive binning the final number of bins for $p = P(H/Hs > 2)$ was 6 (same as before but with slightly different bin edges), and 3 bins for $p = P(H/Hs > 2.2)$. Regardless of the binning, there is similar correlation between p and r and consistent predictive power of 1.1 (Figure 3.18).

While there is plenty of numerical and experimental evidence for modulation instability contributing to extreme waves in wave tanks and simulations, the relative importance of this process in the real ocean is limited. This is evidenced by the low correlation of BFI on rogue wave probability in Figure 3.3. The limited relationship between BFI and rogue wave occurrence supports a probabilistic prediction for rogue waves based on random linear superposition using crest-trough correlation and should replace forecasts based on modulation instability through BFI.

Due to the availability, ease and cost effectiveness, this work uses buoy measurements at single points in space to assess rogue wave probability. There are few caveats that come with these point source measurements. The first being that al-

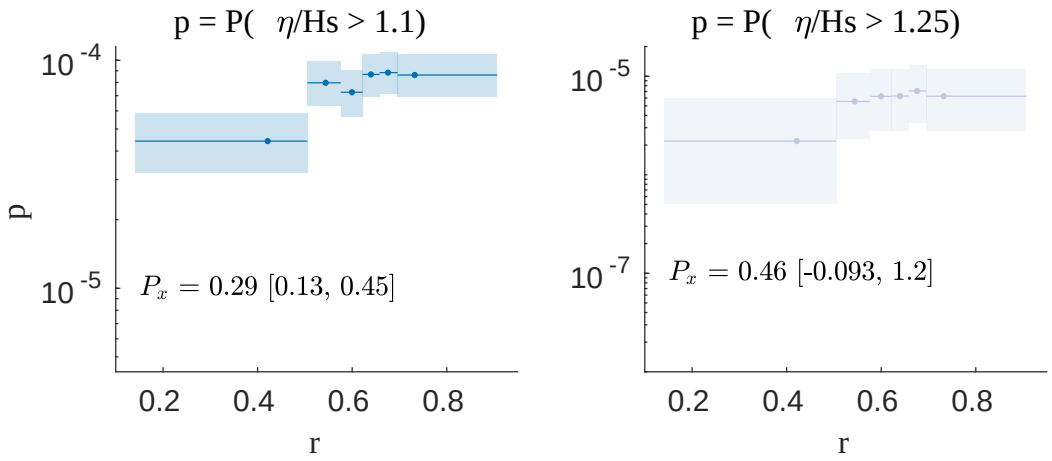

Figure 3.17: Variation in rogue crest probability for $p = P(\eta/H_s > 1.1)$ and $p = P(\eta/H_s > 1.25)$ with crest-trough correlation. The predictive power is listed in the plots.

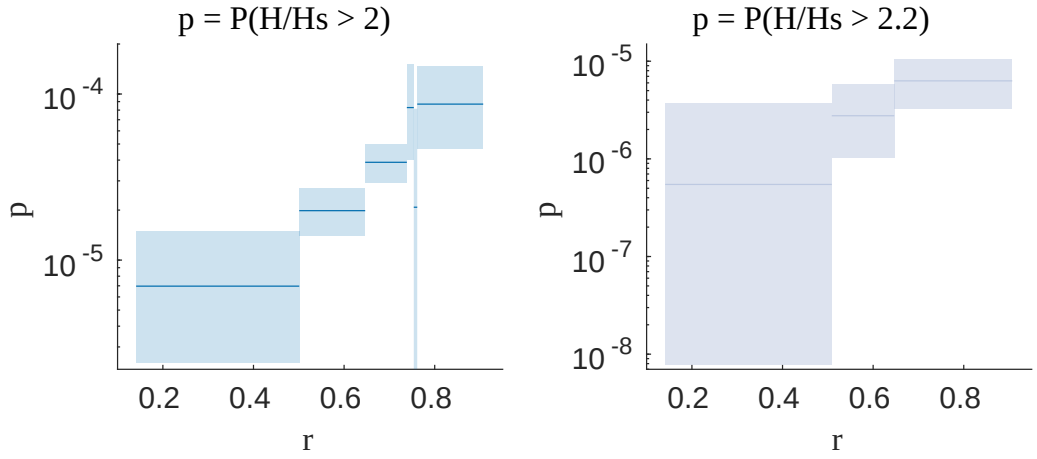

Figure 3.18: Variation in rogue wave probability for $p = P(H/H_s > 2)$ and $p = P(H/H_s > 2.2)$ with crest-trough correlation using an adaptive binning technique.

though surface-following buoys provide reliable estimates of H_s and T_p, they tend to underestimate the crest height of individual waves and overestimate depth of the trough (Casas-Prat and Holthuijsen, 2010). Consequently, the linearization of the surface elevation by the buoys does not affect the distribution of waves heights and the ratio of a wave height to H_s. Therefore this effect will not be influential to the analysis of extreme waves presented here. However, if a similar analysis is pursued for extreme crest heights, rogue crests will be potentially under represented. Another consequence of point source measurements is the maximum wave height over an area in a given length of time will be larger than the maximum wave height at a single point (Forristall, 2006; Fedele et al., 2013). This is due to ocean waves being dispersive and directionally spread. Ultimately ships and off shore platforms are interested in the probability of a rogue wave occurring in their vicinity and it is only in narrow spatially spread sea states that point measurements are exact in predicting the maxima at all points in space. Therefore, the probability of encountering a rogue wave derived from point source measurements is essentially a lower bound for the probability for the surrounding area. Stereo-video observations have been applied to evaluate the changes in the likelihood of oceanic rogue waves, where in an analysis of extreme crests by Benetazzo et al. (2017) the probability of gathering rogue waves in space and time was found to be at least one order of magnitude larger than the probability restricting the analysis to time only. Lastly, buoy measurement do not capture any information of the surrounding seas which would fully elucidate how rogue waves are generated. Theoretically, stereo-video processing techniques like the Wave Acquisition Stereo System (WASS) could be used to elucidate the dynamics of rogue waves in space and time, namely to evaluate if rogue waves are a single wave that grows over time or a random summation of steep waves (Fedele et al., 2013). However, practically there are some challenges. To evaluate the modulation instability in the ocean, one would have to have a sufficient field of view to capture the gradual build up of wave height. With dominant wavelengths in the ocean ranging from 50-250 m and the Benjamin Feir timescale on the order of 10 wave periods (Janssen, 2003), this becomes challenging to capture even with the present 30 m by 1000 m collection area of WASS. Stereo-video data acquisition also come with a higher price tag, higher data storage requirements and more challenging installation and post processing.

