UTRECHT UNIVERSITY

METEOROLOGY, PHYSICAL OCEANOGRAPHY, AND CLIMATE

A MASTER THESIS

Influence of Tropical Cyclones on the Statistical Properties of Extreme Waves in the Tropical Indian Ocean

Author:
Nugrahinggil SUBASITA

Supervisor:
Prof. Dr. Huib DE SWART
and Dr. Andreas STERL

31st January 2018

Utrecht University

Abstract

In this MSc thesis, the climatology of ocean surface waves in the tropical Indian Ocean is investigated. Specific focus is on the probability distributions of the significant wave height, i.e. the mean height of the 1/3 of the waves over a 10-30 minutes interval, and of the mean wave period. Additionally, the characteristics of extreme waves, with heights that occur only once in e.g. 20 years or 50 years and which are generated by tropical cyclones, are analysed.

To that aim, numerical experiments were conducted by using the spectral wave model SWAN, which is forced by winds, that are partly based on data from the ECMWF atmospheric model. The resolution of these wind data is too coarse to properly account for tropical cyclones. The new aspect of this study is that the winds were locally adjusted for the presence of tropical cyclones by including a sub-grid model in which wind speeds of the cyclones are computed from atmospheric pressure data.

It is shown that a 16-year simulation of the SWAN model, forced by winds from the ECMWF model and corrected for the presence of tropical cyclones by means of the sub-grid model, yields a wave climate and statistics of extreme waves that agree with those observed at some buoys. Moreover, they also agree with output of a high resolution atmosphere-ocean wave model, which is much more expensive than the present model. Also, in depth analyses and interpretations of the results are given.

Contents

1 Introduction

1.1 General Remarks

The surface of the ocean is in continuous, wave-like motion. The waves are generated by different forcing agents, such as tides and winds, and have a wide range of time scales (hours to years for tidal waves, seconds to minutes for wind waves (Munk [1951]; LeBlond and Mysak [1978]; Holthuijsen [2007])). Here, wind waves comprise both *sea* and *swell*[1]. An interesting difference between tidal waves and wind waves is that the former are regular and predictable, whilst the latter are quite irregular random features (see Figure 1).
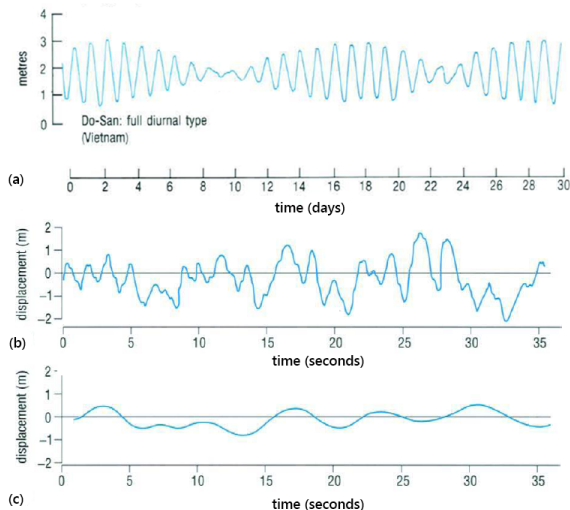

Figure 1: **(a)** Time series of the tide at Do-San, Vietnam, which has a typical period of a day (diurnal). **(b)** The measured surface elevation of wind waves that are under direct influence of wind, which is classified as a sea. **(c)** Same as **(b)**, but for swell, i.e. waves that are no longer directly forced by the wind. Compared with sea, swell usually has a longer period and more regular. Both have typical periods of seconds. Source: Brown [2013] (with modification).

Wind waves are therefore described in terms of their statistical properties, most commonly by means of a wave spectrum (Pierson and Moskowitz [1964]; Hasselmann et al. [1973]; Holthuijsen [2007] and references herein), which gives the distribution of the variance of the surface elevation over the wave frequencies and directions. From the wave spectrum, time-averaged variables can be derived, e wave height, the wave period, and the wave direction. Here the time averages are typically evaluated over 10 minutes intervals, which are long enough to contain many waves and short enough to allow the assumption of statistically stationary conditions (i.e averages are independent of time). On longer time scales, the statistic of properties of wind waves definitely changes in particular due to changes in the wind forcing. The modern prediction models of wind waves are based on the evolution of the wave spectrum due to input wind forcing (WAM by The WAMDI Group [1988],Wavewatch III by Tolman [1991],SWAN by Booij et al. [1999]).

---

[1]a sea is a wave condition under influenced of local wind thus have an irregular pattern. And a swell is a wave condition as a result from distant weather system
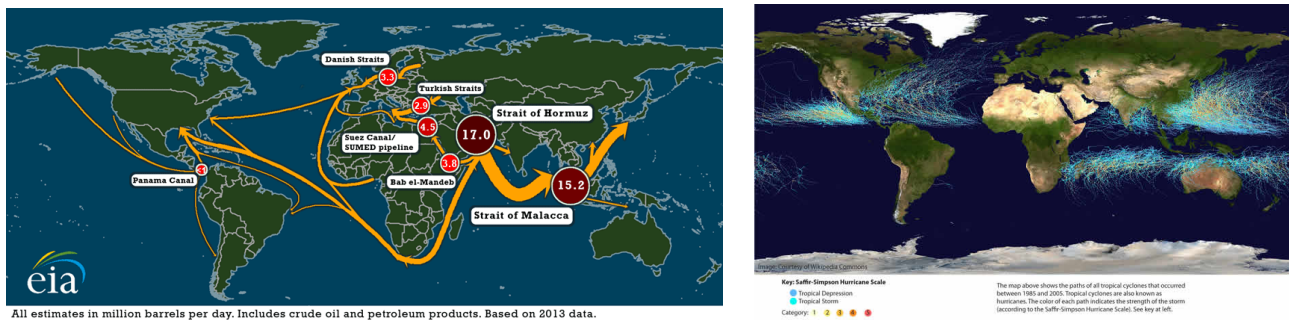
Figure 2: [**left**] Estimation of world's petroleum transportation based on 2013 data. The largest amount of petroleum being transported is from the Middle-East to Asia via tropical Indian Ocean. Source: `http://www.hellenicshippingnews.com`. [**right**] Map of the tracks of all tropical cyclones which formed worldwide from 1985 to 2005. The points show the locations of the storms at six-hourly intervals and the colour shows the intensity (blue-less intense and red-very intense). The North-western Pacific Ocean sees more tropical cyclone than the other basins, with a frequency of ± 27 TC/year, while the North Indian Ocean only has ± 5 TC/year (Joint Typhoon Warning Center [2009]). Source: `https://www.windows2universe.org/`

The focus of this thesis is on the climatology of wind waves and, in particular, on the occurrence of extreme waves, which are loosely defined as the highest wave during certain period, e.g. 20 years (A more precise definition will follow later). Clearly, gaining knowledge about extreme waves is important, e.g. for marine transportation (Soares and Teixeira [2001]), for safety and maintenance of offshore and coastal structures (Goda [1979]), for renewable energy (Falnes and Løvseth [1991]), and for future climate prediction. In that respect, note that approximately two-third of the world's petroleum production is shipped over sea (Rodrigue [2004]). Furthermore, the largest volume are transported from Arabian Peninsula to East Asia via the tropical Indian ocean (see Figure 2 left) in which, extreme waves are frequently observed due to occurrence of tropical cyclones (see Figure 2 right).

Obtaining such knowledge is difficult due to lack of observational data, especially in the Indian Ocean. Moreover, available output from wave prediction models, that cover at least 30 years time period, still have a coarse resolution, both spatially and temporally (Swail and Cox [2000]; Uppala et al. [2005]; Dee et al. [2011]). As a result, these models are not capable of resolving tropical cyclones that have smaller scales than the model resolution. This leads to the underestimation of predicting the extreme wave conditions in the tropical areas. Consequently, resolving a tropical cyclone into wind-waves simulation in a tropical area is very important to properly simulate extreme wave conditions.

1.2 Overview of Available Knowledge

The earliest version of a wave prediction model had been developed by Sverdrup and Munk [1949], which was later revised by Bretschneider [1957]. They made a purely empirical solution to estimate wave parameters from a given wind speed, fetch[2], wind duration, and wind direction. Nevertheless, these empirical models do not account for spatial variation of the wind field, thus they obviously fail to predict wave climate conditions, especially when swell occurs.

From a long period of wave observation, such as dozen of years, time series of wave parameters, such as significant wave height (H_s), defined as $4s$, where s^2 is wave variance, or mean wave period, can be analysed by means of statistics. By a statistical method, a probability density distribution of specific wave parameters can be estimated. Subsequently, statistics of the climate and statistics of extreme values need to be distinguished and analysed. Especially for H_s, extreme values are defined as any value that is higher than the lower bound of the high tail of its distribution ($H_{s,threshold}$). The distribution of extreme values thus differ from the distribution of the climate condition. Hence different approximations are commonly used to explain

Figure 3: Illustration of classification of wave parameter from its probability density distribution, here is for significant wave height (H_s), between climate condition and extreme condition. All H_s higher than $H_{s,threshold}$ are classified as extreme value. Adapted from Holthuijsen [2007].

the characteristic of the distribution. For example, in the application of m-year return period, the lower bound is adjusted as such, so that remainder corresponds to only one event during a continuous observation of m years. Hence, the return value of this period is the minimum threshold value to have this situation.

Caires and Sterl [2005] succeeded to examine the output of ERA-40 model, a coupled atmospheric-ocean wave model from European Centre for Medium-range Weather Forecast (ECMWF). They performed an extreme analysis for wind and wave condition as maps of the 100-year return period. ERA-interim is a newer model than ERA-40 in the ERA-series project. Although having improvement on the spatial resolution, the ERA-interim still has a coarse spatial resolution by 79 km ($\sim 0.70°$) for the atmospheric model and $1°$ for the ocean model (Dee et al. [2011]). This is improved comparing to $1.5°$ resolution of that in ERA-40 (Uppala et al. [2005]). Therefore, these
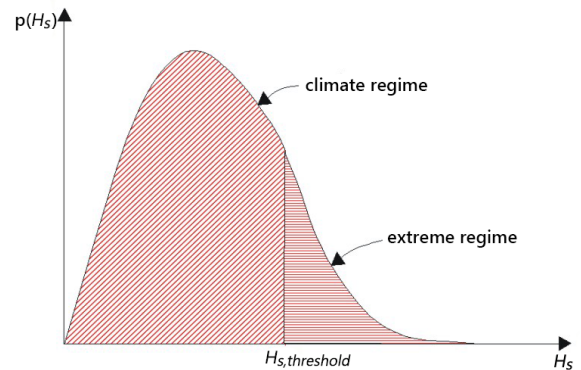
---

[2]the distance to the upwind coastline, see Holthuijsen [2007]

models are not able to correctly resolve the occurrences of tropical cyclones, that have the maximum wind speed in the radius of ∼30 km from its centre (Kimball and Mulekar [2004]). However, the ERA-interim model provides a relatively long time of reanalysis data from 1979 until recent events, which is important in determining the climate condition.

Timmermans et al. [2017] performed numerical simulations using a coupled atmospheric-wave model in high resolution of 0.25°. They used WaveWatch III (Tolman [1991]), instead of WAM (The WAMDI Group [1988]) that is being used in ERA models, as the wave model to simulate the wind waves. This method is able to resolve the importance of tropical cyclones, but it is computationally demanding. Occurrences of tropical cyclones is shown to have a significant impact on the 20 years return value of extreme waves from the distribution of significant wave height in the future climate condition (Timmermans et al. [2017]).

The capability of SWAN (Booij et al. [1999]) to simulate wind waves due to tropical cyclones has been tested and verified (Dietrich et al. [2012] and Drost et al. [2017]). However, Dietrich et al. [2012] only used the wind field retrieved from hindcasts of Hurricanes Katrina and Rita (2005). This wind fields were reproduced using a simple analytical model for a tropical cyclone that was developed by Holland [1980]. Hence, the character of the wind field is very local in regards of its centre. The SWAN wave model will be used for performing wave conditions in this thesis, since this model is available for free.

In conclusion, Timmermans et al. [2017] have shown by increasing resolution from 1°, also used in the ERA-interim, to 0.25°, the occurrences of tropical cyclones can be reproduced. Yet, it demands a higher computational capability by performing two models, i.e. an atmospheric model and an ocean wave model. On the other side, ERA-interim has provided reanalysis products of atmospheric-ocean properties, one of them being the near-surface wind speed (Dee et al. [2011]). However, it has a relatively large grid size (Berrisford et al. [2011]) and does not resolve tropical cyclone correctly. Using input of an empirical wind field due to tropical cyclone from Holland [1980], the extreme waves due to tropical cyclones can be reproduced (Dietrich et al. [2012]; Drost et al. [2017]). In this thesis, the ERA-interim wind products will be integrated with a sub-grid model which resolves wind conditions due to tropical cyclone. Then, the combination of wind fields will be used as an input for the SWAN wave model to generates wind waves with a higher resolution, that can take into account the effects of tropical cyclones. Consequently, instead of performing a couple models of atmospheric and wave, the simulation will be done by only using a wave model, but with modified wind input.

1.3   Research Questions

Based on previous considerations, the following three research questions will be addressed in this thesis:

1. *To what extent does the skill of the SWAN wave prediction model, fed by winds from the ERA-interim model, with regard to the wave characteristics in the tropical ocean, can be improved if winds are locally improved for the occurrence of tropical cyclones by means of a simple sub-grid model?*

   The wave parameters are produced using the spectral wave model, SWAN (Booij et al. [1999]). The input of a combination of the wind field from ERA-interim and that from sub-grid model for tropical cyclones will be used to generate the waves. The skill of the model will be determined based on the correlation factor and the amount of deviation which is measured by root-mean-square error and bias, between the model results and the observational data. The time window for the comparison will be chosen such that the observation contains the extreme waves due to tropical cyclones.

2. *In what way does tropical cyclones influence the shape of the wave spectra in the tropical Indian Ocean?*

   The wave spectra is recorded at the location of buoys for every time step. Evolution of the shape of wave spectra will be investigated with respect to their peaks behaviour, especially during the occurrence of the tropical cyclones. The spectrum is expected to have at least double peaks which are representative of seas and swell.

3. *How does the wave model account for tropical cyclones affect the return periods of 20-year and 50-year extreme waves, especially in the tropical Indian Ocean?*

   The results of wave parameters from the SWAN simulations will be considered for the extreme part only. The occurrences of extreme waves is extracted using the Peak Over Threshold method (Caires and Sterl [2005]). The threshold value in this method works as the lower bound of the extreme regime (see Figure 3). Subsequently, this distribution of extreme values which is extracted by this method will be fitted to a certain commonly used distribution, viz. the Generalized Pareto Distribution. From the fitted distribution, the return period of 20-year and 50-year extreme values will be estimated.

The contents of the forthcoming chapters are as follows. Chapter 2 presents the material and methods that are used in this thesis, including the design experiments and statistical method for the estimation of wave climate and extreme conditions. The model verification and analysis of the

wave climate and the occurrence of extreme conditions will be explained in Chapter 3. Chapter 3 also contains the effect of a tropical cyclone on the wave spectrum in the location of buoys and the comparison the results of the SWAN model to original ERA-interim results. The sensitivity in choosing threshold value for the extreme wave analysis, the comparison of the optimum setting to the other settings and some recommendations for future research are discussed in Chapter 4. Chapter 5 contains the conclusion. Additionally, Appendix A contains descriptions of some wave parameters. The input scripts used in the SWAN simulation are shown in Appendix B. Furthermore, some detail verifications are presented in Appendix C. Finally, Appendix D shows the time series of wave parameters during the occurrence of tropical cyclone Vongfong that will be useful for future research.

2 Material and Methods

2.1 Wind Waves in the Ocean

Surface ocean waves are defined as the up-and-down motion of the free surface of the ocean. They are generated by different forcing agents, such as wind, tides, and buoyancy input. Munk [1951] classified these waves according to their typical wave period. The wave categories range from capillary waves that have periods less than 0.1 second, to planetary waves that have period of months. This thesis focuses on waves with a period between 0.1 - 100 second which are mainly generated by wind (Holthuijsen [2007]). The main restoring force of these wind waves is gravity. Later, they are also referred as wind waves.
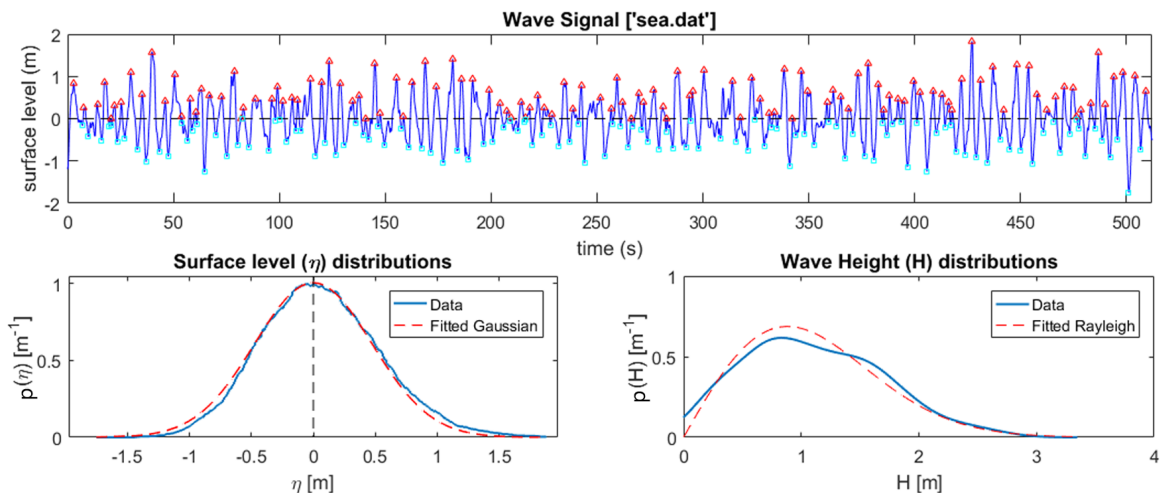

Figure 4: [**top**] Time series of measured sea surface level (blue line); crests and troughs based on up-crossing method are indicated by red-triangles and cyan-squares, respectively; black-dashed line indicates the mean water level. [**bottom-left**] Measured sea surface level distribution (blue) compared to a Gaussian distribution (red-dashed) and [**bottom-right**] distribution of observed wave height (blue) which is approximately following a Rayleigh distribution (red-dashed). Source: WAFO-group [2017]

As is shown in Figure 4, wind waves are to a certain organized features, e.g. clear crests and troughs can be identified. Yet, owing to their stochastic forcing, they show random behaviour, i.e. it is not possible to describe the future evolution of wave pattern in space and time. Therefore, statistical (or probabilistic) methods are required to analyse their properties.

In this context, three approaches are available to describe waves. The first is to use probability density distributions. It is done by assuming that variation of sea surface elevation (η) are statistically stationary, hence time averages can be made. Such conditions typically hold on a time scale

of 15-30 minutes (Holthuijsen [2007]). From the typical time scale, the probability density distributions of sea surface elevation is approximately following Gaussian shape (Figure 4, bottom-left panel). Meanwhile, the wave height distributions can be approximated by a Rayleigh distributions (see Figure 4, bottom-right panel).

The second approach is to describe waves in term of statistical moments. For example, the mean sea level is the average of sea surface elevation during observation. Furthermore, the most important parameter is the variance of wave height (s^2) since it is proportional to the amount of energy included by the waves (Holthuijsen [2007]).