Rogue waves being a random, unpredictable, and scarce phenomenon contribute greatly to the difficulty of effectively studying them. It should also be noted that the relative importance of a generation mechanism of rogue waves is not necessarily

the same everywhere and therefore the same indication parameters might not be as effective in all regions (Cattrell et al., 2018; Adcock and Taylor, 2014). For example the high rogue wave occurrence in the southern California Bight are likely generated by a complex crossing wave fields with large directional spread, whereas in the semi-enclosed seas in Alaska, tidal currents likely play a contributing role. These numerous variations of mechanisms, various conditions and regional differences make studying rogue waves interesting and keeps research ongoing about what takes place in the ocean.

4 Case Study

To evaluate the NEP wave models ability to forecast large storms a case study was performed of the storm event that occurred October 21-22, 2021. This storm made news headlines due to the MV Zim Kingston container ship losing over 100 containers overboard at the mouth of the Juan de Fuca Strait and subsequently catching fire while at anchorage at Constance Bank, just off Victoria. Potential contributing factors to the disaster were evaluated for its approximate location, namely sea height, rogue wave risk, and currents. The model H_s was adequately forecast 24h prior to the storm peak where the winds from HRDPS were accurately forecast. The significant wave height at the mouth of the Juan de Fuca had maximum of 5.7 m. Based on the peak crest-trough correlation the probability of a wave exceeding $2H_s$ was once every 20h. The current conditions were unfavourable as at the storm peak the tidal current was ebbing therefore opposing the incoming wave field.

4.1 Introduction

The storm was caused by pressure system with a central low of 952 hPa from the Pacific moving towards the coast bringing strong winds and large seas, entering the model domain at approximately 21/10/2021 6H and dissipating by 22/10/2021 18H. The maximum wind speed and maximum wave height recorded by the open ocean buoy C46036 was 20.3 m/s and 9.5 m, respectively. It has been reported that the MV Zim Kingston, travelling from South Korea, was drifting at the mouth of the Juan de Fuca Strait when extreme weather caused an excessive listing of the vessel resulting in the collapse of containers, which were lost overboard (`http://www.tsb.gc.ca/eng/enquetes-investigations/marine/2021/M21P0297/M21P0297.html`). A total of 109 containers were lost, two of which contained hazardous chemicals and as of

01/03/2022 only 4 containers have been salvaged with some floating as far as Cape Scott Provincial Park on Vancouver Islands Northwest coast. While the ship was at an anchorage at Constance Bank, just off Victoria, awaiting a berthing slot at the port of Vancouver the vessel reported a fire in some of its damaged containers carrying dangerous goods. The Transportation Safety Board of Canada (TSB) is performing a class 3 investigation, which are generally completed within 450 days. At the time of this analysis an exact location of where the vessel lost the container has not been published, however the automatic identification system (AIS) data published by NOAA (`https://coast.noaa.gov/htdata/CMSP/AISDataHandler/2021/index.html`) indicates the ship drifting at approximately 125.61°W, 48.36°N outside the Juan de Fuca as seen in Figure 4.1, which is consistent with the TSP report and therefore the best estimate for the location of the excessive listing.

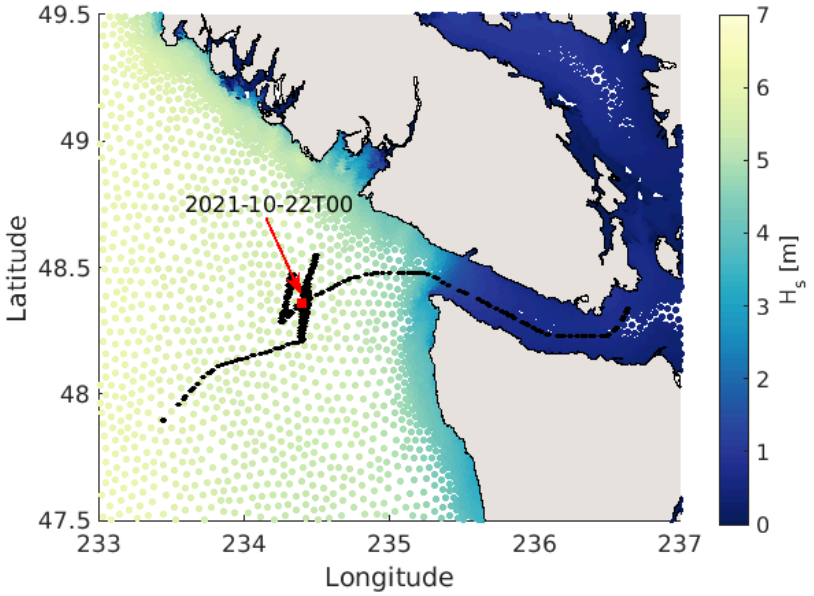

Figure 4.1: AIS data in black for the MV Zim Kingston from 21/10/2021-22/10/2021 with the model output of H_s for 22/10/2021 in colour. The red square marks the approximate location (125.61°W, 48.36°N) where the ship lost the containers.

4.2 Methods

The regional wave model was run from 15/10/2021 to 01/11/2021 with the calibration described in Chapter 2. To evaluate the models forecast with lead time to the storm peak, 48h forecasts were run every 6h with the winds being updated every 6h and the boundary conditions every 12h. For the rogue wave risk analy-

sis the wave spectra were output every 1/4°, south of 53°N within the model domain and r was computed as in Equation (3.8). The sea level and currents for the Juan de Fuca Strait were generated from the modelling tool WebTide from the Bedford Institute of Oceanography (`https://www.bio.gc.ca/science/research-recherche/ocean/webtide/index-en.php`).

For the forecast evaluation three buoys, C46036, C46206 and C46207 in the storms reach were used for comparison (Figure 2.1a). The models forecast of H_s, the frequency spectral shape and the peak period, T_p at the storms peak were evaluated, where the weighted peak period calculation of Equation (3.7) was used for T_p.

4.3 Results

The wave forecast for the storm performs relatively well with the main discrepancies occurring at the peak of the storm with the largest error in the open ocean. Figure 4.2 shows the time series of the "best" modelled data where the wind is updated every 6h and boundary every 12h. The model's 48h forecasts of H_s are plotted in Figure 4.3a where one can see the projected H_s at C46207 and C46206 are well represented by the model. However, the storm peak is greatly underestimated at C46036. The forecast does improve with decreasing lead time to the storm, but still misses the peak. This can be attributed to the winds seen in Figure 4.3b in the top panel, which shows a bust in the wind forecast by HRDPS. To better visualize the relationship between the wind and wave forecast the bias for each 48 hour forecast leading up to the storm peak (vertical dashed lines in Figure 4.3) was calculated. The bias is calculated as the modelled data minus the observation at the storm peak, therefore a positive bias indicates the model is overestimating and a negative bias indicates an underestimation. The H_s and wind speed biases are plotted in Figure 4.4, where the trend in the bias in the wind corresponds to the trend in the H_s bias. The bust in the wind forecast at the open ocean buoy is due to the low pressure system being forecast slightly south than in reality. This is verified by the ERA5 reanalysis data of the wind field on 21/10/2021 20H (Hersbach et al., 2018). The HRDPS wind forecast does approach the ERA5 wind field as shown by the difference between the ERA5 wind field and the wind field for October 21, 2021 20H from the October 21, 2021 00H model forecast and the October 21, 2021 18H forecast (Figure 4.5). Overall, the model forecasts the storm well up to 24h of lead time where the winds are accurately forecast.