The approach of averages can also be derived from the probability distribution. For instance, a significant wave height (H_s) defines as the average of the highest one-third of waves $(H_{1/3})$. The term of significant wave height is often used because it is equivalent to a wave height retrieved from visual estimation (Munk [1951]; Holthuijsen [2007]). Therefore, H_s can be written as

$$H_s = H_{1/3} = \frac{1}{N/3} \sum_{j=1}^{N/3} H_j.$$

Here H_j is a series of wave heights that is ranked based on height, from high waves to small waves. In terms of probability density distribution of wave height $(p(H))$, variable H_s is obtained by

$$H_s = H_{1/3} = \frac{\int_{H^*}^{\infty} Hp(H)dH}{\int_{H^*}^{\infty} p(H)dH} \quad \text{where} \int_{H^*}^{\infty} p(H)dH = \frac{1}{3}.$$
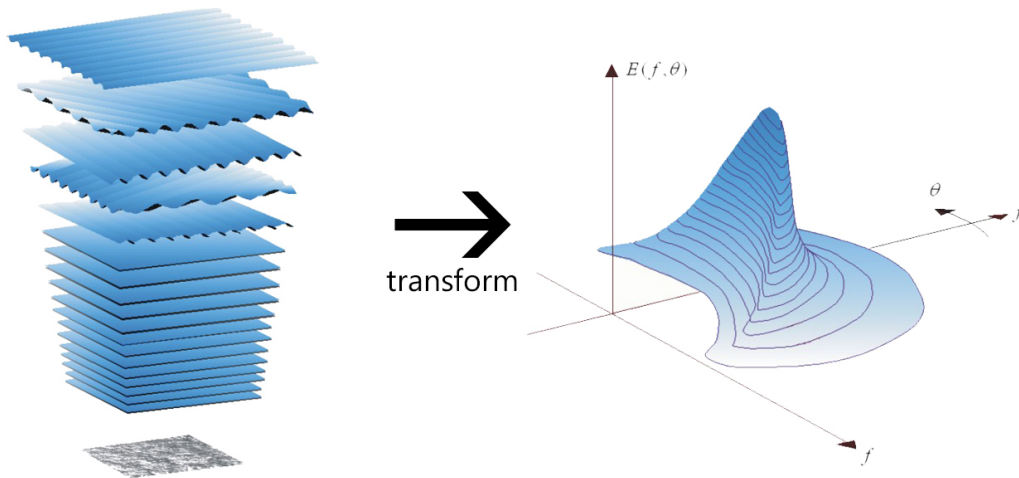

Figure 5:   [**left**] Random seas are approximated by a superposition of many harmonic waves, [**right**] and a frequency-directional wave spectrum $(E(f, \theta))$ is estimated based on obtained wave information from these series of harmonic waves. Source: Holthuijsen [2007]

The third approach is to describe the sea-state or state of random sea waves in term of wave spectra. The basic idea is that any wave profile of ocean waves can be constructed from a superposition of many different individual harmonic waves each having their own wave frequency (f), wave number (k), wave height (H), and wave direction (θ) (see Figure 5). The wave frequency (f) and wave number (k) are related through so-called dispersion relation, which for surface gravity waves in the absence of background current reads

$$\sigma = 2\pi f = \sqrt{gk\tanh kh}.$$

A 2-D wave spectra ($E(f,\theta)$) shows the distribution of wave variance per frequency and angle of direction (unit $m^2/H_z/\text{rad}$) as a function of wave frequency and wave direction. The wave frequency spectrum is retrieved through integration over all directions ($E(f) = \int_0^{2\pi} E(f,\theta)d\theta$). As in the probability density distribution, from the wave frequency spectrum, many wave parameters can also be derived. Before that, it is important to introduce the moments of wave spectrum (m_n) which is formulated as

$$m_n = \int_0^\infty f^n E(f)df, \quad n \in \mathbb{Z}. \tag{1}$$

For example, the significant wave height can be approximated from the wave spectrum information by,

$$H_s = H_{m_0} \approx 4\sqrt{m_0}, \quad \text{where} \quad m_0 = \int_0^\infty E(f)df = s^2.$$

Hence, H_s can be approximated by 4 times of the standard deviation of the wave height distribution. The standard deviation of a distribution is a square root of its variance s^2. And many others wave parameters, such as mean spectral wave period ($T_{m_{01}}$), defined as the ratio of the zeroth moments and the first moment ($\frac{m_0}{m_1}$), and mean zero up-crossing period ($T_{m_{02}}$), can also be obtained from such moments (see Appendix A).

On time scales longer than 30 minutes, the stationary sea-state assumption is no longer valid. Thus wave variables, like H_s and $T_{m_{01}}$ will change in time. These parameters will behave like random variables and again can be analysed by means of probability density distributions. The characteristics of such distributions are expected to depend strongly on the stochastic nature of the local wind field. The analysis of wave parameters is usually done for time series with typical duration of years (Holthuijsen [2007]).
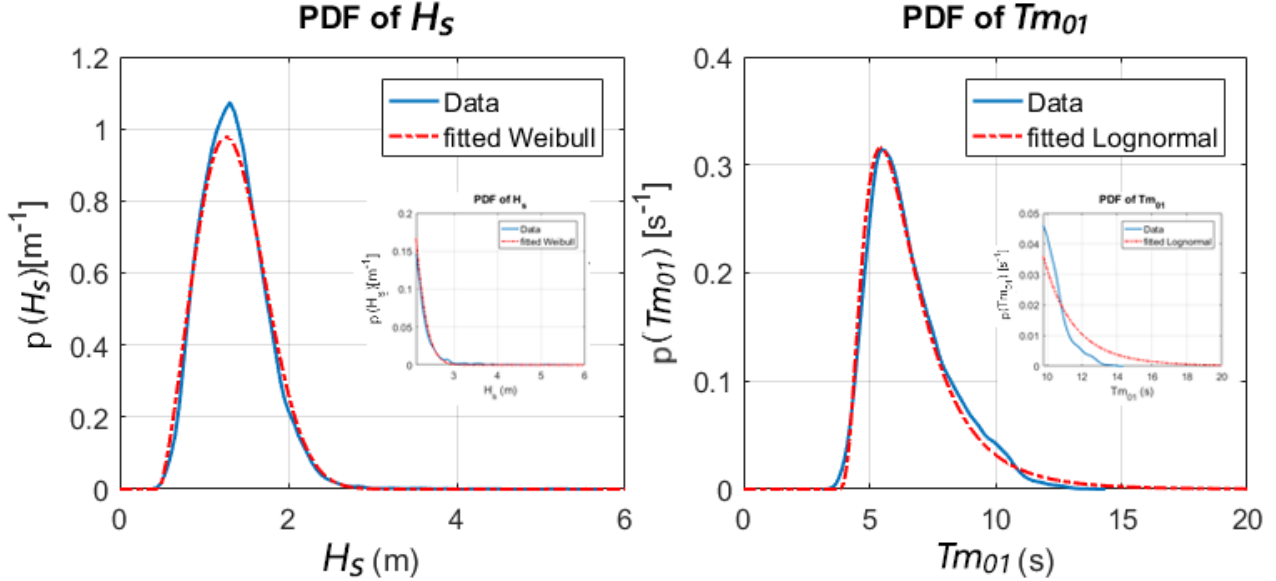
Figure 6: The probability density distribution of H_s (p(H_s)) [top] and $T_{m_{01}}$ (p($T_{m_{01}}$)) [bottom] from buoy measurement at Exmouth, Western Australia (blue histogram). p(H_s) is approximated by 3-parameter Weibull distribution, whilst p($T_{m_{01}}$) is fitted to 3-parameter Lognormal distribution (dash-dotted red line). In insets, the high tail of distributions are shown in zoomed version.

Figure 6, the probability distributions of H_s and $T_{m_{01}}$ during measurement at Exmouth are shown. A 3-parameter Weibull distribution is suggested to fit the distribution of H_s (Mathiesen et al. [1994]). And for the distribution of $T_{m_{01}}$, a 3-parameter Lognormal distribution shows a good fit. However, these choices need further consideration regarding the high tail of the distributions. When the fitted Weibull distribution underestimates the observed high H_s, but the fitted Lognormal overestimates the observed of high $T_{m_{01}}$ (see Figure 6 insets). Hence, a further investigation of extreme value is needed.

The time evolution of wave parameters can be simulated with spectral wave model. Such models are based on physics of waves: propagation of energy by wave group speed[3], local sources of energy due to wind and local sinks due to dissipations. Examples of such wave prediction models are WAM (The WAMDI Group [1988],Komen et al. [1994]), WaveWatch III (Tolman [1991]), and SWAN (Booij et al. [1999]). These models calculate the wave spectrum based on the energy balance equation. The basic principle of this equation will explain further in the following section.

---

[3]wave propagation speed (c) is calculated by $c = \frac{\sigma}{k}$ and wave group celerity (c_g) is defined as $c_g = \frac{\partial \sigma}{\partial k}$

2.2   Spectral Wave Model

The SWAN model is used for simulating wind waves in this thesis. The model uses a term of spectral wave action density (N), i.e. spectral wave variance density (E) in $m^2/\text{H}_z/\theta$ divided by frequency (σ). The reason behind this choice is that, in absence of forcing and dissipation, the wave action balance is conserved even under influence of ambient currents (SWAN team et al. [2016a]). Moreover, the SWAN also estimates physical effects like refraction and diffraction which occur in shallow water which is more convenient to be used in areas that have many islands, such as in the seas of Indonesia.

The action balance equation is formulated as

$$\frac{\partial N}{\partial t} + \vec{\nabla} \cdot (\vec{c_g} \cdot N) = \frac{S}{\sigma} \tag{2a}$$

$$S \equiv S_{in} + S_{ds} + S_{nl} \tag{2b}$$

where $N = N(x, y, \sigma, \theta)$ is the action density function (E/σ), $\vec{c_g}$ is a wave group velocity, and S is source (and sink) of wave energy terms. The term S contains S_{in} which describes input of energy by wind, S_{ds} is dissipation of energy and S_{nl} describes the shift of energy due to non linear interaction between waves (2b). Finally, $\vec{\nabla}$ is a derivative operator with respect to $x, y, \sigma,$ and θ. Equation 2a shows that wave action density propagates spatially in the speed of wave group velocity $(c_{g,x}, c_{g,y})$. During the propagation, the wave action is also experiencing a frequency-dispersion $(c_{g,\sigma})$ and a direction-dispersion $(c_{g,\theta})$ (Holthuijsen [2007]). Frequency-dispersion tells that, in a group of waves, low frequency waves travel faster than high frequency waves. While direction-dispersion is an effect of propagation of individual waves to many different directions and hence spreads the energy.

Figure 7: Illustration of physical processes during wave generation and propagation. In the deep ocean, the main sink is white-capping (spontaneous breaking wave), while both of bottom friction and surf breaking dominate the dissipation of energy in the shallow area.

The source energy of the wave generation is the near-surface wind. This wind is usually measured at 10 m above the surface and known as $\overrightarrow{U_{10}}$. In the computation, the surface wind is transformed into wind friction (U_*) by introducing a drag coefficient (Holthuijsen [2007]), i.e.

$$|\overrightarrow{U_{10}}|^2 = C_D U_*^2. \tag{3}$$

The surface wind transfers its energy to the wave by the resonance mechanism described by Phillips and the shear instability mechanism described by Miles (Booij et al. [1999]). In SWAN, these mechanisms are translated into parametric wave growth function as

$$S_{in}(\sigma, \theta) = \alpha + \beta E(\sigma, \theta). \tag{4}$$

The wave growth function consists of a linear term and an exponential term. The linear wave growth term (α) is approximated by an empirical expression from Cavaleri and Malanotte-Rizzoli (1981), whereas the exponential growth (β) is taken from Komen et al. [1984] (Holthuijsen [2007]). Both variables α and β depend on the wind friction (U_*).
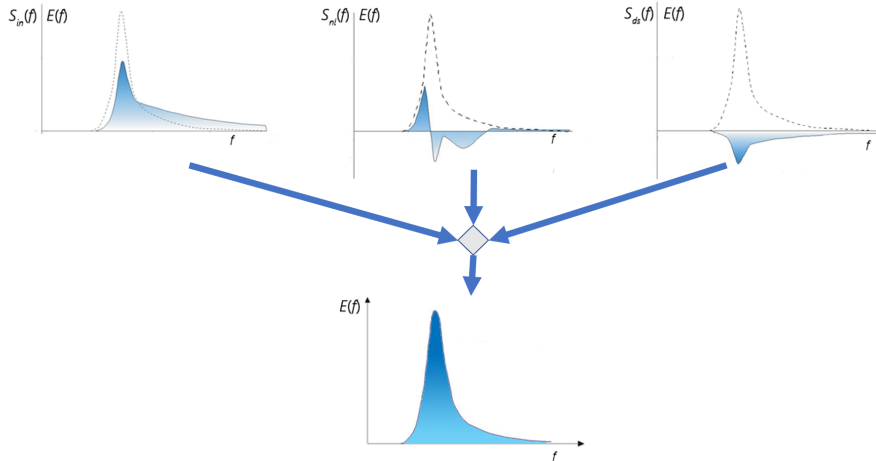

Figure 8: Illustration of physical processes in the action balance equation. The effects of source and sinks of wave variance are illustrated at top figures as $S_{in}, S_{nl},$ and S_{ds}, from the left to the right, respectively. They all combine and create a wave frequency spectrum [bottom figure] (after Holthuijsen [2007])

Waves also interact to each other during the propagation in the ocean. This interaction leads to a redistribution of energy over different of wave components (S_{nl}). Different efficient nonlinear wave interactions occur. Four-wave (*quadruplets*) interactions occur in the deep ocean, while three-wave or *triads* interaction is more important in shallow water area. In SWAN, quadruplets interaction is approximated by the Discrete Interaction Approximation (DIA) of Hasselmann et. al. [1985],

while triads interaction is computed by Lumped Triad Approximation (LTA) derived by Eldeber-sky [1996] (Holthuijsen [2007]; SWAN team et al. [2016a]). As shown in Figure 8, nonlinear wave interaction is introduced as +/-/+ shape (Bouws et al. [1998]; Holthuijsen [2007]).

During the propagation, the waves also can loose their energy because of three different of dissipation mechanisms (S_{ds}). The white-capping is a main source that dissipates the wave energy in the deep ocean. When the waves are entering shallow water, they lose the energy by bottom friction and/or wave breaking.

2.3   Design of Experiments

2.3.1   Numerical Set-up

Two different domains were considered to simulate wave conditions in the tropical Indian Ocean, i.e. a global domain and a tropical domain. The global domain covers almost entire world, from 180°W to 180°E and 81°S to 81°N with a resolution of 1°. In this domain, periodic boundary conditions were applied at the eastern and western boundaries. And open boundary conditions were imposed at both northern and southern boundaries. Later, wave spectra information at the location of boundaries of the tropical domain (see Figure 9) was extracted at every time step.
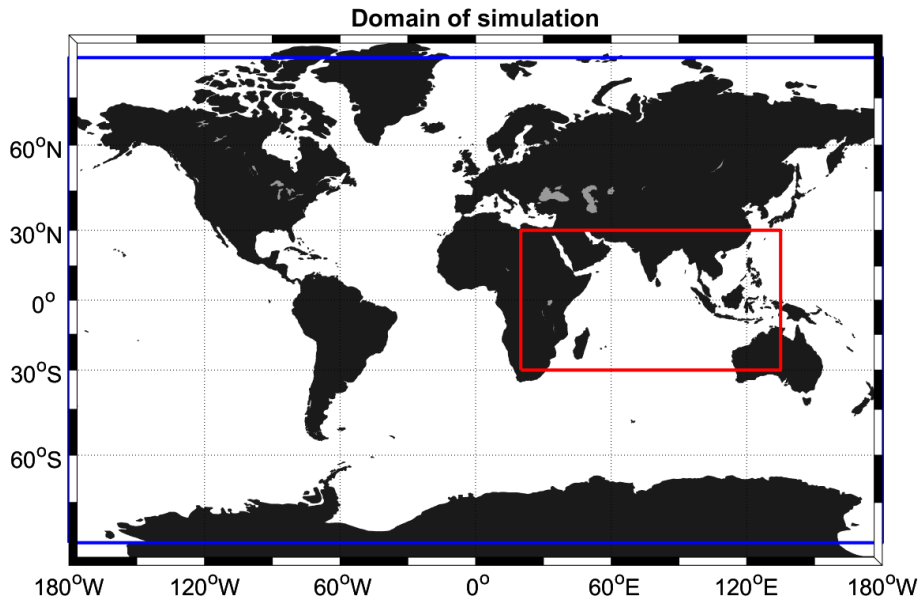

Figure 9:   Domains of the SWAN simulations. The global domain with a grid resolution of 1° is denoted by the blue square, whilst the red square indicates the tropical domain, with a grid resolution of 0.25°.

The tropical domain stretches from 20°E to 135°E and 30°S to 30°. It includes the Indian Ocean and part of the Pacific Ocean. The grid domain has a resolution of 0.25° in order to reproduce the

extreme waves due to tropical cyclones. At the boundaries, wave spectra are imposed that follow from model output of the global domain. There is no feedback of information from the tropical domain into the global domain. Therefore, it is called by *one-way nesting*.

Bathymetry data used for simulations were extracted from the website of ETOPO1 of NOAA (`https://www.ngdc.noaa.gov/mgg/global/`). These bathymetric data were smoothed and values at the grid points of each domain were obtained by linear interpolation.

Three hourly wind data were downloaded from the website of ECMWF (`http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/`). These wind data are the reanalysis products of ERA-interim model (Berrisford et al. [2011]). The values of wind speeds at 10-m height were downloaded in the same resolution as that of the domains.

The wind data due to tropical cyclones can be locally generated based on formulation by Holland [1980]. The variables needed in the formulation have been recorded during the events. This information is available at the websites of NHC and JTWC[4]. NHC (`http://www.nhc.noaa.gov/data`) provides TC information for the region of Atlantic Ocean and the Eastern of Pacific Ocean, while the information for the other regions is provided by JTWC (`https://metoc.ndbc.noaa.gov/web/guest/jtwc`).

Unfortunately, JTWC only records a complete variables since 2001. For that reason, the simulations were done for only the period of 2001-2016. However, for the computational efficiency, the simulations were done for each individual *wave year* that runs from November until October. This time frame was chosen by considering the appearances of tropical cyclone in both hemispheres. November marks the beginning of the tropical cyclone season in the Southern Hemisphere and the end of the tropical cyclone season in the Northern Hemisphere. Therefore, this chosen time frame was slightly different than that in Caires and Sterl [2005] which is October to September. The reason is that the calculation should not cut the summer period in both hemispheres. Before the actual simulations started, a warm-up period of 10 days was given.

Additionally, the internal time step used in the simulations was 1 hour. The frequency domain of the wave spectra has 40 frequency bins with a high and a low cut-off wave frequency values correspond to 1 and 25 sec, respectively. The directional domain (θ) has 10° of bin resolution, which corresponds to a total of 36 bins of wave direction.

2.3.2   Sub-grid Model for the Wind Field of Tropical Cyclone

The ERA-Interim model is obviously unable to resolve the occurrences of tropical cyclones (TC) because of its grid resolution. It leads to an underestimation of the maximum wind speed (Caires

---

[4]National Hurricane Center and Joint Typhoon Warning Center

and Sterl [2005]); and misplacement of the location of the eye of TC. The maximum wind speed commonly occurs within a distance of 30 km from the centre (Kimball and Mulekar [2004]). This value is smaller than the grid resolution of ERA-interim model, i.e. 79 km (Berrisford et al. [2011]). To resolve TC wind field in the wind input, an analytical solution from Holland [1980] was used. The equation of TC wind field (V_H) and the pressure profile (P_H) are written as

$$V_H(r) = \sqrt{V_m^2 A \exp(1 - A) + \left(\frac{rf}{2}\right)^2} - \frac{rf}{2} \tag{5a}$$

$$P_H(r) = P_c + (P_n - P_c) \exp(-A) \tag{5b}$$

where

$$A = \left(\frac{R_w}{r}\right)^B \text{ and } B = \frac{\rho e V_m^2}{P_n - P_c}$$

and ρ is density of air in 1.16 kg m^{-3}, $e = 2.7183$, V_m is maximum sustained wind speed, R_w is radius of maximum wind speed which is also the approximation of the size of the eye, f is Coriolis parameter, r is radius from the centre of TC, P_n is ambient pressure, and P_c is minimal pressure of the centre. The direction of the wind is determined by the cyclonic motion of TC, which is clockwise at the Southern Hemisphere or anti-clockwise at the Northern Hemisphere. Graph of $V_H(r)$ and $P_H(r)$ are shown in Figure 10.
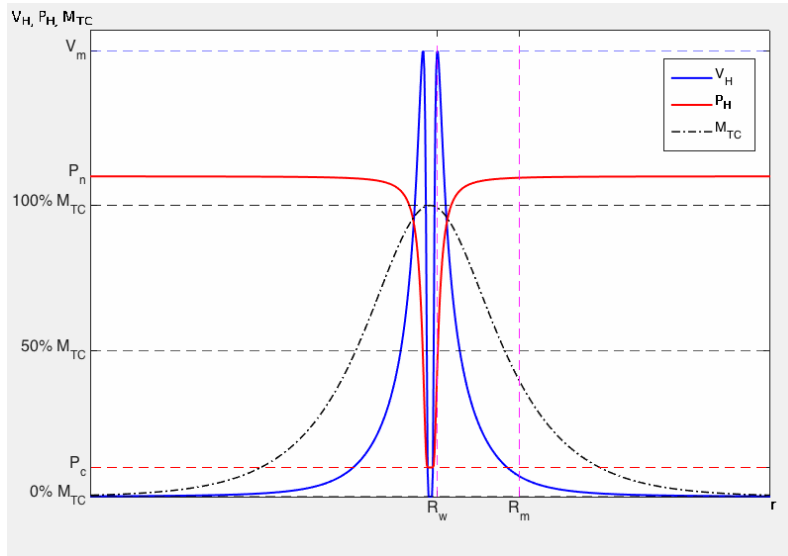

Figure 10: Illustration of tropical cyclone wind (V_H) profile given by Equation 5a (blue line). TC wind is plotted along the pressure profile $P_H(r)$ (red line) and the masking profile for the integration process with the ERA-interim wind field (black dashed-line). **Note:** The axes were normalized as the horizontal axis is the distance from the centre in radius r and the vertical axis determines the normalized values of V_H, P_H and M_{TC}.