The model produces good estimates of T_p at the height of the storm and the frequency spectra spans the same range of observed frequencies (Figure 4.6). However, the model does not capture the secondary wind wave peak at approximately 0.1 Hz. Spectra at the storm peak from the model and the buoys and the respective peak periods are plotted in Figure 4.6.

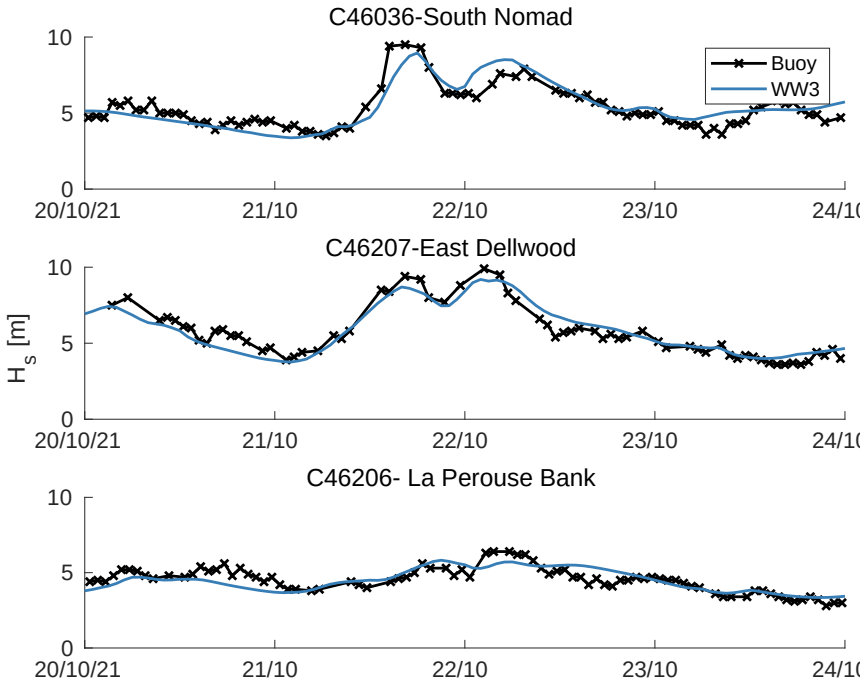

Figure 4.2: Time series of H_s at buoys C46036, C46207 and C46206 during the storm.

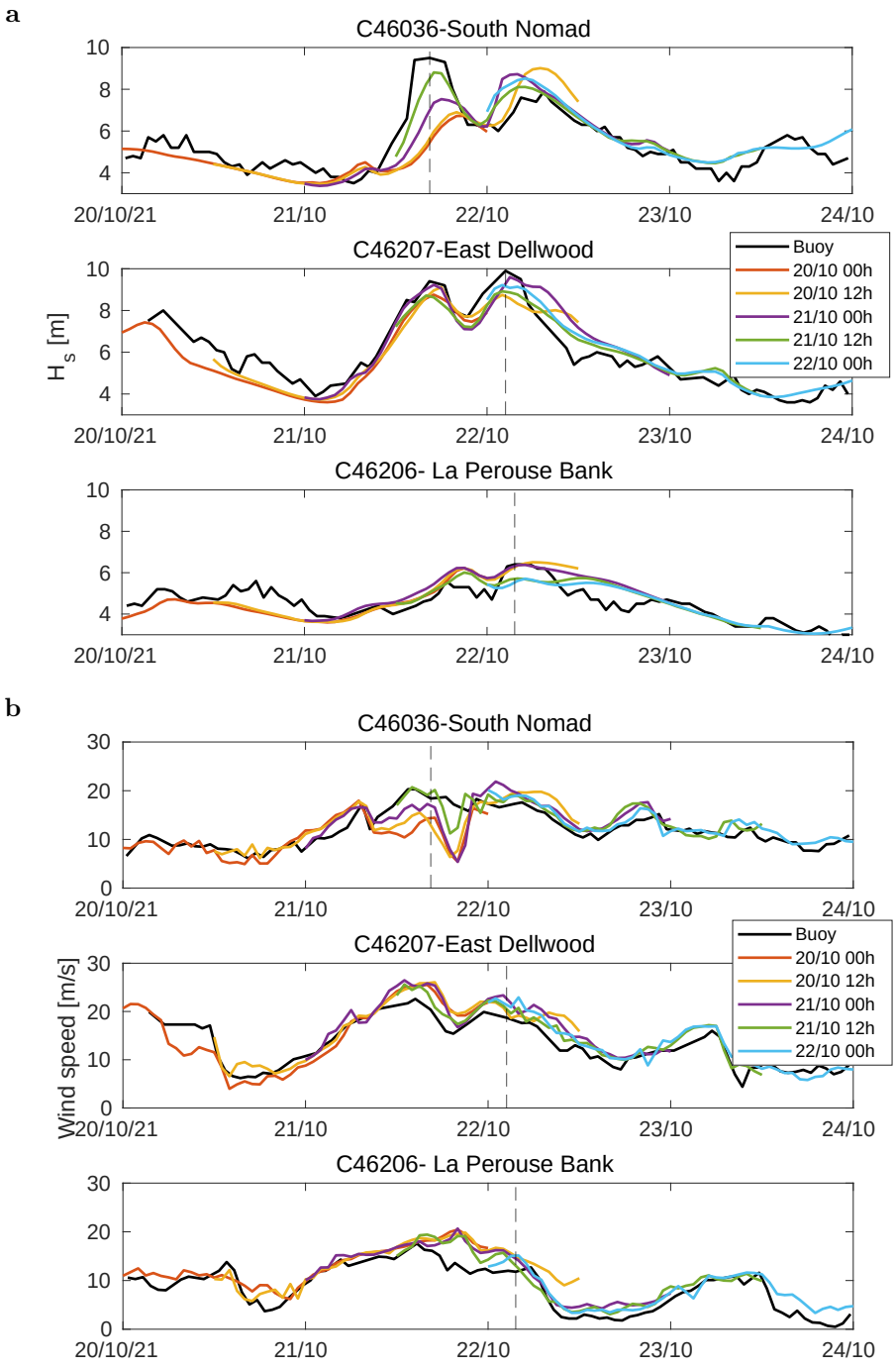

Figure 4.3: 48h forecasts where the wind is updated every 6h and the boundary conditions every 12h. Only every other 48h forecast is plotted for **a)** H_s and **b)** wind speed. Dashed vertical lines corresponds to times of peak H_s reported by the respective buoys.