In the preprocessing, the wind field from a tropical cyclone (V_H) is integrated with ERA-Interim wind field ($\overrightarrow{U_{10,ERA}}$) by using weighting function (M_{TC}) (see Figure 10). The weight criterion accounts for the distance between a certain grid point and the position of minimum pressure of TC (r) and the outer radius of wind speed (R_m). Therefore, at each grid point, a linear combination of the TC wind field and the ERA interim wind field. The integration of sub-grid model of TC into ERA-interim wind field is given as

$$|\overrightarrow{U_{10,\text{SWAN}}}| = (1 - M_{TC})|\overrightarrow{U_{10,\text{ERA}}}| + M_{TC}V_H \tag{6a}$$

$$\text{where} \quad M_{TC} = \text{sech}\left(\frac{\pi}{2R_m}r\right) \tag{6b}$$

is a weight coefficient and R_m is the maximum or outer radius of at which the wind field is affected by that of the TC wind speed. The secant hyperbolic profile was used as a weight function in order to preserve the magnitude of the maximum wind speed (V_m) (Equation 6b). At the same time, it is introduced to the ambient wind field slowly.
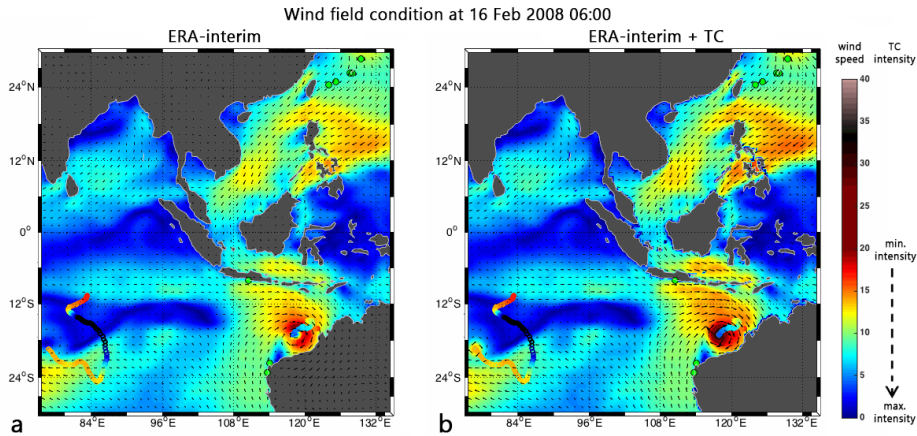

Figure 11:   The wind field of 16 February 2008 at 06:00 UTC are shown for (a) the original ERA-interim wind field and (b) the ERA-interim + tropical cyclone wind field used in the SWAN simulations. Tropical cyclone tracks are plotted in thick coloured lines based on the intensity. A shared colour scale is shown on the right, for the wind speed from 0-40 m/s (from blue-red-cream) and intensity of TC (from red-blue). Arrows in the pictures show the magnitude and the direction of the wind.

According to the Equation (5a), the shape of TC was assumed to be a perfect circle. However, the shape of TC is actually not a perfect circle (see Figure 11). Because the ERA-interim wind field has included TC, although it is underestimated and sometimes misplaced. On the Figure 11, the eye of TC on the right in the original ERA-interim wind field (a) had a larger size and was in a different location than from the integrated wind field (b), but (a) is less intense than (b). The maximum wind of TC on the left is not captured in the original ERA-interim wind field (a).
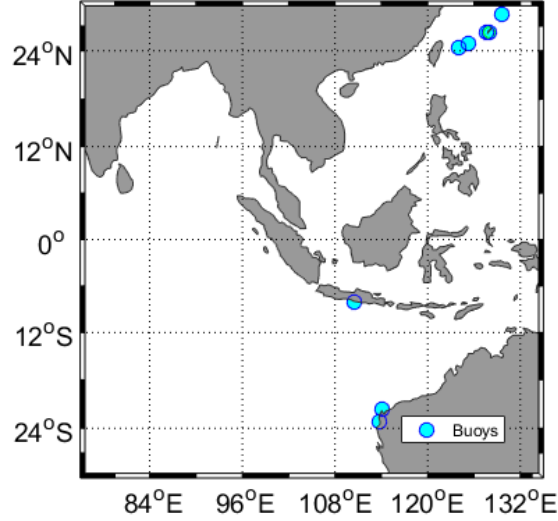
2.3.3 Model Calibration

Figure 12: Map of location of buoys. The buoys are indicated by blue markers.

Table 1: List of downloaded data from measurement buoys.

Buoy(s)	Coral Bay$_a$ Exmouth$_b$	Baron Bay	402_c, 701_d, 702_e, 705_f, 706_g
Provider	Department of Transport, Western Australia♣	Agency for the Assessment and Application of Technology (BPPT), Indonesia♠	NOWPHAS project from Ministry of Land, Infrastructure, Transport and Tourism, Japan★
Coordinates	114.0986°E, 21.6994°S$_a$ 113.7442°E, 23.1743°S$_b$	110.547°E, 8.1362°S	129.5217°E, 28.4519°N$_c$ 127.6475°E, 26.2581°N$_d$ 127.9653°E, 26.2422°N$_e$ 124.1028°E, 24.3653°N$_f$ 125.2356°E, 24.8608°N$_g$
Time of measurement	22/12/2004 - 30/03/2005$_a$ 04/10/2006 - 31/08/2009$_b$ 03/10/2011 - 31/12/2011$_b$	27/05/2014 - 30/09/2014	01/11/2000 - 31/12/2016$^*_{cde}$ 01/01/2005 - 31/12/2016$_{fg}$
Measured Parameter	$H_s, T_{m_{01}}, T_p, \overline{\theta}$	H_s, T_z	$H_s, T_z, \overline{\theta}$

♣ https://www.transport.wa.gov.au/   ♠ https://bppt.go.id/english/   ★ https://nowphas.mlit.go.jp/eng/
[*]The observation data are available since: $c = 1977, d = 1973,$ and $e = 1977.$

The modelled wave parameters were compared with the measured wave parameters at some buoys listed in Table 1. Although only a few, all buoys are located in the tropical domain. These buoys are needed to verify the model performance and to calibrate the model. The source of input wind $(S_{in}(f, \theta))$ is determined by the term of wind friction (U_*), which contains exponential growth term β. The approximation of β in the ECWAM contains a coupling coefficient, which is allowed to vary as a consequence of the air-sea interaction in the model. However, since only a wave model was used in this thesis, i.e. SWAN, the calibration process for the S_{in} was introduced to mimic this air-sea interaction. The calibration processes were done in two steps, i.e. by changing a factor of drag coefficient (ξ) and modifying the variation of wind input (χ). Furthermore, calibrations to other physical processes $(S_{ds}$ and $S_{nl})$ were not considered due to limitation in numbers of observation of the wave parameters.

Drag Coefficient

Both processes were determined to calibrated the input wind (S_{in}). The factor of drag coefficient (ξ) was allowed to vary hence the average of wind friction (U_*) became adjustable. The first method was determined to adjust the average of wave energy resulting by the wind input (U_{10}). This was done by introducing a constant ξ into the estimation of U_* as
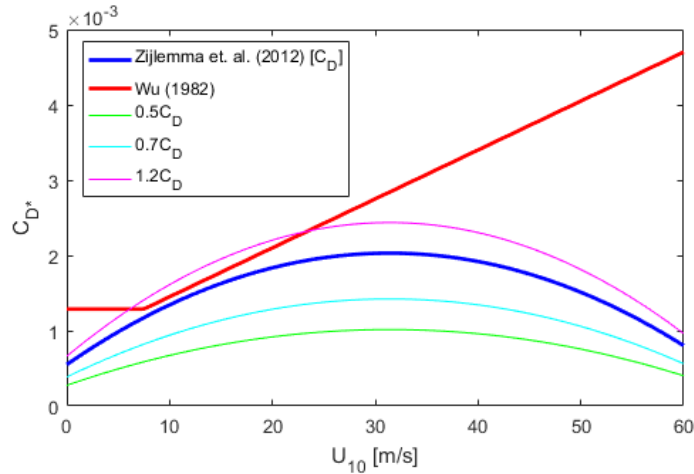
$$C_{D*} = \xi C_D. \tag{7}$$

Figure 13: The recent formulation of drag coefficient in SWAN uses a 2nd order polynomial fit function from Zijlema et al. [2012] $[C_D]$ (blue line), which formulation from Wu [1982] was used for the previous version (red line). Then, cyan and magenta lines are for C_D that are multiplied by factor $\xi < 1$ and $\xi > 1$, respectively.

.

The formulation of the drag coefficient that was used in the SWAN simulations is the one by Zijlema et al. [2012]. They proposed a polynomial function of wind speed to approximate the value of corresponding drag coefficient (Figure 13). The coefficient ξ allows the 2nd order polynomial fit (blue line) to deviate up-and-down linearly (see Figure 13).

Wind variation Enhancement

The coupling coefficient in the ERA-interim works to adjust both surface of the ocean and the wind speed in every 30 minutes of model time step. However, the accessible outputs from this model are 3-hourly datasets as the finest time resolution. Hence, any variations that occur less than 3 hours are not captured in the datasets. The SWAN model only uses the input wind from ERA-interim to generate the waves as a one-way forcing, thus no feedback mechanism is used. Furthermore, the coupling process in the ERA-interim that happened shorter than the temporal resolution of the data needs to be resembled. Therefore, an enhancement to the wind variation was introduced to mimic the coupling mechanism, which is formulated by

$$U_{10,new}(U_{10}; \chi, \overline{U}) = \min \left(\left(\frac{U_{10}}{\overline{U}} \right)^{\chi} U_{10}, 1.5 U_{10} \right). \tag{8}$$

The value of χ is ≥ 0, as it is expected to give a higher variation during strong wind and the other way around. And variable $\overline{U}$ is the mean wind speed during half-year period, i.e. November-April and May-October, of the corresponding wave year. Furthermore, $U_{10,new}$ can not be larger than 1.5 times of the original U_{10}. Finally, the adjusted wind field ($U_{10,new}$) was used as an input for the wave model.

2.4   Statistical Analysis of Wave Parameter

Section 2.1 shows that at time scales less than 30 minutes, statistically stationary condition of sea level variation can be assumed. Theoretically, the probability density distributions of surface elevation, $p(\eta)$, will follow a Gaussian function if the energy spectrum sufficiently narrow (Longuet-Higgins [1975]). Meanwhile, the probability density distributions of wave height, $p(H)$, are approximated to follow Rayleigh distribution (Longuet-Higgins [1957, 1975]). However, there is no certain theory can be applied to the distribution of wave parameters, such as H_s and $T_{m_{01}}$, in the long time series. The statistic of these parameters are determined by the long-term variations of the wind. Thus, an empirical function to fit the distribution is needed.

The SWAN simulation results time series of wave parameters at every grid point in the domain. Consequently, the goodness of the simulation was evaluated. The time series of wave parameters

located at the closest grid point to the measurements were compared. The similarities and the differences between them were analysed. Finally, an analysis of the probability distributions of wave parameters were done.

2.4.1   Goodness of Fit Analysis

The verification was done by comparing two sets of time series of wave parameters, which were obtained from the observation and the SWAN simulation. Two approaches were used for this, which one by studying *quantile-quantile*[5] plot or Q-Q plot and another one by calculating the mean deviations between modelled and observational data.

Q-Q plot compares two sets of time series by means of their cumulative density distributions. The cumulative distribution of a certain threshold value $\hat{x}$, $P_{\hat{x}}(x)$, tells the quantile of any value in the time series of variable x larger than $\hat{x}$. In term of probability density distributions, $P_{\hat{x}}(x)$ means the probability of occurrence of x value that is lower than $\hat{x}$ value ($Pr(x < \hat{x})$), or it can be written as

$$P_{\hat{x}}(x) = Pr(x \leq \hat{x}) = \int_0^{\hat{x}} \mathrm{p}(x)dx. \tag{9}$$

In the Q-Q plot, the cumulative density distributions of the two sets of time series were compared and then plotted each other by their corresponding quantile values or $\hat{x}$.

The second estimation was done to quantify the similarities or differences of the two time series. The similarities of two sets were quantified by means of the correlation coefficient (R). While the differences of two sets of time series were analysed by calculating the root-mean-square error (RMSE) and Bias. The formulations of these calculations are expressed by

$$\mathrm{R} = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2 \sum_{i=1}^{N}(y_i - \overline{y})^2}}, \tag{10a}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2}, \tag{10b}$$

$$\mathrm{Bias} = \overline{x} - \overline{y}. \tag{10c}$$

where x_i are observational data points, y_i are data points of the simulations, and N is the total number data points which should be the same for both datasets. Furthermore, the $\overline{x}$ and $\overline{y}$ are the mean value of each dataset.

---

[5]From cumulative density function, $P_{\hat{x}}(x) = \int_0^{\hat{x}} \mathrm{p}(x)dx = n$, where $0 \leq n \leq 1$ is chosen, as such n^{th}-quantile value of $\mathrm{p}(x)$ or $Q_n(x)$ is $\hat{x}$. $Q_n(x)$ is the inverse function of $P_{\hat{x}}(x)$.

The correlation coefficient (R) yields information about the goodness of the simulation when compared to the observation. The R value ranges from -1 to 1. If the R value goes close to -1, the two datasets are anti-correlated each other. And the other way around, the SWAN simulation is correlated with the measurement data if the R value is close to 1. While, R is 0 means the datasets are uncorrelated at all.

The RMSE analysis tells the difference of the simulation to the observation. In this thesis, the smaller the RMSE value means that SWAN have a better approximation in simulating the real condition. While Bias explains the amount of the difference between average of two time series.

2.4.2 Wave Climate and Extreme Wave Analysis

The probability density distributions of wave parameters were analysed. There are two methods used to estimate characteristics of these distribution, i.e. by considering the full distribution (wave climate) or by considering the part of the distribution for high H_s (extreme waves) analysis. The wave climate statistic used to show the common pattern of wave parameters in a certain region. Whereas, the extreme conditions were further analysed regarding the high tail of the distribution (Figure 3).

Wave Climate Statistic

The wave climate statistic is quite useful to describe wave condition at respective locality (Goda [1979]). In this thesis, wave climate statistic was shown in term of average (μ), variance (s^2), skewness (s_k), and kurtosis (k_u). The formulations are expressed by

$$\mu = \int_{-\infty}^{\infty} x \, \mathrm{p}(x) \, dx, \tag{11a}$$

$$s^2 = \int_{-\infty}^{\infty} (x - \mu)^2 \, \mathrm{p}(x) \, dx, \tag{11b}$$

$$s_k = \int_{-\infty}^{\infty} \left(\frac{(x-\mu)}{\sqrt{s^2}}\right)^3 \mathrm{p}(x) \, dx, \tag{11c}$$

$$k_u = \int_{-\infty}^{\infty} \left(\frac{(x-\mu)}{\sqrt{s^2}}\right)^4 \mathrm{p}(x) \, dx, \tag{11d}$$

respectively. Where x is either H_s or $T_{m_{01}}$.

The mean value or average is closely related to the typical value that can observed. Variance explains the width of the distribution. The skewness describes the tendency of the distribution to have more values for below or above the average. While, kurtosis represents the length of a tail from a distribution. Additionally, the spatial distributions of probability density distributions of H_s, $\mathrm{p}(H_s)$, and of $T_{m_{01}}$, $\mathrm{p}(T_{m_{01}})$, along sectional line in the domain of interest were investigated.

Extreme Wave Analysis

The extreme values are defined as any value that is higher than a certain threshold (Figure 3). However, a common description for this application is by showing the extreme value that is expected to occur only once from a certain period in unit of year m, i.e. the m-year return value. Subsequently, the return period of these extreme values determines the threshold value from the distributions.

Before calculating the return value of H_s, any value that classified as extreme values were extracted from the datasets. There are three common methods to select the extreme values from the distribution, i.e. a total distribution method, annual maxima method and Peak Over Threshold (POT) method (Goda [1979]; Holthuijsen [2007]). The first method considers any values that higher than the threshold value, from the whole dataset. The annual maxima considers only the maximum value of every year as the extreme values. Hence, the time series of extreme values has a time step of a year. Whilst, the peak-over-threshold method considers any peaks value that are highest than a certain threshold $(\hat{x})$ to be formed as a distribution of extreme values. Then, the dataset that contains only the extreme values was formed into a probability density distribution.
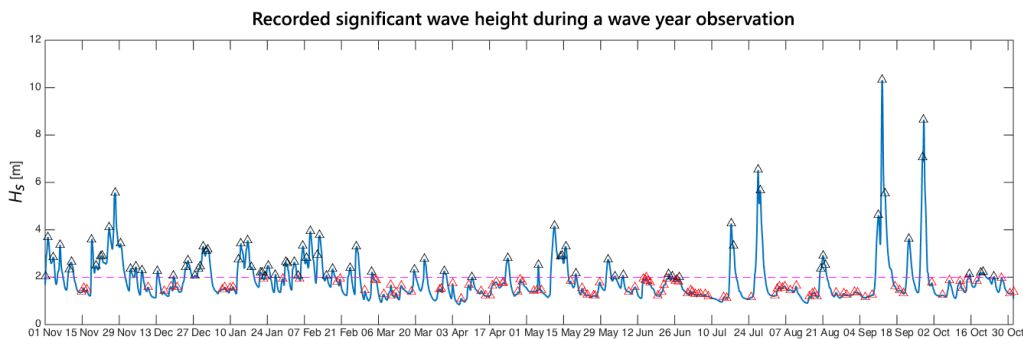

Figure 14: Illustration of an application of the Peak Over Threshold (POT) method. Every peaks that is $\tilde{H}_s$ (triangle markers) higher than a threshold value $(\widehat{H}_s)$, for this example, the threshold value is a 2 meter, is marked by black-triangles. Those peaks are extracted from the total distribution and formed into a new distribution for the extreme wave analysis.

Figure 14 shows an example on the way to select extreme values from a dataset by POT method. In this example, the threshold value $\widehat{H}_s$ is set to 2 meters (magenta dashed line) as such, so that any peaks higher than 2 meters are extracted and formed into a new distribution which contains these peaks $(\tilde{H}_s > \widehat{H}_s)$. However, in the application of POT method, the threshold value is commonly chosen from the nth-quantile[6] of the total distribution (see Figure 3), rather than a certain H_s value, as shown in the illustration in Figure 14.

---

[6]see Footnote 5.

For the extreme wave analysis, especially due to occurrence of TC, the POT method is commonly used (Caires and Sterl [2005]; Caires [2011]; Timmermans et al. [2017]). By this method, the threshold value can be set as such, so that only corresponding extreme events to TC are selected. Hence, the time period between peaks represents as a period between storms $\Delta\tau$[7]. In this study, the POT method is chosen.

The return value is estimated from a probability density distribution that is fitted to the distribution of extreme values, which is the result of POT method. The common probability density distribution that follows from application of POT method is Generalized Pareto Distribution (GPD) (Caires and Sterl [2005]; Holthuijsen [2007]; Caires [2011]). The GPD famous for estimating the tail of distribution, which in this case is the distribution of extreme values. The cumulative density function ($P_{\hat{x}}(x)$) of Generalized Pareto Distribution with two parameters is expressed as

$$
P_{\hat{x}}(\tilde{x}; k, z) = \begin{cases} 1 - \left(1 + \dfrac{k\hat{x}}{z}\right)^{-\frac{1}{k}}; & k \neq 0 \\[2mm] 1 - \exp\left(-\dfrac{\hat{x}}{z}\right); & k = 0 \end{cases} \tag{12}
$$

where $\hat{x}$ is the threshold value for the distribution of peaks $\tilde{x}$, z (scale parameter) > 0 and k (shape parameter) $\in \mathbb{R}$. $0 \leq P_{\hat{x}}(\tilde{x}) \leq 1$. The shape parameter k determines the length of the tail. Subsequently, the m-year return value of wave parameters was estimated by using POT-GPD method in the following chapter.

The return value of certain m-year period, $\hat{x}_m$, is estimated by using corresponding quantile value of GPD and the time step between data in the distribution of peaks ($\tilde{x}$) which has been selected by POT method,

$$
T_{return} = m = \frac{\Delta\tau}{1 - P_{\hat{x}_m}(\tilde{x})} \quad \text{or} \quad 1 - P_{\hat{x}_m}(\tilde{x}) = \frac{\Delta\tau}{m} \tag{13}
$$

where $P_{\hat{x}_m}(\tilde{x})$ is the corresponding CDF value of m-year return period, $\Delta\tau$ is a average time between the successive peaks in the distributions of extreme waves ($Pr(\tilde{x} > \hat{x})$), which were extracted by POT method and m is the determined return period in unit of years. Consequently, the threshold value, by means of m-year return value ($\hat{x}_m$), was approximated by taking an inverse of Equation (13) and inserting to Equation (12). Caires and Sterl [2005] formulated the m-year return value of wave parameters ($\hat{x}_m$) as,

$$
\hat{x}_m = \begin{cases} \hat{x} + \dfrac{z}{k}\left[\left(\dfrac{m}{\Delta\tau}\right)^k - 1\right]; & k \neq 0 \\[3mm] \hat{x} + z\log\left(\dfrac{m}{\Delta\tau}\right); & k = 0 \end{cases} \tag{14}
$$

---

[7]By assuming that any peaks higher than thethreshold value in the time series of H_s are due to storm events.

However, the method for choosing the threshold value ($\hat{x}$) that is used in POT method is still uncertain. Caires and Sterl [2005] suggested to use a threshold value between 0.93^{th}-0.97^{th} quantile of the total distribution such that the GPD distribution gives a best fit to the actual distribution. The threshold value for POT method can be as high as the 0.997^{th}-quantile in some location to retrieve the best fit ([Timmermans et al., 2017]). Therefore, there is no definite formulation of $\hat{x}$ that could be applied to every ocean. Consequently, an estimation of this value needs further consideration.

3 Results

SWAN simulations were done for 16 years in both the global domain and the tropical domain with spatial resolutions of $1°$ and $0.25°$, respectively (see Figure 9). The modified ERA-interim wind field by including the Holland sub-grid model was used as the input forcing. From these simulations, 3-hourly map of wave parameters for a whole domain and, particularly, the 2-D wave spectrum data at the location of buoys were recorded. Finally, recorded wave parameters, especially for H_s and $T_{m_{01}}$, were analysed based on their climate and extreme conditions.

3.1 Model Verification

The model verification was done to reveal the optimum calibration coefficients (see section 2.3.3), i.e. C_D calibration factor (ξ) and wind variation enhancement (χ). The contents of this subsection are related to the first research question, as given in the Introduction. Consequently, the SWAN simulations were performed only during the wave year of 2008 to compare with the wave measurement located at Exmouth, Western Australia (see Table 1). Additionally, the distribution of wave parameters from both measurement and SWAN simulations were compared.

Table 2: Goodness of fit analysis for SWAN simulations results with respect to the measurement data at Exmouth and ERA-interim reanalysis product. Wave parameters to be compared are H_s and $T_{m_{01}}$, which the average values (μ) during the wave year of 2008 are indicated below.

Wave Parameter	Calibration Coefficient		SWAN vs. Buoy of Exmouth			SWAN vs. ERA-Interim			ERA-Interim vs. Buoy of Exmouth		
	ξ	χ	R	RMSE [m]	Bias[♠] [m]	R	RMSE [m]	Bias[♠] [m]	R	RMSE [m]	Bias[♠] [m]
Wave Height $[H_s]$ μ_{H_s} : 1.75 m*	0.7	0	0.75	0.52	-0.43	0.88	0.40	-0.35	0.75	0.31	-0.09
	0.5	0.125	0.76	0.30	0.11	0.87	0.27	0.20			
	0.7	0.125	0.77	0.40	-0.30	0.89	0.27	-0.20			
	1.2	0.125	0.66	1.17	-1.10	0.81	1.05	-1.00			
	0.7	0.225	0.76	0.52	-0.44	0.89	0.39	-0.34			
Wave Period $[T_{m_{01}}]$ $\mu_{T_{m_{01}}}$: 7.14 s*	0.7	0	0.73	1.22	-0.45	0.84	1.17	-0.82	0.78	1.02	0.36
	0.5	0.125	0.72	1.40	0.88	0.84	0.84	0.50			
	0.7	0.125	0.74	1.07	-0.04	0.84	0.84	-0.40			
	1.2	0.125	0.64	2.00	-1.48	0.75	2.14	-1.84			
	0.7	0.225	0.74	1.19	-0.45	0.84	1.15	-0.81			

*The average values (μ) of H_s and $T_{m_{01}}$ were calculated from the time series of observed H_s and $T_{m_{01}}$.
♠ Bias is calculated based on Eq. 10c.

The optimum values for the calibration coefficients were determined by goodness of fit analysis. With this method, the goodness of SWAN simulation in reproducing the time series of H_s and $T_{m_{01}}$ was quantified. Besides the Exmouth buoy, retrieval values from ERA-interim were also used for the verification. The goodness of model output with respect to observational data was quantified

by means of value for correlation coefficient (R), root-mean-square error (RMSE), and Bias (see section 2.4.1).

From Table 2, the coefficient of $\xi = 0.7$ and $\chi = 0.125$ were found to be the optimum values because these give the highest correlation to the buoy and the lowest value for the RMSE and the Bias, especially for the $T_{m_{01}}$. The correlation value of H_s modelled by SWAN with observation is higher than that of ERA-interim (Table 2). This is probably because the SWAN simulations with the sub-grid model are able to capture the extreme wave conditions generated by TC.

The RMSE of modelled H_s shows a remarkable deviation from measured H_s, viz. 0.4 m. The bias shows a negative value of -0.3 m, which means that SWAN gives a higher average value of H_s. Because of that, the SWAN simulation results are expected to have better agreement with observation by assuming that the RMSE occurred due to the Bias. Nevertheless, these values are lower when they are analysed for the entire domain. Therefore, these values were used to simulate the 16-years wave conditions.

Another way to verify the model performance is by examining Q-Q plots. Figure 15 shows the Q-Q plots of H_s and $T_{m_{01}}$ at Exmouth, Western Australia. The nth-quantile values from modelled wave parameters over a 16-year period were compared to those observed parameters. Comparison for other buoys are shown at Appendix C.

Q-Q plots of H_s show that ERA-interim is unable to capture high waves events (Figure 15c). The blue markers are shown to deflect to the left from the 1-1 line (black line). Obviously, ERA-interim data will not produce a good result on the estimation of return values of extreme waves. SWAN results, on the other hand, show a good agreement with the measurements (see Figure 15a). A regression line qualitatively has a similar slope as that of the 1-1 line.

Q-Q plots of $T_{m_{01}}$ show similar features, but both SWAN and ERA-interim are unable to fit the

Figure 15: Q-Q plots of H_s [**a-b**] and $T_{m_{01}}$ [**c-d**]. Recall that $Q_n(x)$ is the inverse of fraction events for which $x < \hat{x}$ (see Footnote 5), the cumulative density distributions of wave parameters from simulations of the SWAN [**(a-c)**] and the ERA-interim [**(b-d)**] are compared to ~3-years of measurement data from wave buoy at Exmouth, Western Australia. The black line is the 1-1 line and the red-dashed line is the linear regression line.

high periodic waves (10 s and more). In Figure 15b, the regression line (red-dashed-line) has a very small deviation from the 1-1 line (black-line). This is also indicated by the value of the Bias of $T_{m_{01}}$ between SWAN and observations, which has a value of -0.04 m (see Table 2).

3.1.1   Differences of Wave Climate Statistics: SWAN vs. ERA Interim

On section 3.3.1, an analysis of wave climate statistic is done by using SWAN simulation result. It is important also to understand the wave characteristics in the tropical Indian Ocean from the ERA-interim results because ERA-interim has shown a fairly well performance to estimate the climate conditions, despite having a weakness on the estimation of the extreme condition (Dee et al. [2011]). Subsequently, differences between the output of SWAN simulation and ERA-interim are analysed.

Figure 16: Map **a** shows differences between mean significant wave height obtained with SWAN and with ERA interim. **b** is same as **a**, but for spectral mean wave period ($T_{m_{01}}$). The location of the buoys are marked in green-dots. Again, the numbers between brackets on titles indicate the mean value from a whole domain. **Note:** colour scales are used together for both panels. The red colour means that the output from SWAN simulation has higher value than the results of ERA-interim model.

Figure 16a shows the difference of the average value of H_s from the output of SWAN simulations and the ERA-interim model. Overall, SWAN is able to produce a similar characteristics as those obtained with ERA interim. From Figure 16a-b, the mean of the differences for a whole domain are remarkably small, as -0.12 m for H_s and -0.21 s for $T_{m_{01}}$, compared to their domain averaged value that occurs in the domain (see Figure 20a and Figure 21a). However, both differences on average of H_s and $T_{m_{01}}$ show a specific pattern. In the northern Indian Ocean and the eastern Indonesian sea, SWAN yields higher value than ERA interim. In contrast, ERA interim gives a higher value in the South China Sea, in the western Indonesian sea, and in *shadow zones*[8] (e.g. Gulf of Persia,

---

[8]shadow zone in ocean waves is an area behind the obstacle in which the wave propagates into. (Holthuijsen [2007])

north of Australia and Madagascar). SWAN also underestimates H_s in the Southern Ocean. A similar look is also shown for $T_{m_{01}}$ in Figure 16b. In both figures, the SWAN results a small deviation from the ERA-interim, especially at the location of buoys (green-dot markers).

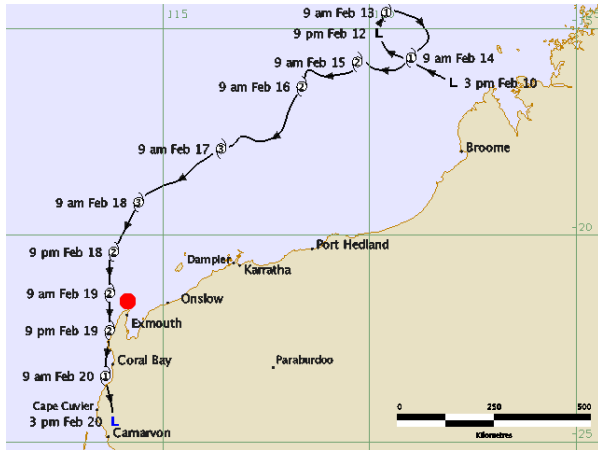## 3.2 Analysis of Wave Evolution during Tropical Cyclone Nicholas

The purposes of this subsection are mainly to answer of the second research question and partly to emphasize the explanation in the previous subsection (section 3.1). For these purposes, the analysis of wave conditions was done by investigating time series of wave parameters and wave spectra during TC Nicholas. The modelled, both SWAN and ERA-interim, and observed H_s and $T_{m_{01}}$ were compared. From the comparison, the goodness of SWAN with a sub-grid model of TC wind field in reproducing the wave conditions during TC was examined.

Figure 17: Storm track of tropical cyclone Nicholas (10-21 February 2008). The buoy was located near the city of Exmouth (red-dot). Source : `www.bom.gov.au` (timestamps in Figure are in UTC+9)

Figure 17 shows tracks of TC Nicholas which occurred in the eastern Indian Ocean near Australia between 10-21 February 2008. The timestamps in the figure are in UTC+9. TC Nicholas reached the maximum intensity on 16 February 2008 06:00 with a maximum wind speed of 80 knots (or $\sim$41.2 m/s). The radius of maximum wind speed at that time was only 10 nautical miles (18.5 km). It maintained its intensity until 18 February 06:00 at $\sim$200 km to the north off-coast Exmouth, before it weakened while moving southward. At midnight of 19 February, the eye of TC Nicholas was located at the closest buoy measurement at Exmouth, Western Australia (see Figure 17). At that time, the maximum wind speed had been reduced into 55 knots (29 m/s).
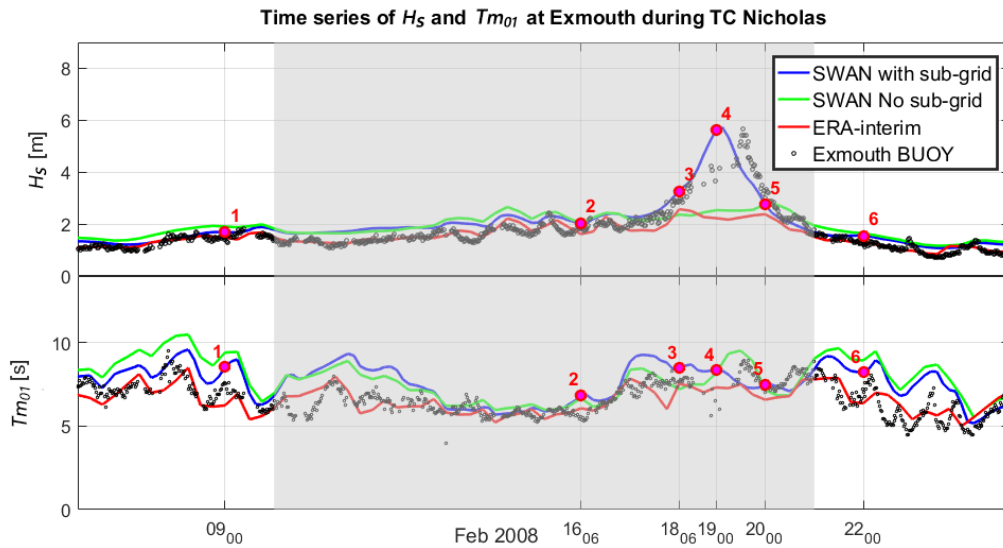
Figure 18:   Time series of H_s [**top**] and $T_{m_{01}}$ [**bottom**] at Exmouth during TC Nicholas on 10-21 February 2008 (grey-shaded) from 3 sources: SWAN with sub-grid model (blue line), SWAN without sub-grid model (green line), ERA-Interim (red line), and wave buoys from Department of Transport, Western Australia (black dots). The red dot markers with numbers indicate the time when the wave spectra were analysed.

Figure 18 shows time series of H_s and $T_{m_{01}}$ obtained from SWAN, ERA-interim and a measurement buoy. A maximum H_s of almost 6 meters was recorded at 19 Feb 2008 00:00 (see marker no. 4 in the top panel of Figure 18). At this time, the TC was above the location of the buoy (see Fig. 17). A different behaviour is shown by $T_{m_{01}}$ as the observed mean period dropped slightly after the H_s reached a maximum. The SWAN is showed this drop in period only after some time, while the observation show a significant drop of $T_{m_{01}}$ from 7 to 5 sec (see marker no. 4 Figure 18). The peak period (T_p)[9] was not really well calculated by SWAN, especially between 13-16 February 2008.

Before TC Nicholas occurred (see marker no. 1 and 6 Figure 18), a relatively high $T_{m_{01}}$ were recorded. However, the measurement of H_s recorded a relatively low H_s. This condition probably occurred due to an influence of swell from the southern ocean, which annually average propagates to the north-east direction in Exmouth (Zheng and Li [2017]). Further examination will be taken into the wave spectra.

The green line in Figure 18 is the output of the SWAN model without the sub-grid model. By using the same input wind, H_s of the SWAN simulations (green line) closely follows that of ERA-interim (red line). Figure 18 shows that SWAN can reproduce similar wave characteristics as ERA-interim by adjusting the calibration coefficients to mimic the air-sea interaction process.

---

[9]see Appendix A.

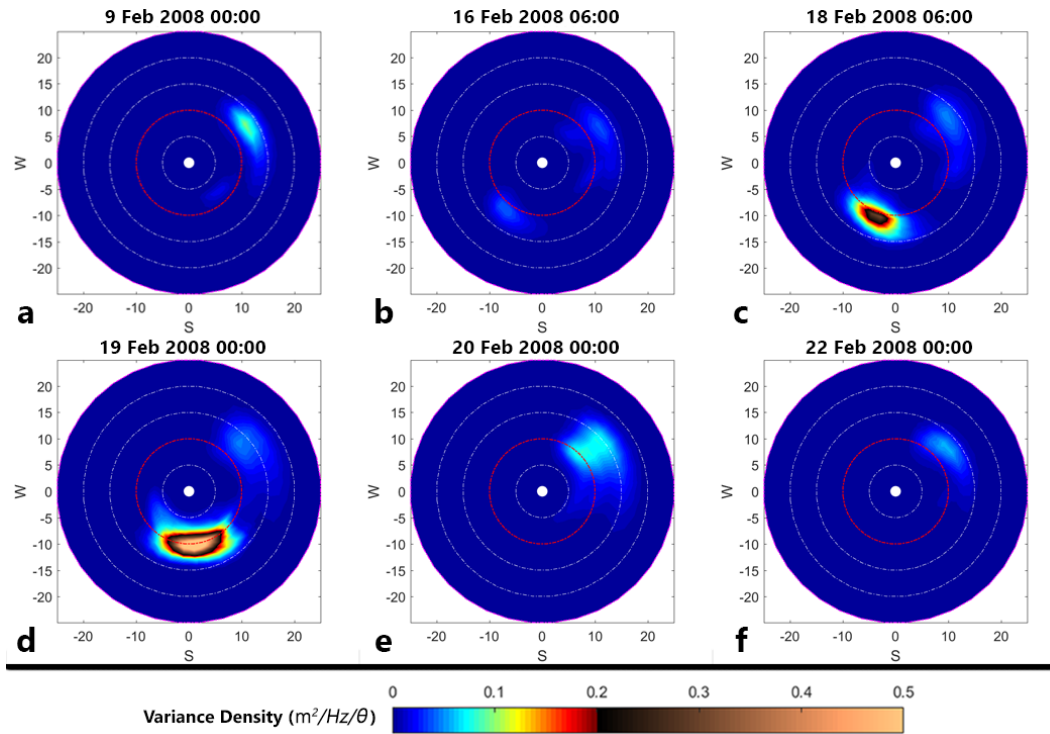## Time evolution of spectrum profiles during TC Nicholas

Figure 19: (a-f) Plot of time evolution of 2D spectrum profile at a certain time during the event (indicated on the title). A common scale is used. **Note**: the axis indicates the wave period [s] and white lines are corresponding to 5-sec contour lines, except red line for 10 sec and magenta line for 25 sec. The S stands for southward direction and W for westward direction. The colour indicates the amount of wave energy in a variance density.

Figure 19 shows the wave spectra, calculated by SWAN at location Exmouth, at different times, as marked by red-dots in Figure 18. The colours indicate values of wave variance, which is proportional to the wave energy density, per frequency per angle. In Figure 19, the radial axis indicates the wave period ($T = f^{-1}$) rather than frequency f, as in Figure 5. This facilitates the interpretation of shifts of peaks in wave variance density during the events.