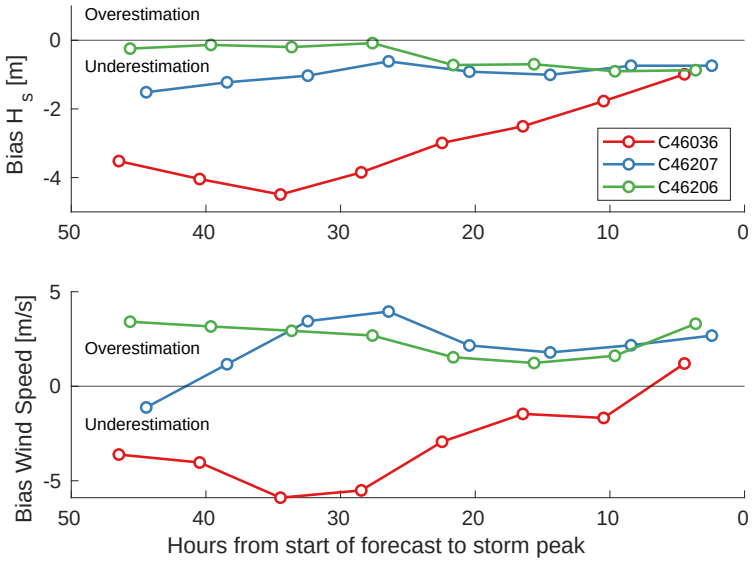

Figure 4.4: The bias (model output - observation) at the storm peak for each 48 hour forecast leading up to the storm peak for each buoy.

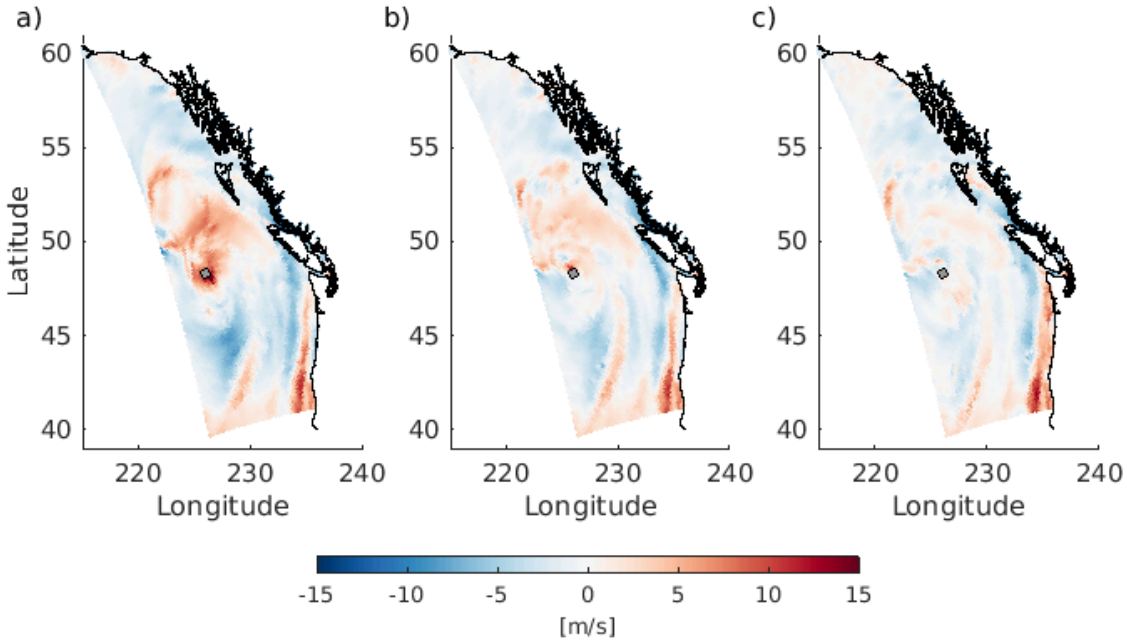

Figure 4.5: Difference between ERA5 reanalysis wind field and HRDPS forecast wind field wind for October 21, 2021 20H from the **a)** October 21, 2021 00H model forecast, **b)** October 21, 2021 12H model forecast, and **c)** October 21, 2021 18H model forecast. Positive values mean the ERA5 wind speed is larger than what was predicted by the HRDPS. Grey marker is location of buoy C46036.

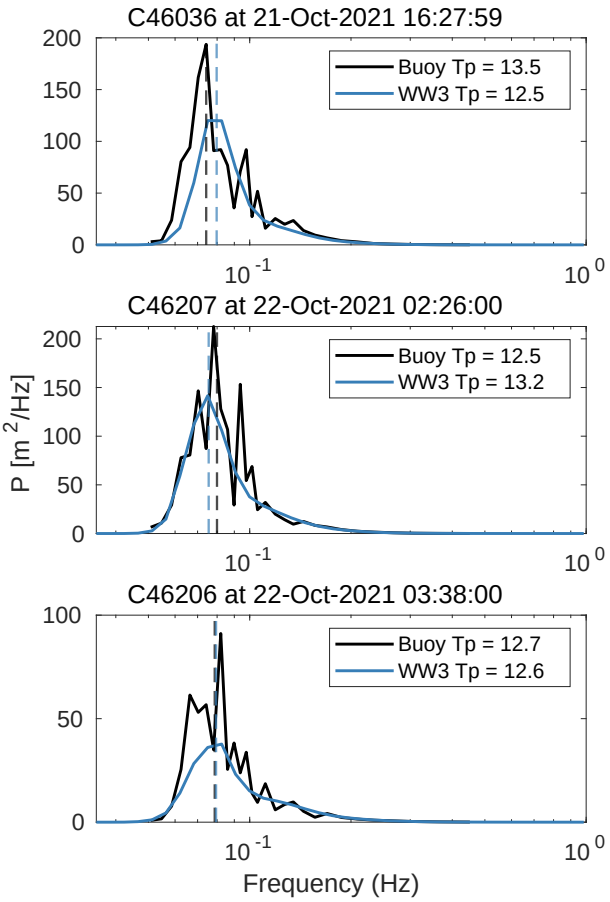

Figure 4.6: Wave frequency spectra at the storm peak as marked in Figure 4.3 from buoys C46036, C46207 and C46206 compared to the model output, with vertical dashed lines at the respective peak periods.

The Neah Bay buoy 46087 was unfortunately non-operational at the time of the storm so an observational data set of the storm near the mouth of the Juan de Fuca is unavailable. Regardless, Figure 4.7 shows the time series of the modelled H_s and r at the mouth of the Juan de Fuca (125.3 °W, 48.4 °N) which would be similar condition that the MV Zim Kingston would have been subject to although the exact location of the ship when it lost the containers is unknown. Also included is the 48h forecast of H_s and r from October 21, 2021 6H which is the forecast 24 hours prior to the modelled storm peak. The max H_s and r during the storm at this location was 5.7 m and 0.73, respectively. Using the fitted equations for rogue wave probability in Equation (3.12) an r value of 0.73 corresponds to a rogue wave $H > 2H_s$ every 20h and $H > 2.2H_s$ every 3.5 days, assuming an average wave period of 10 s. This evaluation does not attribute this container ship disaster to rogue waves in any way

but simply presents information that could be available in future forecasts if this probabilistic prediction for rogue waves is adopted.

Other influential factors in the Juan de Fuca are strong tidal currents. At the mouth of the Juan de Fuca (125.61 °W, 48.36 °N) the maximum current speed is 0.15 m/s during the storm and moving into the strait (124.50 °W, 48.45 °N) the maximum current speed increases to 0.79 m/s (Figure 4.8). These maxima correspond to ebbing tides meaning that the dominant wave direction and current speed are opposing which can be a dangerous combination due to wave steepening.

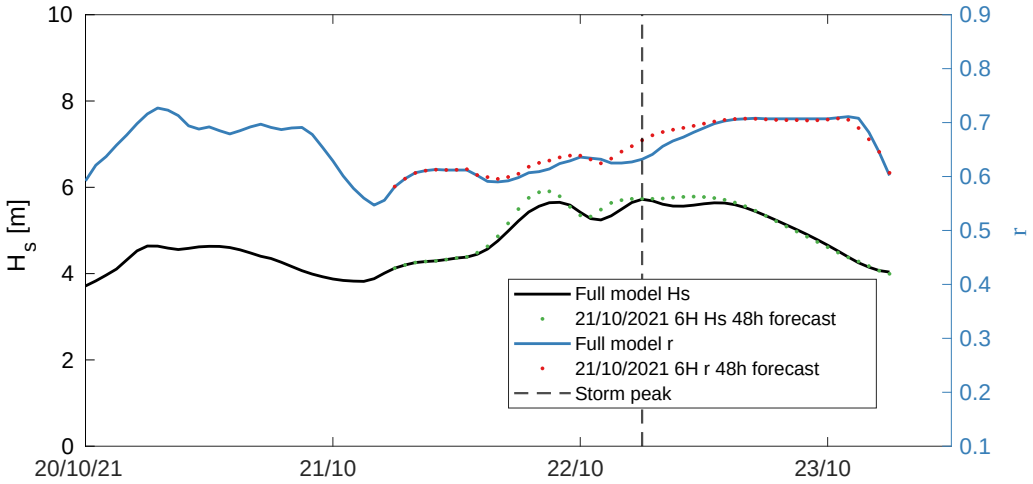

Figure 4.7: Time series of H_s and crest-trough correlation r outside the Juan de Fuca Strait (125.61 °W, 48.36 °N). Vertical dashed line is the storm peak from the model H_s output. Dotted lines are the H_s and r forecasts from 21/10/2021 6H which is approximately 24h prior to the storm peak.

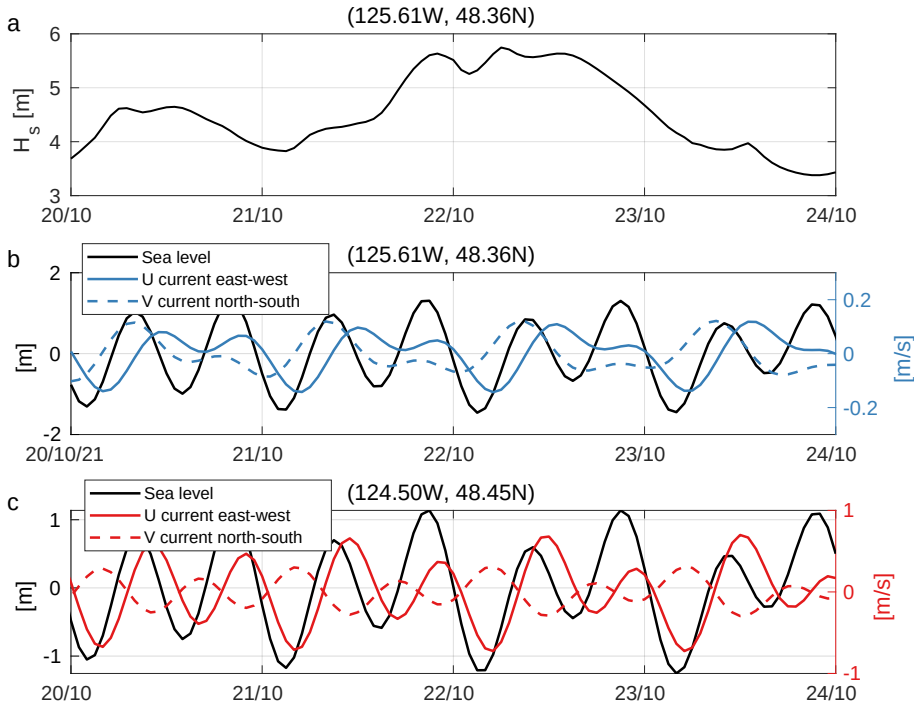

Figure 4.8: Time series from 20-24 October, 2021 of **a)** H_s from the wave model and **b)** the mean sea level and directional current outside the mouth of the Juan de Fuca Strait at 125.61 °W, 48.36 °N. **c)** The mean sea level and currents for a location in the Juan de Fuca Strait at 124.5 °W, 48.45 °N.

4.4 Discussion

The storm on Oct 21-22, 2021 in the Northeast Pacific, that resulted in the MV Zim Kingston container ship loosing over 100 containers, was modelled with the WW3 region wave model. The goal of this analysis is not to attribute the incident to a specific cause, but to evaluate the models capacity to forecast large storms and to identify potentially contributing factors to the incident.