The important feature in Figure 19 concerns the spectral location of the peak energy. Before and after the occurrence of TC Nicholas, the wave spectra shows that most of the energy is in waves that propagate to the north-east direction (Fig 19a&f). Swell from the southern ocean has a significant impact to this area which always propagates to the north-east (Fig 19a-f). In Figure 19a, a weak sea is observed that has a peak period (T_p) less than 10 sec and those waves mainly propagate to the south-east.

Figure 19b shows the wave spectrum at the time that TC reached its maximum intensity where is

far to the north-east of the location in Exmouth. The spectrum is characterized by two peaks that have a similar magnitude of energy density, but correspond to waves that propagate to opposite directions. Two days later, the wave components propagating to the south-west became stronger as the TC came closer to the buoy location (Figure 19c). Additionally, the sea from TC had a slightly lower T_p value than of swell from the Southern Ocean.

On the next day, the peak spectral energy density of sea from TC Nicholas with the highest value was found for waves that propagated southward (see Figure 19d). This happened because the TC was located to the west of the buoy at that time, hence the wind blew from the north direction (Figure 17). At 20 Feb 2008, the TC made landfall to the south of Exmouth. The veering wind from TC Nicholas generated a less strong wave to the north-east direction, whilst there was a already swell in that same direction. This resulted in a wave spectra with two peaks that correspond to waves with different periods, but propagating in the same direction. After TC Nicholas disappeared, the north-easterly swell was still observed.

3.3   Analysis of Wave Climate

3.3.1   Average, Variance, Skewness and Kurtosis
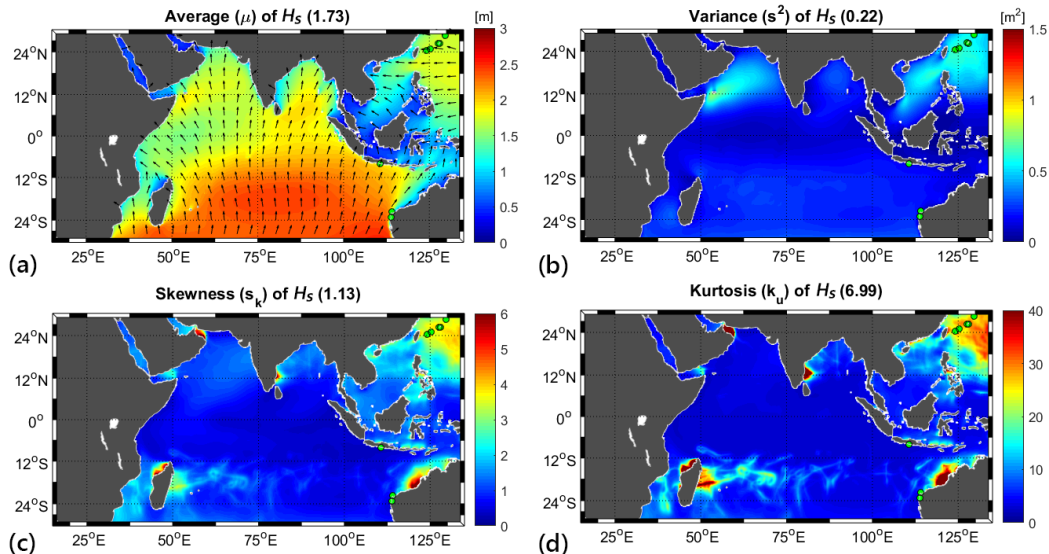
Significant Wave Height (H_s)

Figure 20: Statistical moments of H_s (a) Average [m], (b) Variance [m^2], (c) Skewness, and (d) Kurtosis from a 16 years SWAN simulation. The average of main wave direction ($\bar{\theta}$) is also plotted in panel a as a vector. **Note:** The colour scales are different. The numbers between brackets indicate the mean values from a whole domain.

Figure 20 consists of four panels which show the spatial distribution of four statistical moments of H_s, viz. average $(\overline{H_s})$, variance, skewness, and kurtosis. Figure 20a also shows the average of main wave direction as vectors. The locations of the buoys are indicated by green-circle markers, the same as Figure 12. The mean values for the whole domain are indicated by the number between brackets in the figure titles.

Figure 20a reveals an area with high averaged H_s, i.e. $\overline{H_s} \geq 2$ meter, in the Indian Ocean. This pattern is evenly stretched from the Southern Ocean across the equator into the northern Indian Ocean, especially in the eastern part of the ocean. A similar area with high $\overline{H_s}$ values can be observed at the north of Philippine. The outer islands of Indonesia, e.g. Sumatra, Java, etc., act as a barrier to protect the inland sea, which subsequently results in low $\overline{H_s}$ values this sea.

Figure 20b reveals areas of high variance in H_s near Gulf of Persia and at South China Sea. This possibly occurs due to the monsoon system. The spatial patterns of skewness and kurtosis show similar features with high values on the areas where TC occurred (Figure 20c-20d).

There are five spots with a remarkably high value of both the skewness and the kurtosis. Those spots which are in the north-western of Australia, to the eastern of India, near to the Madagascar and in the north-western of Pacific are because tropical cyclones hit this area. Where, the one which is in the seas of Indonesia perhaps due to strong wind events.
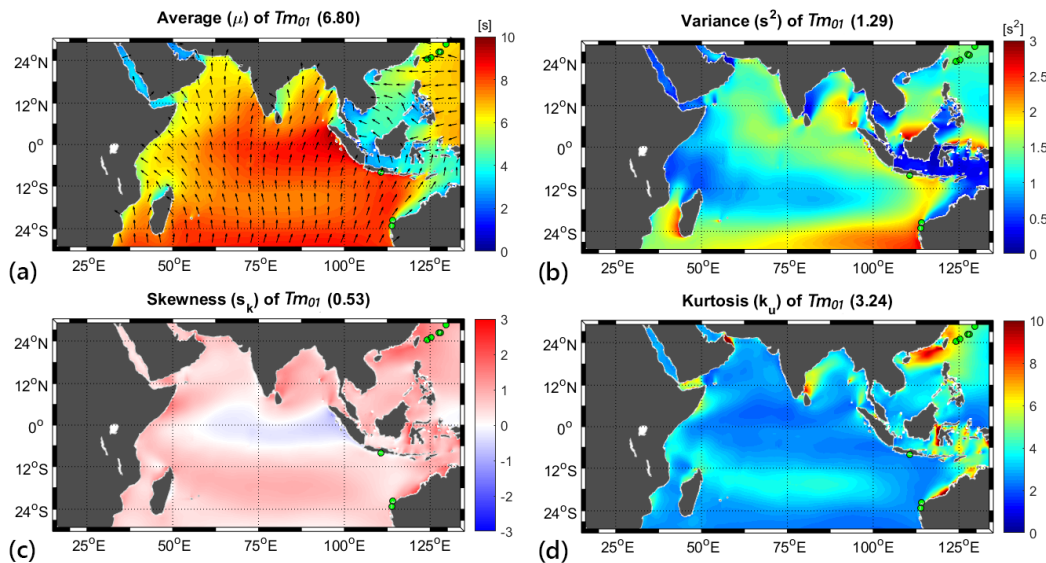
Spectral Mean Wave Period $(T_{m_{01}})$

Figure 21: Same as Figure 20, but for $T_{m_{01}}$. The map of skewness uses a different colour scale to emphasize the negative value near the Equator.

Figure 21 shows four maps of average, variance, skewness, and kurtosis of $T_{m_{01}}$. As in Figure 20, the numbers between brackets indicate the mean value for the whole domain. The spatial distribution of the average value of spectral mean wave period $T_{m_{01}}$ shows a similar pattern as that of H_s (compare Figure 21a with Figure 20a). The waves with a characteristic of long period propagate from the Southern Ocean and Pacific Ocean .

However, the distribution of variance of $T_{m_{01}}$ shows a different pattern to that of H_s (see Figure 21b). The variance of $T_{m_{01}}$ seems to have positive correlation with the average. This can be observed from the areas with a higher value than the domain averaged value (Figure 21a) where it is also higher than averaged variance value (Figure 21b). Along the latitude of 24°S and along the equator both the mean and variance of $T_{m_{01}}$ are high in value. However, along the 12° latitude, small values of both variance and the mean of $T_{m_{01}}$ occur.

3.3.2   Probability Density Functions of Wave Parameters along Sections Line
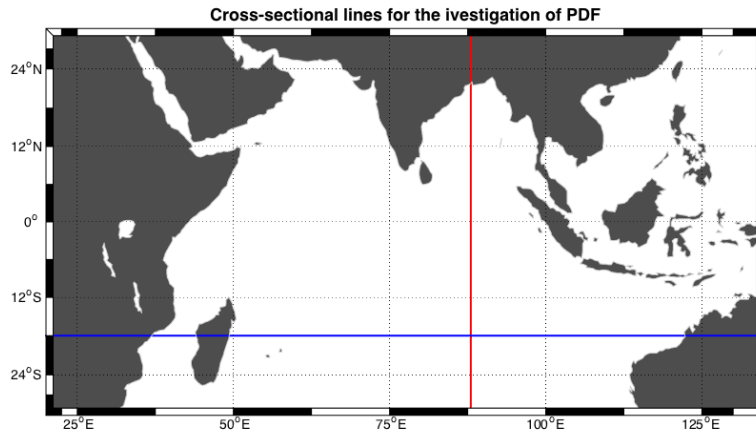

Figure 22: Probability density functions of wave parameters are analysed along longitude of 88°E (red line) and latitude of 18°S (blue line).

Another way to investigate the wave climate is by examining the spatial pattern of probability density function (PDF) of the wave parameters. These were extracted for H_s and $T_{m_{01}}$ along longitudinal and latitudinal lines across of the Indian Ocean (see Figure 22).
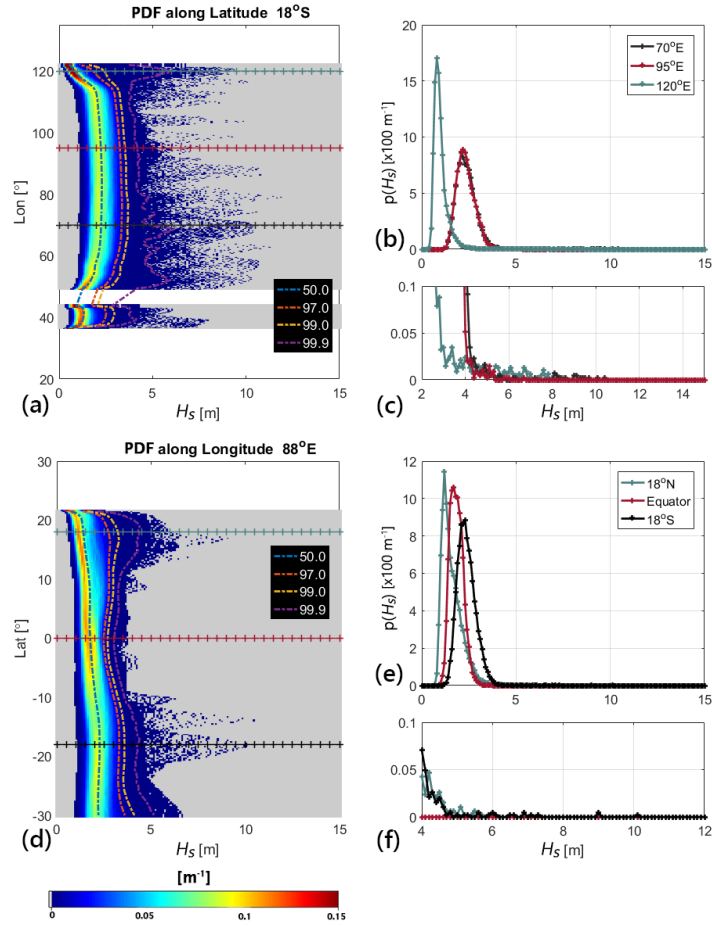
Figure 23: The spatial patterns of PDF of H_s along the latitude of 18°S (**a**) are plotted together with the contour lines of $50^{th}, 90^{th}, 99^{th}$, and 99.9^{th} percentiles. At specific points which are indicated by horizontal lines at panel **a**, i.e. at 70°E, 95°E, and 120°E, PDF of H_s are extracted and plotted in panel **b**. Panel **c** shows the same plot as panel **b**, but in zoomed-in version, to show the tail of distribution. Panel **d** show similar to panel **a**, but along the longitude of 88°E. Panels **e** and **f** work the same as panels **b** and **c**, but for along longitude of 88°E and latitude of 18°N, 18°S, and the equator. Panels **a** and **d** use the same colour scale at the bottom of the figure.

Figure 23 shows probability density functions of H_s along the lines showed in Figure 22. Panels (**a**) and (**d**) show them along latitude 18°S and along longitude 88°E, respectively. Together with those plots, contour lines of $50^{th}, 90^{th}, 99^{th}$, and 99.9^{th} percentiles[10] are shown by dashed lines. On these panels, three specific locations are indicated by horizontal lines. The corresponding PDFs for these specific locations are plotted in panels **b** and **e**, respectively. Panel **c** shows the same PDFs as panel **b**, but focus only for the highest tail with a maximum $p(H_s)$ value of 0.001 m^{-1}. Panel **f** is the same as panel **c**, but for panel **e**.

---

[10]n^{th}-percentile value determines $\hat{x}$ is chosen as such, so that $\int_0^{\hat{x}} p(x)dx = \frac{n}{100}$. Where x is either H_s or T_{m01}.

In panel **a** and **c** of Figure 23, p(H_s) values are showing to have long tails. This is due to high waves that are generated during tropical cyclones. These characteristics can be observed for the 0.1^{th} highest percentiles. This is indicated by the contour line of 99.9^{th}-percentile, which has a different behaviour with the other contour lines, especially in Figure 23a. For example, PDF of H_s for 70°E shows the similar behaviour to that for 95°E (see Figure 23**b**), but the tail of the distribution of H_s for 70°E has more events (Figure 23**c**).

Near the equator, the map of PDF suggest that the wave conditions are relatively calm compared to those at the higher latitudes (Figure 23d). In panel **f**, the occurrence of H_s with values higher than 4 meter can not be observed. Although having a lower average value, the distribution of H_s at latitude 18°N reveals occurrence of high H_s.

Figure 24: Same as Figure 23, but for T_{m01}.

Figure 24 shows the same plots as in Figure 23, but for the probability density function of $T_{m_{01}}$. The distribution of high values shows having a regular pattern, i.e. all contour lines of percentiles of $T_{m_{01}}$ show the similar behaviour, including the 99.9^{th} percentile line. There are no outliers as found in the distribution of H_s (Figure 23). Interestingly, a peculiar pattern is shown near the equator, as the most occurrences of the wave period are shifted to the longer period. Because of this, on panel **e** Figure 24, the PDF of $T_{m_{01}}$ at the equator implies that the skewness of this distribution has a negative value, which is consistent with the map of skewness of $T_{m_{01}}$ (see Figure 21c).

3.4   Analysis of Extreme Waves

The contents of this subsection are meant to answer the third research question on the extreme waves characteristics under influence of tropical cyclones in term of its return value of 20-year and of 50-year.
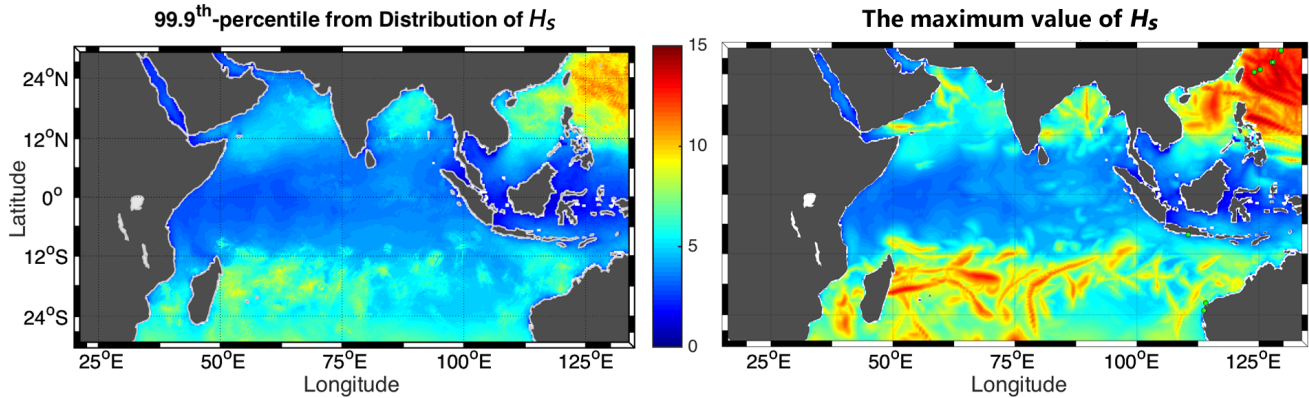

Figure 25: [**Left panel**] Map of height of extreme waves, where the height is defined as the 99.9^{th} percentile of the probability density distribution, p(H_S). [**Right panel**] The maximum value of H_s computed from the time series. **Note:** Both figures use a same colour scale in meters.

To investigate the statistics of extreme waves, the tails of the probability density distributions of H_s were examined. The threshold value for the extreme regime (see Figure 3) that corresponds to the 99.9^{th} percentile is plotted (Figure 25, left panel). In the right panel of Figure 25, the maximum value of 16-years simulations is plotted. Comparing both figures, many similarities and some differences can be observed, mainly in the magnitude of H_s. For some areas, viz. the north Indian Ocean, a difference of magnitude appears when tropical cyclones occur. In these areas, the lower bounds of extreme regime which correspond to 0.1% highest wave events are not the part of high waves which were generated by tropical cyclones. These plots are similar to the findings of Timmermans et al. [2017].
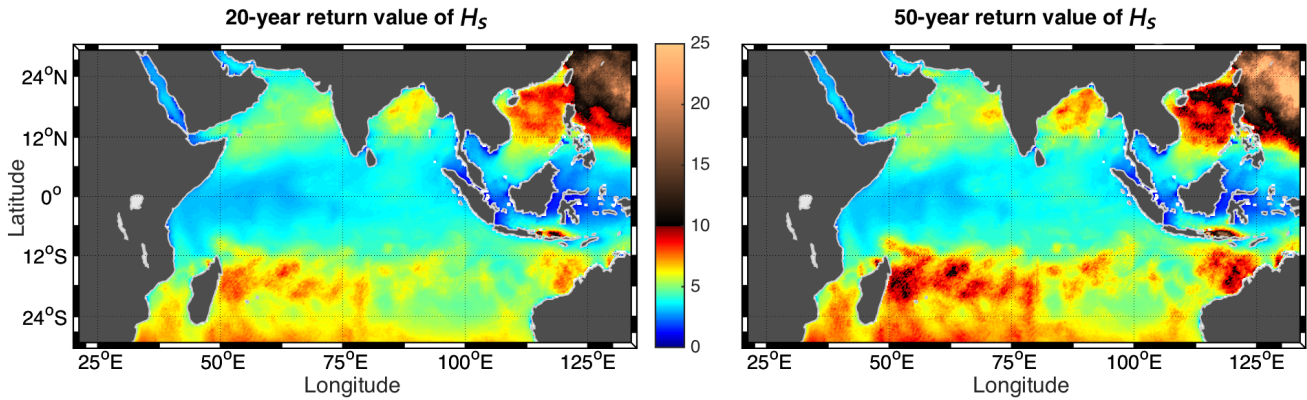
Figure 26: Map of return value of H_s for a period of 20-year [**left**] and 50-year [**right**]. **Note:** Both figures use a same colour scale in meter.

The map of 20-year and 50-year return value of H_s in the tropical Indian Ocean are presented in Figure 26. These values are retrieved using Equation 14 which is derived from the Generalized Pareto Distribution (Equation 12). A significant wave height of ∼20 m is expected to be observed at the north-western Pacific Ocean in the period of 20 years (see left panel Figure 26). While for a duration of 50 years observation, a higher value of ∼25 m significant wave height is suggested to be observed in this area (Figure 26, right panel). Other areas where the high waves are at least 7 meter are mainly in the areas which are affected by tropical cyclones, except one spot to the south of the Celebes island (approximately at 120°E, 8°S). Tropical cyclones do not occur in this area due to its geographical position.
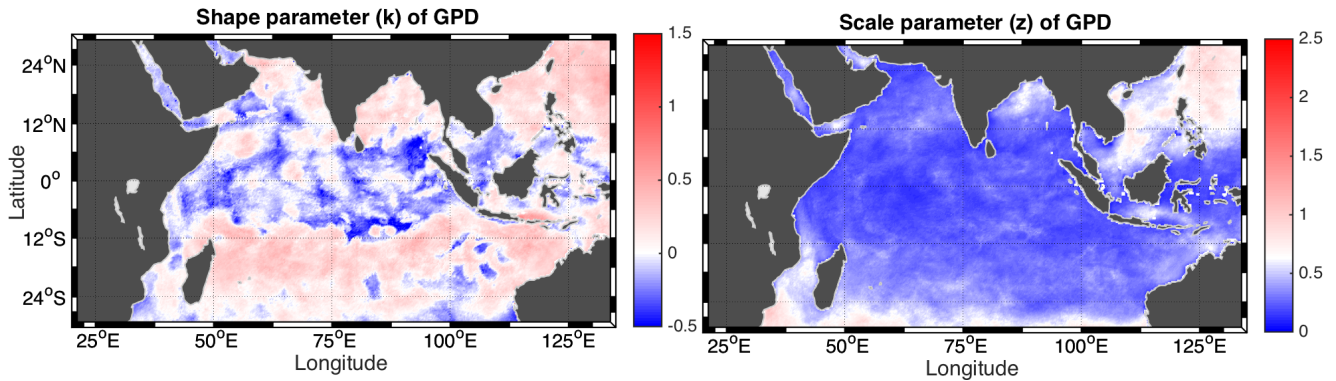

Figure 27: Map of Generalized Pareto Distribution parameter: shape parameter (k) [**left**] and scale parameter (z) [**right**]. **Note:** These figures use different colour scales.

Figure 27 show the spatial distribution of the shape parameter (k) and the scale parameter (z) of GPD. These parameters were chosen so that the GPD could give the best fit to estimate the actual distribution. The signatures of tropical cyclones again can be observed at the map of the

shape parameter. When the tropical cyclones are crossing over the ocean, they produce high waves events. As a consequence, the shape parameters have positive values. Along the equator, the distributions of H_s are suggested to have characteristics of a short tail, as observed with value of $k < 0$, because tropical cyclone have never crossed this area.

From these results, the extreme wave characteristics, by means their return value of a certain year, in the tropical ocean really depend on the occurrence of any single tropical cyclone. For a different time period of simulation will result in the different of map of return value. To make the estimation of return value is independent from the occurrence of tropical cyclones, a longer period of wave simulation is needed.

4 Discussion

16 years wave simulations were done by using input wind of ERA-interim, but corrected locally in the areas which are influenced by tropical cyclones. As a result, an improvement from the results of Caires and Sterl [2005] is shown by capturing tropical cyclones and the resulting extreme waves. Furthermore, the method of using the SWAN model, which is forced by available ERA-interim wind and enhanced at specific locations of tropical cyclones by sub-grid model, is much faster than simulating two models, i.e. atmospheric-ocean model in high resolutions as that of Timmermans et al. [2017]. Additionally, the simulations could cover much larger area than that of Dietrich et al. [2012]. However, for application of extreme wave analysis at coastal area, a further nested domain is needed.

The SWAN simulations show reasonably good results, compared to the observational data from buoys and ERA-interim re-analysis products. The results of wave climate statistics suggest that both of the SWAN simulations and of ERA-interim model are fairly good to estimate the climate conditions. However, in this study, the SWAN simulations is limited to a 16 years period, therefore it is still too uncertain to estimate the 50-year return period. Hence, long time series are needed to estimate the extreme wave conditions better.

4.1 Choosing the Threshold Value for the Peak Over Threshold Method

In the analysis of extreme values, the distribution is filtered, by only considering the peaks which exceed certain threshold, so called the Peak Over Threshold (POT) method. It results in a new distribution which contains only high waves. Then, the probability density function from this new distribution is fitted to a certain distribution function, i.e. the Generalized Pareto Distribution (GPD). In this procedure, there is one uncertain step, which is choosing the threshold value (see section 2.4.2).
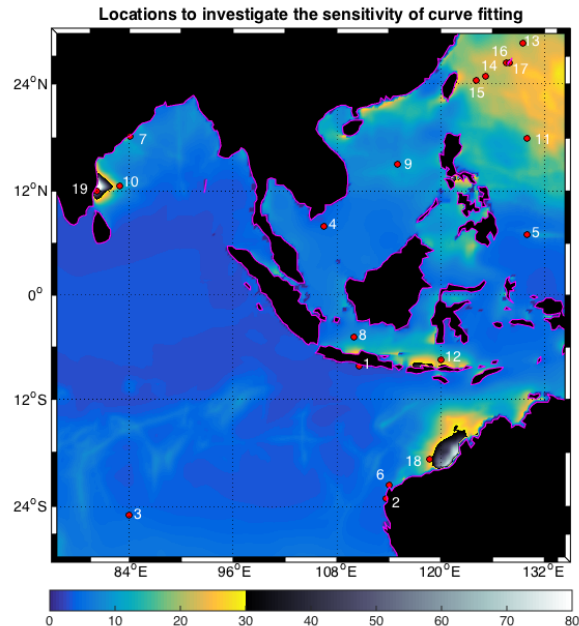
Figure 28: Red-numbered dots represent 19 locations at which probability density distributions of H_s were analysed for the sensitivity of their extreme wave characteristics to the choice of the threshold value for POT. In the background is the map of the kurtosis value from H_s distribution.

Following Bouws et al. [1998], the threshold value is chosen between the 93th and 97th-percentile of the total distribution. This procedure was followed by Caires and Sterl [2005] in their estimation of the 100-year return value of significant wave height. The result shows that the fitted distributions yield a good agreement with 90% of a whole domain. In Timmermans et al. [2017], this threshold was up to 99.97^{th}-percentile at a specific location in order to get the best fit. The corresponding threshold values to that of 93th and 97th-percentile can be very different, especially if the distribution has a long tail. This kind of distribution is commonly observed in the areas frequently influenced by tropical cyclones (see Figure 23). Therefore, a further analysis on the sensitivity of choosing the threshold value was done.

The investigation was done by considering many distributions in the locations which have different values of kurtosis for the distribution of H_s. The reason behind this choice is that the kurtosis is measured on the length of a tail from a distribution (see section 2.4.2). Therefore, it is expected that the threshold value in the POT method will be a function of kurtosis. Consequently, cumulative density distribution of H_s from 19 locations were analysed with regard of the sensitivity of fitting their high tails for different threshold values, i.e. in the range of 0.8 to 1 in term of their quantiles[11] (see Figure 28). From the results of different threshold values, the root-mean-square

---

[11]The value of quantile of probability distribution equals to its percentile value divided by 100; Comparing

errors (RMSE) of the fitted of cumulative density distribution (CDF) curve of the GPD and of the actual distribution were computed. Similar to Equation 10b, the RMSE between these two CDFs is computed by,

$$\text{RMSE}_\text{Q} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(P_{\widehat{H_s}}(H_{s,i})_\text{fit} - P_{\widehat{H_s}}(H_{s,i})_\text{obs} \right)^2},$$

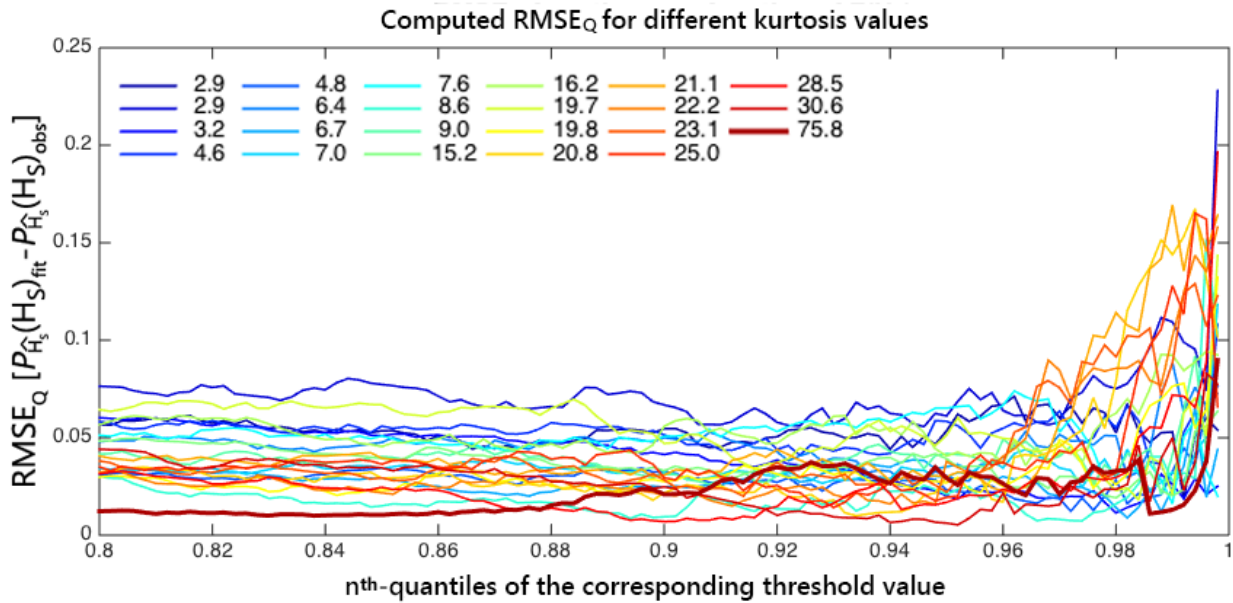where $\widehat{H_s}$ was determined from CDF of the actual distribution.

Figure 29: Root-mean-square error (RMSE) between the actual cumulative density distribution and the fitted GPD are plotted as a function of the threshold value, by means of the quantile value from the actual distribution, for any different kurtosis value. The different colour represent the corresponding kurtosis values.

In Figure 29, the RMSE values of the curve fitting between the actual cumulative density distribution and the fitted GPD are plotted as a function of the quantile of the corresponding threshold value. From this figure, the suggestion of selecting threshold value from $93^{th} - 97^{th}$-percentiles (or $0.93^{th} - 0.97^{th}$-quantiles) will indeed give a minimum error, especially for distributions with low kurtosis values. However, the RMSE value for the threshold value that is higher than 0.97^{th}-quantile is shown with high error value, because there is not enough data in the high tail of the distribution. The RMSE value above can be as high as 0.2.

The results of Figure 29 suggest that the higher the kurtosis value, the lower the RMSE value. The

Footnote 5 with Footnote 10.

distribution with the highest kurtosis (red-thick-line on Figure 29) shows that, for the threshold value corresponding 88th-percentile or lower, the error values are less than 0.02. Instead, the bluish-lines, which correspond to low kurtosis values, are likely to have the minimum error value for the threshold value corresponding larger quantile, up to 0.97.
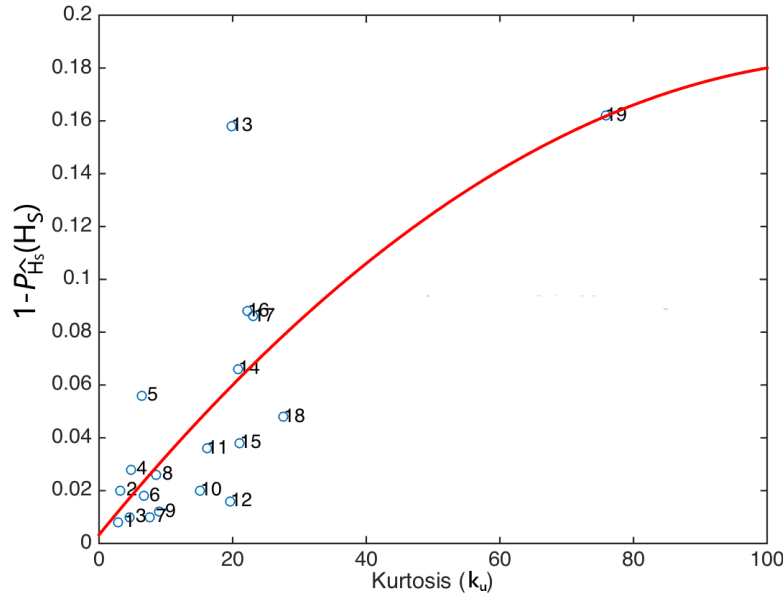

Figure 30: The fraction $1 - P_{\widehat{H}_s}(H_s)$ of extreme waves that correspond to a given the threshold values that is chosen to give the minimum RMSE value are plotted as a function of kurtosis for 19 locations.

Figure 30 shows the fraction $1 - P_{\widehat{H}_s}(H_s)$ (Equation 13), which will result in the optimum threshold values $(\widehat{H}_s)$, as a function of the kurtosis. The 19 corresponding locations picked from the domain (see Figure 28) are plotted in blue markers, together with their numbers. The polynomial curve to fit these points are plotted by the red line. The fit function for the relationship between higher quantile fraction that correspond to the threshold value and kurtosis values is written as,

$$1 - P_{\widehat{H}_s}(H_s) = -0.000013k_u^2 + 0.0031k_u + 0.0033. \tag{15}$$

From the new formulation of threshold value (Equation 15), the maximum RMSE value of 0.04 is found. This is better than choosing a fix threshold value of $93^{th} - 97^{th}$-percentile, as was suggested by Bouws et al. [1998] and Caires and Sterl [2005]. Based on this outcome, the goodness of fit is lower by using RMSE values between two cumulative density distributions of 0.06 (see Figure 29). In this thesis, Equation 15 is used to estimate the threshold values, by means of the quantile of the distribution, for the POT method.

4.2   Comparison of Model Results with Dataset of Other Settings

Comparison between SWAN simulations: $1°$ vs. $0.25°$

Simulations for $1°$ grid resolution are done for the global domain and mainly for generating input waves for the boundary conditions of the tropical domain (section 2.3). The differences between the two models only concern the resolution of the wave model and the size of the domain. These models used the same input wind field that is from combination of wind from ERA-interim and the sub-grid model, thus the tropical cyclones were resolved in the wind data.

Figure 31: Map of maximum significant wave height obtained from 16-years simulations in resolution of $1°$ on the global view [**left panel**] and its differences to the outputs of $0.25°$ resolution for the tropical domain [**right panel**], red colour means $0.25°$ resolution model produced a higher maximum value than that in the $1°$ resolution model.

Figure 31 shows maps of the maximum value of H_s from the results of the SWAN simulations in the global domain and their difference with respect to the $0.25°$ SWAN simulations, which were focused on the tropical domain. In the left panel, the global view of the maximum H_s is shown. The high values of the maximum significant waves heights are observed near latitude $60°$ in both hemispheres. High wave events due to tropical cyclones can be seen near the latitude of $15°$ in both hemisphere, except in the South Atlantic and the South-east Pacific Ocean.

By comparing the output of SWAN with the $1°$ resolution with that of SWAN simulations with a resolution of $0.25°$ (see Figure 31, right panel), some differences can be spotted that are mainly showing underestimation of $1°$ resolution's model in the areas where tropical cyclones occur. Although using the same wind forcing, the SWAN simulations with a lower resolution were unable to produce high waves because of they do not resolve.

4.3   Recommendation for Future Research

First of all, the simulation for 16 years is relatively too short to determine the extreme wave, by means its return value of 20-year or 50-year. Thus, simulations with a longer period are proposed. Although, the primary data for generating tropical cyclone wind field is not complete, by means all parameters that are needed in Equation 5a are recorded for the period before 2001. The other parameters can be approximated from the available parameters by some empirical relationships.

Comparison to ERA5

Nowadays, a new brand of ERA series has been announced by ECMWF, namely ERA-5. The ERA-5 has a higher resolution than ERA-interim by 31 km and 0.36° of spatial resolution for atmospheric and wave model, respectively (`https://software.ecmwf.int/wiki/display/CKB/ERA5+data+documentation`). Regarding the temporal resolution, ERA5 provides hourly data, instead of the 3-hourly data that are provided by ERA-interim. Since July 2017, ERA-5 has completed wave simulations the global earth surface and atmospheric phenomena from 2010 onwards. ERA-5 will be continued this calendar year to provide data which cover similar time period as ERA-interim (1979-present).

Using the high resolution of ERA-5, a relatively small and intense phenomena, such as tropical cyclones and related extreme waves, are captured. Therefore, it is very important to use ERA-5 in the future. Especially using ERA5 to investigate the effectiveness of the coupling method of ERA-interim wind field and parametric TC wind field to simulate wave, as was done in this thesis.
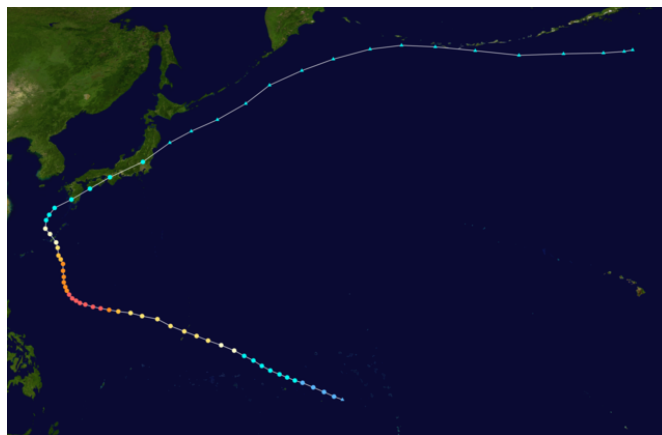

Figure 32: The track of TC Vongfong occurred in the north-western Pacific Ocean during 2-16 October 2014. Its intensity increased up to category 5 (Super Typhoon) when the eye located closely to Okinawa, Japan. Source: `https://en.wikipedia.org/wiki/2014_Pacific_typhoon_season`

One proposed interesting case is to simulate tropical cyclone Vongfong that occurred at October 2014. The TC Vongfong was passing closely to wave buoys in the Okinawa, Japan.

Statistical Analysis on Extreme Wave Periods and Wind Speeds

Another suggestion for the future work is to conduct the statistical analysis of the wave period and the wind speed. In this thesis, the statistical analysis of the extreme events has been done only for the significant wave height (H_s). It is also interesting to investigate at extreme wave periods and wind speeds. As mentioned in section 2, the behaviour of wind waves on long time scales of order months or years, will be the response to the random behaviour of the wind. Therefore, the extreme events of H_s, presented in the section 3.4, will probably relate to extreme wind speeds events.

Additionally, the only wave period that is considered in this thesis is the spectral mean wave period ($T_{m_{01}}$). Meanwhile, there are many definitions of wave periods that could be adopted (see Appendix A) and analysed by statistical method. Although it is not common, the extreme analysis also can be done for these parameters.

5 Conclusions

The main findings of this study are the following:

1. The SWAN wave model with a resolution of 0.25°, being forced by winds with a resolution of 1° from the ECMWF atmospheric model, but locally corrected for the presence of tropical cyclones as a sub-grid model, yields a wave climatology and statistics of extreme waves that are comparable with those obtained from the high resolution wave-atmosphere model of Timmermans et al. [2017]. The main advantage of the SWAN model is that it is much faster since it runs only for one model with high spatial resolution.

2. The time evolution of wave spectra during the tropical cyclones, whose track is crossing close to the location of a buoy, is observed with a characteristic of wave energy with two peaks, i.e. for seas and swell. The peak of wave energy, which corresponds to the seas of the tropical cyclone, has a shorter wave period compared to the one that of corresponds to propagating swell from the ocean.

3. The maps of 20-year and 50-year return value of the extreme significant wave height are shown to have high values in the area where tropical cyclones occur. The distributions of high values in these maps depend on the occurrence of any individual tropical cyclones. For this reason, a longer period of wave simulation is needed.