The model performs well for this storm, despite the common occurrance of wave models missing the peaks of storms (Cavaleri, 2009). The output of H_s was adequately forecast 24h prior to the storm peak where the winds from HRDPS were accurately forecast. This is based of comparison with three buoys in the storm track. A large underestimation of H_s at the storm peak at one of the buoys was due to inaccurate localization of the wind, emphasizing the importance of the wind forecast for wave forecasting. A qualitative comparison of the frequency spectra at the storm peak at the buoy locations show the frequency spectrum span the same range and have

comparable peak periods however, the model does not capture the secondary wind wave peak. Currents in the vicinity of the mouth of Juan de Fuca Strait are also assessed during the storm and found currents up to 0.15 m/s, increasing to 0.8 m/s inside the strait. High current speeds combined with opposing large waves and winds would increase the risk level in that area (Zippel and Thomson, 2017; Saetra et al., 2021). A sample rogue wave forecast using crest-trough correlation was produced for the storm at the mouth of the strait, which shows an increasing rogue wave probability during storm development and continuing even after storm dispersal (Figure 4.7). This case study demonstrates the capacity of the regional model to forecast large storms and exemplifies a possible rogue wave forecast through the output of crest-trough correlation.

5 Summary and Future work

As the main discussions for the wave model, rogue wave analysis and storm case study are located in their respective chapter, this chapter provides a brief summary of the results and then focuses on future work stemming from this research due to the present limitations.

5.1 Summary

In this work a rogue wave probability forecast was developed from the spectral output of a standard wave model for the NEP. Firstly, a WW3 spectral wave model was calibrated for the Northeast Pacific Ocean to provide a higher resolution forecast to the region in comparison to the relatively low resolution 1/4° global model. The high resolution unstructured regional model (5 km offshore to 1 km nearshore) has been implemented into ECCC wave forecast as of December 2021. The model achieves high correlation and low bias when compared to long time series of buoy observations of significant wave height and wave period measurements. While the wave model is calibrated to the bulk of the observations, which consist of moderate to low sea states, the model does produce an accurate forecast for a large storm where the winds are adequately predicted. For the rogue wave analysis, long term sea surface elevation records from research buoys and hourly bulk statistics from monitoring buoys were compiled within the models domain. From these observations, sea state parameters that have been proposed to be linked to rogue wave generation were assessed as indicators. Among others, these parameters included BFI, stemming from the modulation instability effect and crest-trough correlation, proposed as a front runner by Häfner et al. (2021b). Crest-trough correlation showed the highest correlation with rogue wave probability and is consistent to what is reported in Häfner et al.

(2021b). Furthermore, BFI demonstrated little predictive power in the real ocean data. Crest-trough correlation was then implemented into the model output and was found to be moderately well predicted. The correlation of rogue wave probability with crest-trough correlation over BFI supports a probabilistic prediction of rogue waves based on random linear superposition and should replace forecasts based on modulation instability. In the case study of a large fall storm event, a potential rogue wave forecast that could be provided for such events was demonstrated.

5.2 Future Work

5.2.1 Wave model development

The regional model improves the wave modelling capacity for the NEP, due to the high resolution compared to the global model. Future model development could include incorporating current forcing into the model to better represent the sea state in tidally effected areas and at the mouths of rivers. Opposing waves and currents are known to be important in tidal inlets (Rapizo et al., 2017). That is, the wavelength will increase for waves propagating into a current heading in the same direction and shorten for waves opposing a current and the wave amplitude will be modulated. As demonstrated in the case study, current within the model domain such as the Juan de Fuca Strait can be significant and current forcing has already been incorporated into the UBC-run 500 m Strait of Georgia model with success (Gemmrich and Pawlowicz, 2020). When the current velocity opposes and exceeds in magnitude one-quarter of the initial phase velocity of the waves, the wave energy can no longer be propagated against the current (Longuet-Higgins and Stewart, 1961). In reality the waves tend to break before this point is reached, but regardless currents will effect wave height distributions. This effect would be influential in the Strait of Georgia as the dominant wave period ranges from 0.5-8 s corresponds to phase speeds of 0.8-12.5 m/s and the velocity of the tidal stream can reach speeds of 0.75-1.3 m/s. Furthermore, the outflow of the Fraser River at its mouth tend to be around 1 m/s during slack water depending on wind speed and direction and can reach upwards of 2.5 m/s at low water (Thomson, 1981). To properly calibrate current forcing in the model reliable buoys in the Salish Sea would be required.

The resolution of the regional grid could also be increased further. Abdolali et al. (2020) demonstrated that with an unstructured grid and using the implicit solving

scheme resolutions can be on the order of tens of meters without inflating the computational cost. Theoretically, the NEP grid resolution could be increased, but to properly calibrate the model with such resolution a higher density of observations in the nearshore would be required. Having high resolution for high resolution sake is impractical if it is not verified with measurements, although increasing the resolution likely does not detract from the model performance as a wave forecast. There is also the possibility of defining higher resolution areas within the larger domain without nesting grids. This would allow the introduction of higher resolution in more populated areas like the surrounding waters of Vancouver Island and Hecate Strait.

5.2.2 Rogue wave prediction

In the domain presented crest-trough correlation r has proven to be a valuable parameter as a rogue wave risk predictor and should be considered to be included in standard wave forecasts. In the case of WW3 this would be a simple addition to the models calculations of spectral parameters.

Of course, rogue wave research can always be improved with more data and therefore a larger sample size of rogue events. This is what makes Häfner et al. (2021b) compilation of seas surface elevation data so valuable and important in identifying sea state parameters influential to rogue wave formation. This allowed us to use those results as a starting point for this analysis and corroborate the findings as this data set is not as extensive. The limited surface elevation records available is one of the main limitations of this work. Adapting the analysis to include hourly bulk statistics of the MEDS buoys to supplement the time series data sets was beneficial, but also came with assumptions that skews towards underestimating rogue wave occurrence. Further validation of r as a prediction parameter would require more sea surface elevation records over longer time periods. In this work no attempt was made to calibrate the wave model for r. Future work could include identifying WW3 parameterization settings that modify r with the purpose of improving the forecast for r and therefore the rogue wave forecast.

Bibliography

Abdolali, A., A. Roland, A. van der Westhuysen, J. Meixner, A. Chawla, T. J. Hesser, J. M. Smith, and M. D. Sikiric, 2020: Large-scale hurricane modeling using domain decomposition parallelization and implicit scheme implemented in WAVEWATCH III wave model. *Coastal Engineering*, **157**, 103 656, https://doi.org/10.1016/j.coastaleng.2020.103656.

Adcock, T. A. A., and P. H. Taylor, 2014: The physics of anomalous ('rogue') ocean waves. *Reports on Progress in Physics*, **77 (10)**, 105 901, https://doi.org/10.1088/0034-4885/77/10/105901.

Ardhuin, F., and Coauthors, 2010: Semiempirical dissipation source functions for ocean waves. Part I: Definition, calibration, and validation. *Journal of Physical Oceanography*, **40 (9)**, 1917–1941, https://doi.org/10.1175/2010jpo4324.1.

Battjes, J., and J. Janssen, 1978: Energy loss and set-up due to breaking random waves. *Proceedings of 16th Conference on Coastal Engineering*, 569–587.