Acknowledgement

A   Appendix A : Wave Parameters

From output of SWAN (wave spectra $E(\sigma, \theta)$) mainly based on moment of spectrum (see Equation 1):

- **Significant Wave Height** (H_s) In the discrete sample data, it is used that H_s is the average height of the one-third highest waves on the record. Thus,

$$H_s = \frac{1}{N/3} \sum_{j=1}^{N/3} H_j,$$

where H_j are wave heights ordered from large to small values and N is the number of wave heights in the record. Significant wave height can also be calculated from the probability density distribution of wave height (p(H)):

$$H_s = H_{1/3} = \frac{\int_{H^*}^{\infty} \mathrm{Hp(H)}d\mathrm{H}}{\int_{H^*}^{\infty} \mathrm{p(H)}d\mathrm{H}} \quad \text{where} \int_{H^*}^{\infty} \mathrm{p(H)}d\mathrm{H} = \frac{1}{3}.$$

In SWAN, it is formulated as

$$H_s = 4\sqrt{\int \int E(\sigma, \theta)d\sigma d\theta}.$$

- **Mean zero up-crossing wave period** (T_z) is defined as the average of wave period that is calculated by zero up-crossing method. So, it defines a wave period as the length of time needed to the surface elevation up-crossing a consecutive zero level, i.e. from a lower to a higher value. From wave spectra, it can be calculated as the ratio of zeroth order moment (m_0) and second-order moment (m_2), also called by $T_{m_{02}}$, where

$$T_z = T_{m_{02}} = 2\pi \left(\frac{\int \int E(\sigma, \theta)d\sigma d\theta}{\int \int \sigma^2 E(\sigma, \theta)d\sigma d\theta} \right)^{\frac{1}{2}}.$$

- **Mean spectral wave period** $(T_{m_{01}})$ is obtained from the integral of the first-order moment of the wave spectrum:

$$T_{m_{01}} = 2\pi \left(\frac{\int \int E(\sigma, \theta)d\sigma d\theta}{\int \int \sigma E(\sigma, \theta)d\sigma d\theta} \right).$$

- **Peak period** (T_p) is the inverse of f_p, which is frequency band at which the highest energy of frequency wave spectrum $(E(\sigma) = \int E(\sigma, \theta)d\theta)$ is located, or

$$T_p = f_p^{-1}.$$

- **Mean spectral wave period** $(T_{m_{-10}})$ is obtained by comparing the integral over frequency and direction of the moment of -1 order and the zeroth order moment:

$$T_{m_{-10}} = 2\pi \left(\frac{\int \int \sigma^{-1} E(\sigma, \theta) d\sigma d\theta}{\int \int E(\sigma, \theta) d\sigma d\theta} \right).$$

In ERA-interim, $T_{m_{-10}}$ is defined as the mean wave period.

B   Appendix B : SWAN settings

Script for Global Domain

```
PROJECT 'GLO_2008' 'GL08'
$
$ Script of SWAN Simulation written in GLO_2008.swn
$ ─────────────────────────────────────────────── $
$
$ ──────── SETTING
SET NAUT
MODE DYN TWOD
COORD SPHE CCM  REP
$
$ ────────── DOMAIN SPECIFICATION
CGRID REG −180.000 −81.000 0.0 360 162  361  162 CIRCLE 36 flow=0.04 fhigh=1.00
$
$ ───────────── BATHYMETRY
INPGRID BOT REG −180.125 −81.125 0 1441  649 0.250 0.250
READINP BOT 1.0  '..\XX_BATH\save_ETOPO1_25.txt' 2 0 FREE
$
$   ────────── WIND
INPGRID WIND REG −180.00 −81.00 0  479  216 0.750 0.750 NONSTAT 20071020.000000 3 HR
READINP WIND 1.0 SERIES '..\YY_WIND\ListWind2008.txt' 2 0 FREE
$
$ ─────── BOUNDARY SPECIFICATION
BOUND SHAP JON 3.3 PEAK DSPR POW
$
$ ────── INITIAL STATE
INITIAL DEFAULT
$
$ ────── PHYSICAL SETTING
GEN3 KOMEN AGROW
TRIADS
BREAKING
DIFFRAC
FRICTION
$
$ ── NUMERICAL SETTING
PROP GSE 1 DAYS
NUM ACCUR npnts=99.5 NONSTAT mxitns=5 DIRIMPL cdd=1
$ Cut−off swell for > 10 s
QUANTITY Hswell fswell=0.1
$
$ ──── NEST DOMAIN
$ WIO
NGRID 'NWIO' 20.0 −30.0 0.0 60.0 60.0 40 40
NESTOUT 'NWIO' '../02_WIO/BND_WIO.nc' OUTPUT 20071101.000000 3 HR
$ EIO
NGRID 'NEIO' 75.0 −30.0 0.0 60.0 60.0 40 40
NESTOUT 'NEIO' '../02_EIO/BND_EIO.nc' OUTPUT 20071101.000000 3 HR
$ ──── OUTPUT SPECIFICATION
BLOCK 'COMPGRID' NOHEADER 'BASE000.mat' LAY 1 XP YP DEPTH OUTPUT 20071101.000000 900 DAY
BLOCK 'COMPGRID' NOHEADER 'MAIN000.mat' LAY 1 HSIG TPS TM01 DIR OUTPUT 20071101.000000 6 HR
BLOCK 'COMPGRID' NOHEADER 'WIND000.mat' LAY 1 WIND OUTPUT 20071101.000000 6 HR
$
COMPUTE NONSTAT 20071020.000000 1 HR 20081101.000000
$
TEST 1,0
STOP
```

Script for East Indian Ocean (EIO) Domain

```
PROJECT 'EIO_2008' 'EI08'
$
$ Script of SWAN Simulation written in EIO_2008.swn
$ ———————————————————————————————————————— $
$
$ ———————— SETTING
SET NAUT
MODE DYN TWOD
COORD SPHE CCM
$
$ ———————— DOMAIN SPECIFICATION
CGRID REG 75.000 −30.000 0.0  60  60  240  240 CIRCLE 36 flow=0.04 fhigh=1.00
$
$ ———————————— BATHYMETRY
INPGRID BOT REG −180.125 −81.125 0 1441  649 0.250 0.250
READINP BOT 1.0 '..\XX_BATH\save_ETOPO1_25.txt' 2 0 FREE
$
$ ——————— WIND
INPGRID WIND REG 15.00 −30.00 0  480  240 0.250 0.250 NONSTAT 20071020.000000 3 HR
READINP WIND 1.0 SERIES '..\ZZ_WIND\ListWind2008_C.txt' 2 0 FREE
$
$ ——————— BOUNDARY SPECIFICATION
BOUND SHAP JON 3.3 PEAK DSPR POW
BOUND NEST 'BND_EIO.nc' CLOSED
$
$ —————— INITIAL STATE
INITIAL DEFAULT
$
$ —————— PHYSICAL SETTING
GEN3 KOMEN AGROW
TRIADS
BREAKING
DIFFRAC
FRICTION
$
OFF BNDCHK
$
$ —— NUMERICAL SETTING
PROP GSE 1 DAYS
NUM ACCUR npnts=99.5 NONSTAT mxitns=5 DIRIMPL cdd=1
$ Cut−off swell for > 10 s
QUANTITY Hswell fswell=0.1
$ ———— OUTPUT SPECIFICATION
$ Exmouth Buoy
POINTS 'POIN' 114.0986 −21.6995
$
SPECTOUT 'POIN' 'EIO.sp2' OUTPUT 20071101.000000 1 HR
$
TABLE 'POIN' HEAD 'EIO_1.tbl' HSIG HSWELL DIR TM01 TPS TM02 WIND OUTPUT 20071101.000000 1 HR
$
TABLE 'POIN' HEAD 'EIO_2.tbl' PTHSIG OUTPUT 20071101.000000 1 HR
$
TABLE 'POIN' HEAD 'EIO_3.tbl' PTRTP OUTPUT 20071101.000000 1 HR
$ Saveas BLOCK
BLOCK 'COMPGRID' NOHEADER 'BASE001.mat' LAY 1 XP YP DEPTH OUTPUT 20071101.000000 900 DAY
BLOCK 'COMPGRID' NOHEADER 'MAIN001.mat' LAY 1 HSIG TM01 DIR OUTPUT 20071101.000000 3 HR
BLOCK 'COMPGRID' NOHEADER 'FREQ001.mat' LAY 1 TPS TM02 TMM10 OUTPUT 20071101.000000 3 HR
BLOCK 'COMPGRID' NOHEADER 'WIND001.mat' LAY 1 WIND OUTPUT 20071101.000000 3 HR
$
COMPUTE NONSTAT 20071020.000000 1 HR 20081101.000000
$
TEST 1,0
STOP
```

Script for West Indian Ocean (WIO) Domain

```
PROJECT 'WIO_2008' 'WI08'
$
$ Script of SWAN Simulation written in WIO_2008.swn
$ ----------------------------------------------------- $
$
$ -------- SETTING
SET NAUT
MODE DYN TWOD
COORD SPHE CCM
$
$ ---------- DOMAIN SPECIFICATION
CGRID REG 20.000 -30.000 0.0  60  60  240  240 CIRCLE 36 flow=0.04 fhigh=1.00
$
$ ------------- BATHYMETRY
INPGRID BOT REG -180.125 -81.125 0 1441  649 0.250 0.250
READINP BOT 1.0  '..\XX_BATH\save_ETOPO1_25.txt' 2 0 FREE
$
$  ---------- WIND
INPGRID WIND REG 15.00 -30.00 0  480  240 0.250 0.250 NONSTAT 20071020.000000 3 HR
READINP WIND 1.0 SERIES  '..\ZZ_WIND\ListWind2008_CS.txt' 2 0 FREE
$
$ ------- BOUNDARY SPECIFICATION
BOUND SHAP JON 3.3 PEAK DSPR POW
BOUND NEST 'BND_WIO.nc' CLOSED
$
$ ----- INITIAL STATE
INITIAL DEFAULT
$
$ ------ PHYSICAL SETTING
GEN3 KOMEN AGROW
TRIADS
BREAKING
DIFFRAC
FRICTION
$
OFF BNDCHK
$
$ -- NUMERICAL SETTING
PROP GSE 1 DAYS
NUM ACCUR npnts=99.5 NONSTAT mxitns=5 DIRIMPL cdd=1
$ Cut-off swell for > 10 s
QUANTITY Hswell fswell=0.1
$ ---- OUTPUT SPECIFICATION
$ Durban Buoy
POINTS 'POIN' 32.2500 -28.8300
$
SPECTOUT 'POIN' 'WIO.sp2' OUTPUT 20071101.000000 1 HR
$
TABLE 'POIN' HEAD 'WIO_1.tbl' HSIG HSWELL DIR TM01 TPS TM02 WIND OUTPUT 20071101.000000 1 HR
$
TABLE 'POIN' HEAD 'WIO_2.tbl' PTHSIG OUTPUT 20071101.000000 1 HR
$
TABLE 'POIN' HEAD 'WIO_3.tbl' PTRTP OUTPUT 20071101.000000 1 HR
$ Saveas BLOCK
BLOCK 'COMPGRID' NOHEADER 'BASE001.mat' LAY 1 XP YP DEPTH OUTPUT 20071101.000000 900 DAY
BLOCK 'COMPGRID' NOHEADER 'MAIN001.mat' LAY 1 HSIG TM01 DIR OUTPUT 20071101.000000 3 HR
BLOCK 'COMPGRID' NOHEADER 'FREQ001.mat' LAY 1 TPS TM02 TMM10 OUTPUT 20071101.000000 3 HR
BLOCK 'COMPGRID' NOHEADER 'WIND001.mat' LAY 1 WIND OUTPUT 20071101.000000 3 HR
$
COMPUTE NONSTAT 20071020.000000 1 HR 20081101.000000
$
TEST 1,0
STOP
```

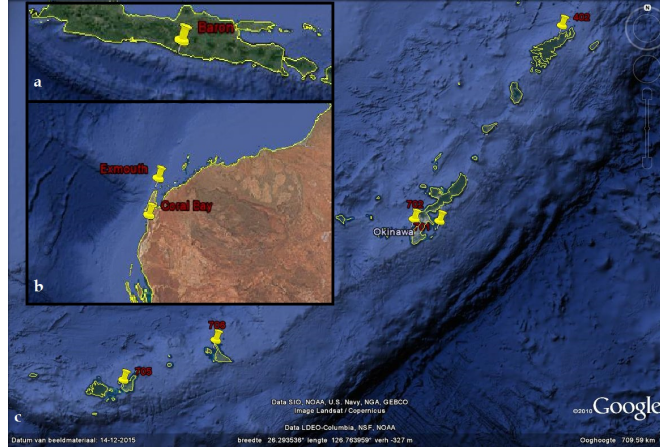# C   Appendix C : Verification for Other Buoys

Figure 33: Map of the locations of buoys similar to Figure 12, but more details geographically. Source: Google Earth.

The SWAN simulations were also verified with respect to other buoys. Most buoys from Japan, which measured near Okinawa islands, are located at the shadow zones of the islands, except Buoy No. 701. Since these islands were not captured in the simulations, the outputs from the SWAN simulations are expected to be higher than observational data in average.

Table 3: Goodness of SWAN simulations with respect to observational data measured from many buoys and ERA-interim re-analysis products, for optimum values of ξ and χ.

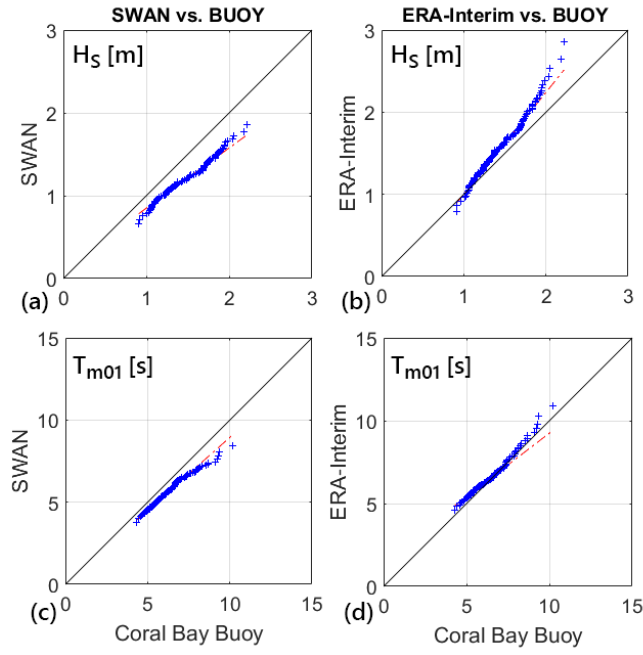| Wave Parameter	Buoy Name	SWAN vs. Buoy of Exmouth			SWAN vs. ERA-Interim			ERA-Interim vs. Buoy of Exmouth		
		R	RMSE [m]	Bias [m]	R	RMSE [m]	Bias [m]	R	RMSE [m]	Bias [m]
Wave Height $[H_s]$	Coral Bay	0.773	0.240	0.138	0.889	0.430	0.390	0.754	0.339	-0.252
	Baron Bay	0.559	0.417	-0.083	0.816	0.277	-0.185	0.360	0.502	0.102
	402	0.635	0.841	-0.507	0.925	0.319	0.017	0.740	0.796	-0.524
	701	0.660	0.712	-0.488	0.923	0.291	-0.07	0.724	0.654	-0.411
	702	0.691	0.900	-0.700	0.947	0.298	-0.102	0.691	0.859	-0.599
	705	0.633	1.385	-1.273	0.897	0.363	0.053	0.717	1.464	-1.326
	706	0.611	1.212	-1.085	0.901	0.399	0.093	0.740	1.320	-1.178
	Buoy Name	R	RMSE [s]	Bias [s]	R	RMSE [s]	Bias [s]	R	RMSE [s]	Bias [s]
Wave Period	Coral Bay	0.828	1.028	0.723	0.846	0.969	0.800	0.732	0.848	-0.077
	Baron Bay	0.555	1.806	-1.478	0.787	0.900	0.487	0.287	2.421	-1.964
	402	0.417	1.168	-0.170	0.777	0.761	-0.046	0.631	0.789	-0.124
	701	0.548	1.085	-0.001	0.772	0.776	-0.175	0.688	0.823	0.173
	702	0.384	1.401	-0.508	0.769	0.781	-0.234	0.526	1.097	-0.274
	705	0.295	1.988	-1.566	0.768	0.749	0.076	0.458	1.878	-1.642
	706	0.306	1.647	-0.989	0.765	0.794	0.160	0.519	1.504	-1.149

Q-Q plots

Coral Bay

Figure 34: Same as Figure 15, but for Coral Bay Buoy.
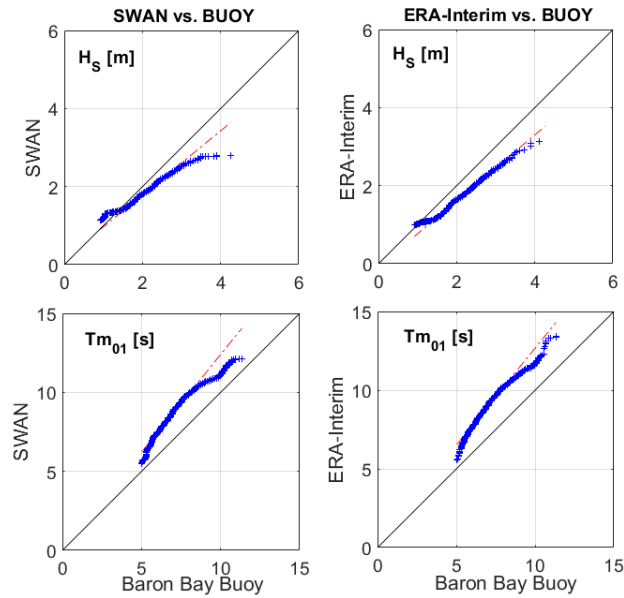
Baron Bay

Figure 35: Same as Figure 15, but for Baron Bay Buoy.
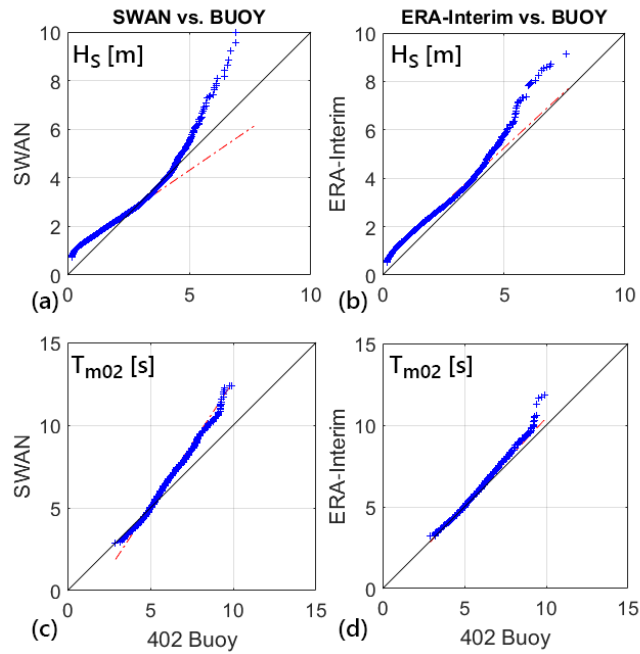
Buoy No.402, Japan

Figure 36: Same as Figure 15, but for Buoy No. 402, Okinawa.
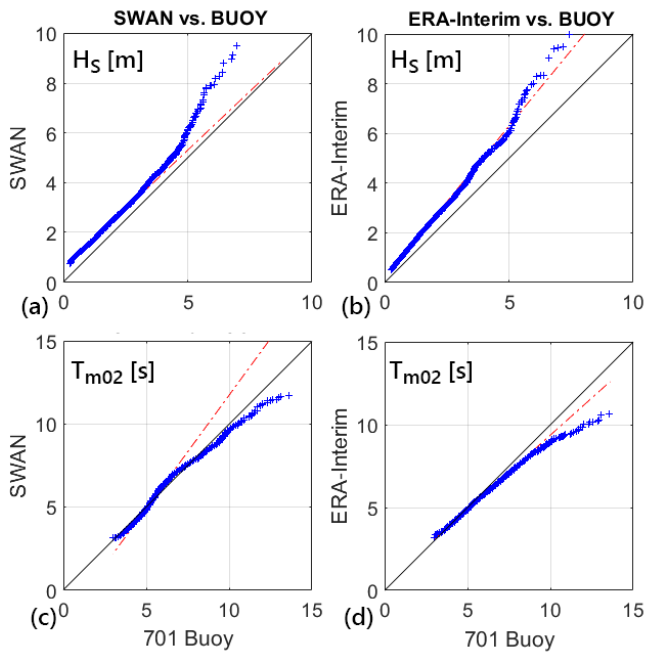
Buoy No.701, Japan

Figure 37: Same as Figure 15, but for Buoy No. 701, Okinawa.
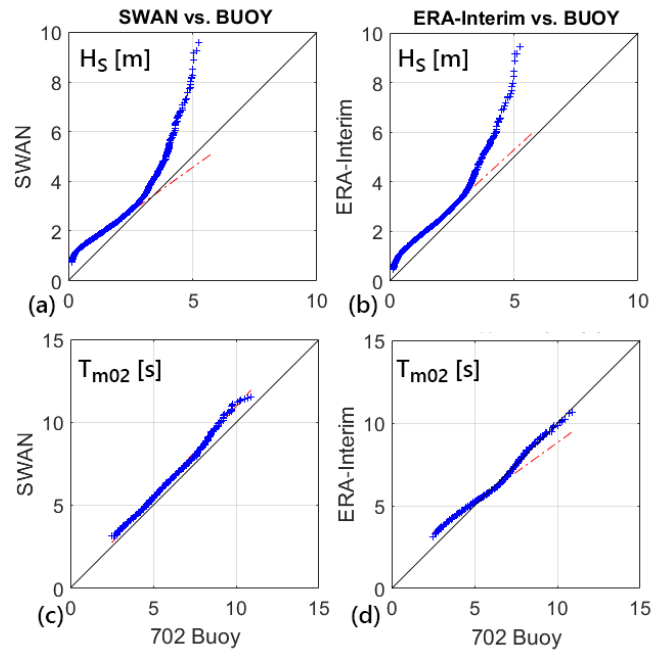
Buoy No.702, Japan

Figure 38: Same as Figure 15, but for Buoy No. 702, Okinawa.
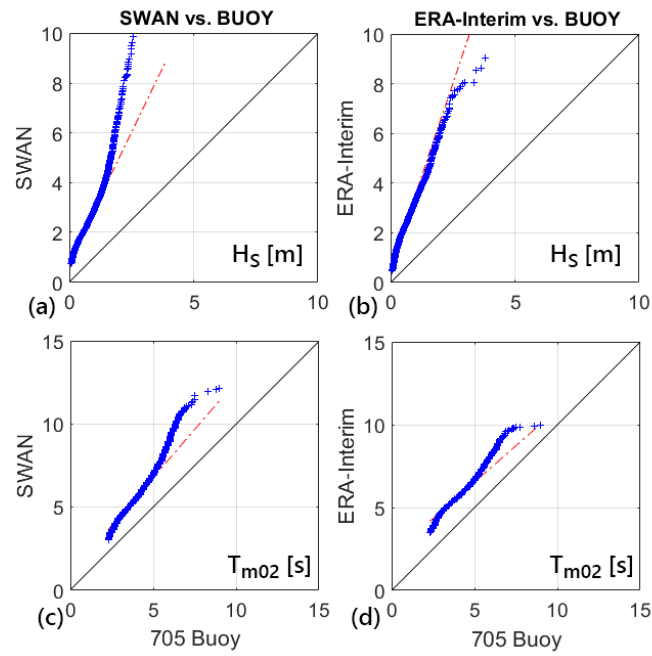
Buoy No.705, Japan

Figure 39: Same as Figure 15, but for Buoy No. 705, Okinawa.
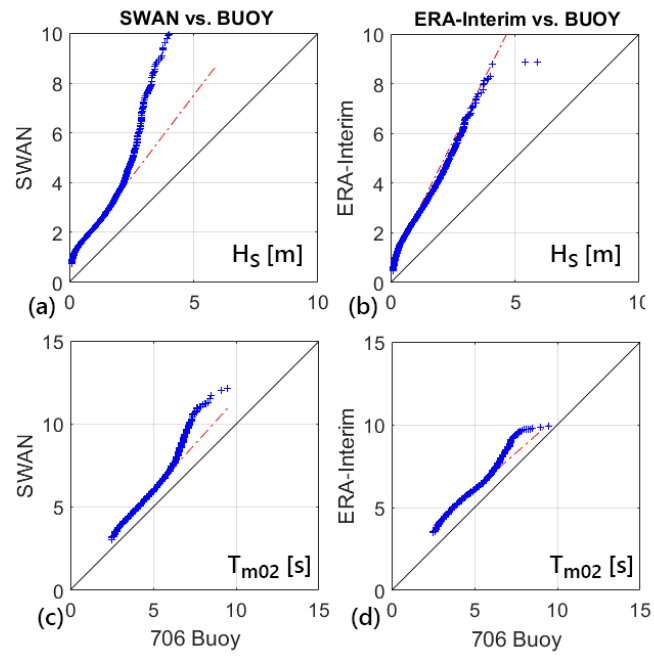
Buoy No.706, Japan

Figure 40: Same as Figure 15, but for Buoy No. 706, Okinawa.

D   Appendix D : Time Series of Wave Parameters During Tropical Cyclone Vongfong

In this section, any information related to the tropical cyclone Vongfong, which was occurred in the North-western Pacific Ocean at 2-16 October 2014, is presented,
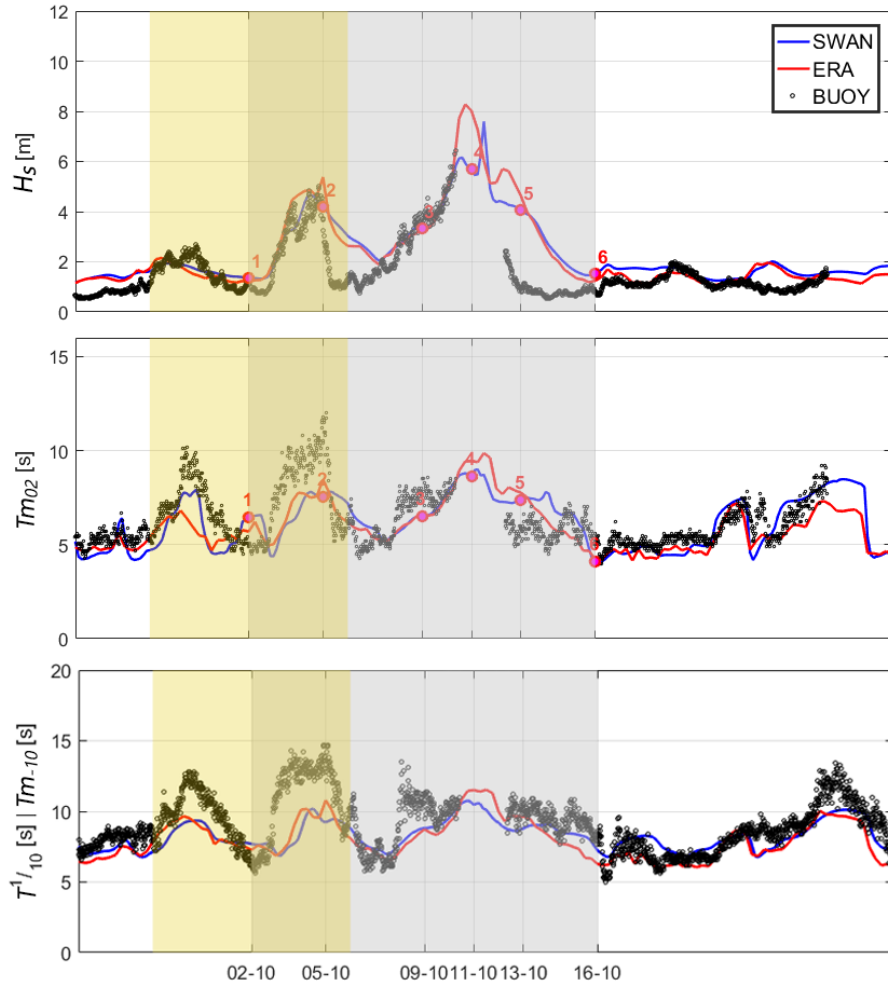

Figure 41: Time series of [**top**] significant wave height H_s, [**middle**] mean zero up-crossing wave period $T_{m_{02}}$ and [**bottom**] mean wave period by means of $T_{m_{-10}}$ are plotted for output of the SWAN simulations (blue line), ERA-interim (red line), and observational data from Buoy No. 701, Okinawa, Japan. The grey shaded area is the time period when TC Vongfong occurred. The red markers with numbers indicate the time when the wave spectra is analysed. The horizontal axis indicates the time of the events, 2-16 October 2014.
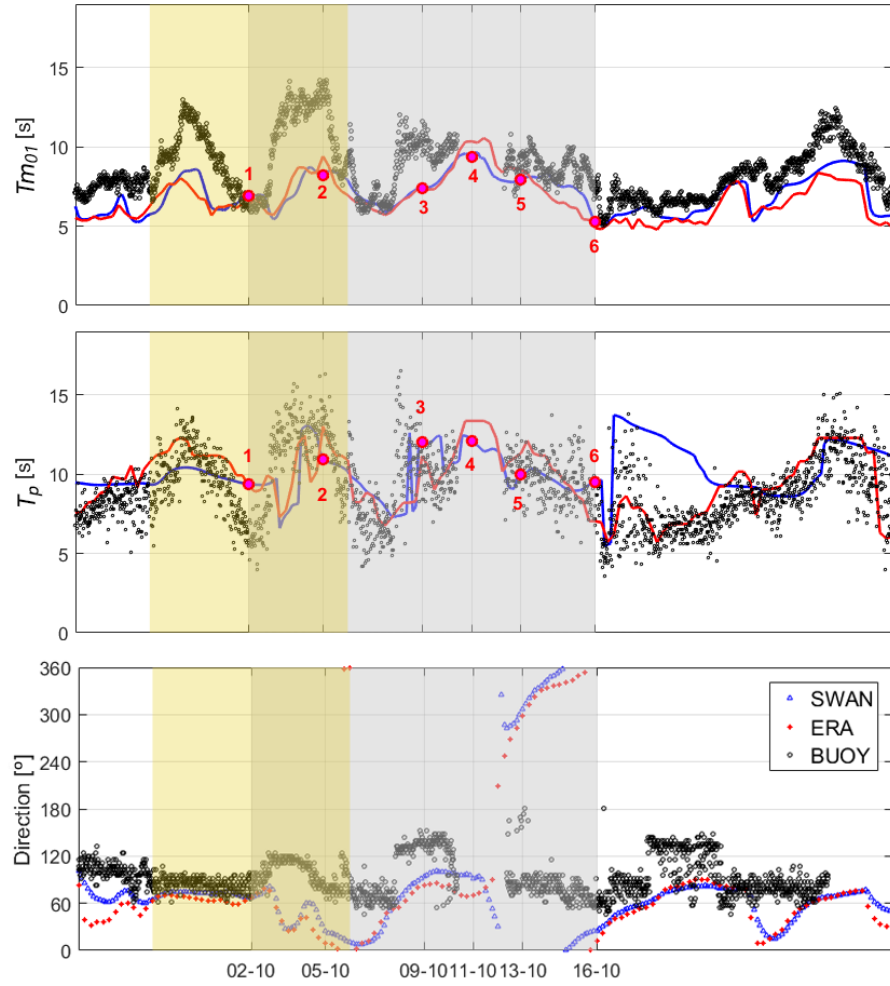
Figure 42: Similar to Figure 41, but for [**top**] mean spectral wave period $T_{m_{01}}$, [**middle**] peak wave period T_p and [**bottom**] main wave direction.
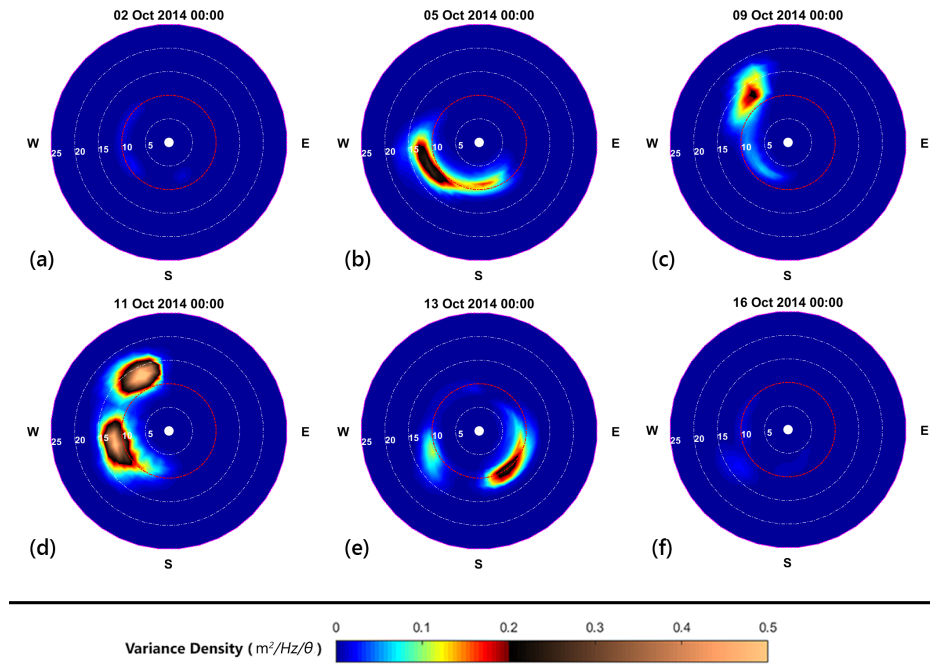
Figure 43: Similar to Figure 19, but for TC VongFong. Panels **a-f** correspond to numbers $1-6$ in Figure 41 and Figure 42, respectively.

References

Berrisford, P., Dee, D., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S., Uppala, S., and Simmons, A. (2011). The ERA-Interim archive, version 2.0.

Booij, N., Ris, R., and Holthuijsen, L. H. (1999). A third-generation wave model for coastal regions: 1. model description and validation. *Journal of geophysical research: Oceans*, 104(C4):7649–7666.

Bouws, E., Draper, L., Shearman, E. D. R., Laing, A. K., Feit, D., Mass, W., Eide, L. I., Francis, P., Carter, D. J. T., and Battjes, J. A. (1998). Guide to wave analysis and forecasting. WMO-no. 702. *World Meteorological Organization.*

Bretschneider, C. L. (1957). Revisions in wave forecasting: deep and shallow water. *Coastal Engineering Proceedings*, 1(6):3.

Brown, J. (2013). *Waves, Tides and Shallow-Water Processes: Prepared by an Open University Course Team.* Elsevier.

Caires, S. (2011). Extreme value analysis: wave data. *Joint WMO/IOC Technical Commission for Oceanography and Marine Meteorology (JCOMM) Technical Report*, (57).

Caires, S. and Sterl, A. (2005). 100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 data. *Journal of Climate*, 18(7):1032–1048.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Holm, E., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., AP, M., Monge-Sanz, B., Moncrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N., and Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, 137(656):553–597.

Dietrich, J. C., Tanaka, S., Westerink, J. J., Dawson, C. N., Luettich, R. A., Zijlema, M., Holthuijsen, L. H., Smith, J. M., Westerink, L. G., and Westerink, H. J. (2012). Performance of the unstructured-mesh, SWAN+ ADCIRC model in computing hurricane waves and surge. *Journal of Scientific Computing*, 52(2):468–497.

Drost, E. J. F., Lowe, R. J., Ivey, G. N., Jones, N. L., and Péquignet, C. A. (2017). The effects of tropical cyclone characteristics on the surface wave fields in Australia's North West region. *Continental Shelf Research*, 139:35–53.

Falnes, J. and Løvseth, J. (1991). Ocean wave energy. *Energy policy*, 19(8):768–775.

Goda, Y. (1979). A review on statistical interpretation of wave data. *Report of the Port and Harbour Research Institute, Japan*, 18:5–32.

Hasselmann, K., Barnett, T., Bouws, E., Carlson, H., Cartwright, D., Enke, K., Ewing, J., Gienapp, H., Hasselmann, D., Kruseman, P., et al. (1973). Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). *Deut. Hydrogr. Z.*, 8(12):1–95.

Holland, G. J. (1980). An analytic model of the wind and pressure profiles in hurricanes. *Monthly weather review*, 108(8):1212–1218.

Holthuijsen, L. H. (2007). *Waves in oceanic and coastal waters*. Cambridge university press.

Janssen, P. (2004). *The interaction of ocean waves and wind*. Cambridge University Press.

Janssen, P. A. (1989). Wave-induced stress and the drag of air flow over sea waves. *Journal of Physical Oceanography*, 19(6):745–754.

Joint Typhoon Warning Center (2009). cited 2009a: 2008 Annual Tropical Cyclone Report.

Kimball, S. K. and Mulekar, M. S. (2004). A 15-year climatology of North Atlantic tropical cyclones. Part I: Size parameters. *Journal of Climate*, 17(18):3555–3575.

Komen, G. J., Cavaleri, L., and Donelan, M. (1994). *Dynamics and modelling of ocean waves*. Cambridge university press.

Komen, G. J., Hasselmann, K., and Hasselmann, K. . (1984). On the existence of a fully developed wind-sea spectrum. *Journal of physical oceanography*, 14(8):1271–1285.

LeBlond, P. H. and Mysak, L. A. (1978). *Waves in the Ocean*. Elsevier.

Longuet-Higgins, M. S. (1957). The statistical analysis of a random, moving surface. *Phil. Trans. R. Soc. Lond. A*, 249(966):321–387.

Longuet-Higgins, M. S. (1975). On the joint distribution of the periods and amplitudes of sea waves. *Journal of Geophysical Research*, 80(18):2688–2694.

Mathiesen, M., Goda, Y., Hawkes, P. J., Mansard, E., Martín, M. J., Peltier, E., Thompson, E. F., and Van Vledder, G. (1994). Recommended practice for extreme wave analysis. *Journal of hydraulic Research*, 32(6):803–814.

Munk, W. H. (1951). Origin and generation of waves. Technical report, Scripps Institution of Oceanography La Jolla California.

Pierson, W. J. and Moskowitz, L. (1964). A proposed spectral form for fully developed wind seas based on the similarity theory of S. A. Kitaigorodskii. *Journal of Geophysical Research*, 69(24):5181–5190.

Rodrigue, J.-P. (2004). Straits, passages and chokepoints: a maritime geostrategy of petroleum distribution. *Cahiers de géographie du Quebec*, 48(135):357–374.

Soares, C. G. and Teixeira, A. P. (2001). Risk assessment in maritime transportation. *Reliability Engineering & System Safety*, 74(3):299–309.

Swail, V. R. and Cox, A. T. (2000). On the use of NCEP–NCAR reanalysis surface marine wind fields for a long-term North Atlantic wave hindcast. *Journal of Atmospheric and oceanic technology*, 17(4):532–545.

SWAN team et al. (2016a). *SWAN technical documentation*. Delft University of Technology., Delft, Holland.

SWAN team et al. (2016b). *SWAN user manuals*. Delft University of Technology., Delft, Holland.

The WAMDI Group (1988). The WAM model—A third generation ocean wave prediction model. *Journal of Physical Oceanography*, 18(12):1775–1810.

Timmermans, B., Stone, D., Wehner, M., and Krishnan, H. (2017). Impact of tropical cyclones on modeled extreme wind-wave climate. *Geophysical Research Letters*, 44(3):1393–1401.

Tolman, H. L. (1991). A third-generation model for wind waves on slowly varying, unsteady, and inhomogeneous depths and currents. *Journal of Physical Oceanography*, 21(6):782–797.

Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V. d., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., et al. (2005). The era-40 re-analysis. *Quarterly Journal of the royal meteorological society*, 131(612):2961–3012.

WAFO-group (2017). *WAFO - A Matlab Toolbox for Analysis of Random Waves and Loads - A Tutorial*. Math. Stat., Center for Math. Sci., Lund Univ., Lund, Sweden.

Zheng, C. W. and Li, C. Y. (2017). Analysis of temporal and spatial characteristics of waves in the Indian Ocean based on ERA-40 wave reanalysis. *Applied Ocean Research*, 63:217–228.

Zijlema, M., Van Vledder, G. P., and Holthuijsen, L. (2012). Bottom friction and wind drag for wave models. *Coastal Engineering*, 65:19–26.