Benetazzo, A., and Coauthors, 2017: On the shape and likelihood of oceanic rogue waves. *Scientific Reports*, **7 (1)**, https://doi.org/10.1038/s41598-017-07704-9.

Casas-Prat, M., and L. H. Holthuijsen, 2010: Short-term statistics of waves observed in deep water. *Journal of Geophysical Research*, **115 (C9)**, https://doi.org/10.1029/2009jc005742.

Cattrell, A. D., M. Srokosz, B. I. Moat, and R. Marsh, 2018: Can rogue waves be predicted using characteristic wave parameters? *Journal of Geophysical Research: Oceans*, **123 (8)**, 5624–5636, https://doi.org/10.1029/2018jc013958.

Cavaleri, L., 2009: Wave modeling—missing the peaks. *Journal of Physical Oceanography*, **39 (11)**, 2757–2778, https://doi.org/10.1175/2009jpo4067.1.

Cavaleri, L., F. Barbariol, and A. Benetazzo, 2020: Wind–wave modeling: Where we are, where to go. *Journal of Marine Science and Engineering*, **8 (4)**, 260, https://doi.org/10.3390/jmse8040260.

Cavaleri, L., F. Barbariol, A. Benetazzo, L. Bertotti, J.-R. Bidlot, P. Janssen, and N. Wedi, 2016: The draupner wave: A fresh look and the emerging view. *Journal of Geophysical Research: Oceans*, **121 (8)**, 6061–6075, https://doi.org/10.1002/2016jc011649.

Cavaleri, L., and Coauthors, 2018: Wave modelling in coastal and inner seas. *Progress in Oceanography*, **167**, 164–233, https://doi.org/10.1016/j.pocean.2018.03.010.

Didenkulova, I. I., A. V. Slunyaev, E. N. Pelinovsky, and C. Kharif, 2006: Freak waves in 2005. *Natural hazards and earth system sciences*, **6 (6)**, 1007–1015.

Donelan, M. A., and A.-K. Magnusson, 2017: The Making of the Andrea Wave and other Rogues. *Scientific Reports*, **7 (1)**, https://doi.org/10.1038/srep44124.

Dysthe, K., H. E. Krogstad, and P. Müller, 2008: Oceanic rogue waves. *Annual Review of Fluid Mechanics*, **40 (1)**, 287–310, https://doi.org/10.1146/annurev.fluid.40.111406.102203.

Eldeberky, Y., V. Polnikov, and J. Battjes, 1996: A statistical approach for modeling triad interactions in dispersive waves. *Coastal Engineering Proceedings*, **1 (25)**, https://doi.org/10.9753/icce.v25.%p, URL https://icce-ojs-tamu.tdl.org/icce/article/view/5289.

Fedele, F., A. Benetazzo, G. Gallego, P.-C. Shih, A. Yezzi, F. Barbariol, and F. Ardhuin, 2013: Space–time measurements of oceanic sea states. *Ocean Modelling*, **70**, 103–115, https://doi.org/10.1016/j.ocemod.2013.01.001.

Fedele, F., J. Brennan, S. P. de León, J. Dudley, and F. Dias, 2016: Real world ocean rogue waves explained without the modulational instability. *Scientific Reports*, **6 (1)**, https://doi.org/10.1038/srep27715.

Fenton, J., 1988: The numerical solution of steady water wave problems. *Computers & Geosciences*, **14 (3)**, 357–368, https://doi.org/10.1016/0098-3004(88)90066-0.

Forristall, G. Z., 2006: Maximum crest heights over an area and the air gap problem. *Volume 3: Safety and Reliability; Materials Technology; Douglas Faulkner Symposium on Reliability and Ultimate Strength of Marine Structures*, ASMEDC, https://doi.org/10.1115/omae2006-92022.

Gemmrich, J., and L. Cicon, 2022: Generation mechanism and prediction of an observed extreme rogue wave. *Scientific Reports*, **12 (1)**, https://doi.org/10.1038/s41598-022-05671-4.

Gemmrich, J., and C. Garrett, 2011: Dynamical and statistical explanations of observed occurrence rates of rogue waves. *Natural Hazards and Earth System Sciences*, **11 (5)**, 1437–1446, https://doi.org/10.5194/nhess-11-1437-2011.

Gemmrich, J., and R. Pawlowicz, 2020: Wind waves in the strait of georgia. *Atmosphere-Ocean*, **58 (2)**, 79–97, https://doi.org/10.1080/07055900.2020.1735989.

Gemmrich, J., and J. Thomson, 2017: Observations of the shape and group dynamics of rogue waves. *Geophysical Research Letters*, https://doi.org/10.1002/2016gl072398.

Gorman, R. M., and H. J. Oliver, 2018: Automated model optimisation using the Cylc workflow engine (Cyclops v1.0). *Geoscientific Model Development*, **11 (6)**, 2153–2173, https://doi.org/10.5194/gmd-11-2153-2018.

Hasselmann, K., and Coauthors, 1973: Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). *Hydraulic Engineering Reports*, Deutches Hydrographisches Institut.

Hasselmann, S., K. Hasselmann, J. H. Allender, and T. P. Barnett, 1985: Computations and parameterizations of the nonlinear energy transfer in a gravity-wave specturm. Part II: Parameterizations of the nonlinear energy transfer for application in wave models. *Journal of physical oceanography*, **15 (11)**, 1378–1391.

Haver, S., 2004: A possible freak wave event measured at the Draupner jacket January 1 1995. *In Proceedings of Rogue Waves*.

Hersbach, H., and Coauthors, 2018: ERA5 hourly data on pressure levels from 1979 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, https://doi.org/10.24381/cds.bd0915c6, (Accessed on 15 -Feb-2022),.

Hiles, C., A. D. de Andres, S. Beatty, and B. Buckham, 2015: A case study on the matrix approach to WEC performance characterization. *In Proceedings of the 11th European Wave and Tidal Energy Conference, Nantes, France*.

Häfner, D., 2021: Bayesian Histograms v0.2 [Source code]. URL https://github.com/dionhaefner/bayesian-histograms, Github.

Häfner, D., J. Gemmrich, and M. Jochum, 2021a: FOWD: A free ocean wave dataset for data mining and machine learning. *Journal of Atmospheric and Oceanic Technology*, https://doi.org/10.1175/jtech-d-20-0185.1.

Häfner, D., J. Gemmrich, and M. Jochum, 2021b: Real-world rogue wave probabilities. *Scientific Reports*, **11 (1)**, https://doi.org/10.1038/s41598-021-89359-1.

Janssen, P., and J.-R. Bidlot, 2009: On the extension of the freak wave warning system and its verification. *ECMWF Technical Memo*, https://doi.org/10.21957/UF1SYBOG.

Janssen, P. A. E. M., 2003: Nonlinear four-wave interactions and freak waves. *Journal of Physical Oceanography*, **33 (4)**, 863–884.

Jensen, R. E., V. Swail, and R. H. Bouchard, 2021: Quantifying wave measurement differences in historical and present wave buoy systems. *Ocean Dynamics*, **71 (6-7)**, 731–755, https://doi.org/10.1007/s10236-021-01461-0.

Kharif, C., and E. Pelinovsky, 2003: Physical mechanisms of the rogue wave phenomenon. *European Journal of Mechanics - B/Fluids*, **22 (6)**, 603–634, https://doi.org/10.1016/j.euromechflu.2003.09.002.

Kinsman, B., 1965: *Wind waves, their generation and propagation on the ocean surface*. Prentice-Hall, Englewood Cliffs, N.J.

Kirezci, C., A. V. Babanin, and D. V. Chalikov, 2021: Modelling rogue waves in 1D wave trains with the JONSWAP spectrum, by means of the High Order Spectral Method and a fully nonlinear numerical model. *Ocean Engineering*, **231**, 108 715, https://doi.org/10.1016/j.oceaneng.2021.108715.

Longuet-Higgins, M. S., 1952: On the statistical distribution of the heights of sea waves. *Journal of Marine Research*, **11**, 245—-266.

Longuet-Higgins, M. S., 1980: On the distribution of the heights of sea waves: Some effects of nonlinearity and finite band width. *Journal of Geophysical Research*, **85 (C3)**, 1519, https://doi.org/10.1029/jc085ic03p01519.

Longuet-Higgins, M. S., 1984: Statistical properties of wave groups in a random sea state. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **312 (1521)**, 219–250, https://doi.org/10.1098/rsta.1984.0061.

Longuet-Higgins, M. S., and R. W. Stewart, 1961: The changes in amplitude of short gravity waves on steady non-uniform currents. *Journal of Fluid Mechanics*, **10 (04)**, 529, https://doi.org/10.1017/s0022112061000342.

Marine Environmental Data Section Archive, Ecosystem and Oceans Science, Department of Fisheries and Oceans Canada, 2021: URL https://meds-sdmm.dfo-mpo.gc.ca, data obtained on 2021/08/15.

Naess, A., 1985: On the distribution of crest to trough wave heights. *Ocean engineering*, **12 (3)**, 221–234.

NCEP, 2016: ETOPOI global relief model. URL https://www.ngdc.noaa.gov/mgg/global/global.html.

NDBC, 2021: National Oceanic and Atmospheric Administration's National Data Buoy Center. URL https://www.ndbc.noaa.gov/, Data obtained on 2021/08/15.

Rapizo, H., A. V. Babanin, D. Provis, and W. E. Rogers, 2017: Current-induced dissipation in spectral wave models. *Journal of Geophysical Research: Oceans*, **122 (3)**, 2205–2225, https://doi.org/10.1002/2016jc012367.

Roland, A., 2009: Development of WWM II: Spectral wave modelling on unstructured meshes. Ph.D. thesis, Technische Universit at Darmstadt, Instituteof Hydraulic and Water Resources Engineering.

Saetra, Ø., T. Halsne, A. Carrasco, Ø. Breivik, T. Pedersen, and K. H. Christensen, 2021: Intense interactions between ocean waves and currents observed in the Lofoten Maelstrom. *Journal of Physical Oceanography*, https://doi.org/10.1175/jpo-d-20-0290.1.

Saulnier, J.-B., A. Clément, A. F. de O. Falcão, T. Pontes, M. Prevosto, and P. Ricci, 2011: Wave groupiness and spectral bandwidth as relevant parameters for the performance assessment of wave energy converters. *Ocean Engineering*, **38 (1)**, 130–147, https://doi.org/10.1016/j.oceaneng.2010.10.002.

Serio, M., M. Onorato, A. R. Osborne, and P. A. E. M. Janssen, 2005: On the computation of the benjamin-feir index. *Il Nuovo Cimento C*, **28**, 893–903, https://doi.org/10.1393/ncc/i2005-10134-1.

Smith, J. M., T. Hesser, A. Roland, and M. Bryant, 2018: Validation of unstructured WAVEWATCHIII for nearshore waves. *Coastal Engineering Proceedings*, 55, https://doi.org/10.9753/icce.v36.waves.55.

Tayfun, M. A., 1990: Distribution of large wave heights. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, **116 (6)**, 686–707, https://doi.org/10.1061/(asce)0733-950x(1990)116:6(686).

Tayfun, M. A., and F. Fedele, 2007: Wave-height distributions and nonlinear effects. *Ocean Engineering*, **34 (11-12)**, 1631–1649, https://doi.org/10.1016/j.oceaneng.2006.11.006.

The Wamdi Group, 1988: The WAM model-a third generation ocean wave prediction model. *Journal of Physical Oceanography*, **18 (12)**, 1775–1810, https://doi.org/10.1175/1520-0485(1988)018⟨1775:twmtgo⟩2.0.co;2.

The WAVEWATCH III ® Development Group (WW3DG), 2019: User manual and system documentation of WAVEWATCH III ® version 6.07. Tech. Note 333, NOAA/NWS/NCEP/MMAB. College Park, MD, USA, 465 pp. + Appendices.

Thomson, J., J. Gemmrich, W. E. Rogers, C. O. Collins, and F. Ardhuin, 2019: Wave groups observed in pancake sea ice. *Journal of Geophysical Research: Oceans*, **124 (11)**, 7400–7411, https://doi.org/10.1029/2019jc015354.

Thomson, R., 1981: Oceanography of the British Columbia Coast. *Canadian Special Publication of Fisheries and Aquatic Sciences*, **56**.

Tillotson, K., and P. D. Komar, 1997: The wave climate of the Pacific Northwest (Oregon and Washington): A comparison of data sources. *Journal of Coastal Research*, **13 (2)**, 440–452.

Tolman, H. L., 1994: Wind waves and moveable-bed bottom friction. *Journal of Physical Oceanography*, **24 (5)**, 994–1009, https://doi.org/10.1175/1520-0485(1994)024⟨0994:wwambb⟩2.0.co;2.

Young, I., 1995: The determination of confidence limits associated with estimates of the spectral peak frequency. *Ocean Engineering*, **22**, 669–686, https://doi.org/10.1016/0029-8018(95)00002-3.

Zakharov, V. E., 1968: Stability of periodic waves of finite amplitude on the surface of a deep fluid. *Journal of Applied Mechanics and Technical Physics*, **9 (2)**, 190–194, https://doi.org/10.1007/bf00913182.

Zippel, S., and J. Thomson, 2017: Surface wave breaking over sheared currents: Observations from the Mouth of the Columbia River. *Journal of Geophysical Research: Oceans*, **122 (4)**, 3311–3328, https://doi.org/10.1002/2016jc012498.