Wave Zonation for the Maltese Islands

Student: Ing. Edward Sultana

Supervisor: Professor Aldo Drago

A dissertation presented to the Faculty of Science of the University of Malta for the Masters of Science in Geoscience

January 2021

Abstract:

Malta is a small island in the middle of the Mediterranean. The sea is an integral part of our daily life and also an important economic sector. Notwithstanding all the advances in technology, the sea also poses its hazards. Studies have shown that worldwide 0.3% of all commercial vessels are lost at sea each year claiming the lives of an average of 1,000 people per year. Though not the only cause of the incidents, waves and bad weather could be the instigators for the incidents to occur. It was also found that in most cases (approximately 66%) the incidents occurred in less than extreme conditions.

The weather forecasts for mariners are mainly based on Significant Wave Height H_s and the Mean Wave Period T_z. Based on this, the study focused on identifying other wave parameters that adversely affect the vessel stability. Besides H_s, wave length, wave steepness and individual wave height were identified to impact vessel stability. The Linear Wave Theory was subsequently applied to the data generated by the *SWAN* wave model, to provide operational and forecasted data for these parameters around the Maltese territorial waters. These two-dimensional plots are easy to interpret and provide the vessel master with added information to enable calculated decisions to be taken based on forecasted operational data.

To guarantee the long-term sustainability of coastal and off-shore projects, structures should be designed to withstand the most extreme wave conditions during their life time. This study carried out an Extreme Value Analyses (EVA) to determine the highest Significant Wave Height H_s that can hit the Maltese islands with a return period of N years, N being 50, 100 or more years depending on the end application. For this purpose, various distributions and extreme parameter analysis were used. Results were compared to determine the most accurate estimate of the *100-year* return period significant wave height, Hs_{100}. The H_{s100} was ultimately calculated to be 8.8 m. This value compares very favourably to other estimates obtained for other areas in the Mediterranean Sea. Besides the H_{s100}, two methods were used to determine also the highest individual wave height h that has an annual exceedance probability of 0.01. The two methods gave different values, 15 m and 13 m, with the former estimate considered to be more reliable.

Through the use of wave data, this project addressed the important issue of safety at sea. This project also carried out an assessment of the extreme wave conditions. By means of forecasts in the form of 2-D plots of significant wave height, wave length, wave steepness and individual wave height, together with the EVA, this study opens the way to new products that can be offered operationally at the service of mariners and structural engineers.

Acknowledgements

A massive thank you to Professor Aldo Drago for introducing me to the beauty of oceanography, and by virtue of his contagious enthusiasm and dedication showed me the importance of data and modelling. Thank you for recommending this project which I found tremendously interesting and for your constant support and direction throughout. Also, I am grateful to Dr. Anthony Galea for his constant advice and methodical approach to guide me through to the end.

But a really heartfelt thankyou goes to Dr. Adam Gauci. His constant availability, advise and knowledge especially on data and programming all delivered with his characteristic humility and charisma has saved me and my computer many times throughout the past year. Thank you. Really one of a kind.

My dear friend Mark Gaglione. I cannot be grateful enough for all you have done for me over the many years we've known each other, always there with your smile and wise advice. One of a kind. So proud to call you my friend.

My family. My pillar of support, my source of inspiration, I owe you all. I cannot thank you enough for being there when I hit the low points and in my moments of joy. This is dedicated to you. Love always.

Contents:

Index of Tables:

Index of Figures:

Chapter 1 : Wave Zonation for the Maltese Islands

1.1 Introduction

Throughout history, seas and coastal areas have been pivotal in the development and progress of civilisations, providing a medium of transport, a means of recreation and a source of subsistence amongst others. With a growing global emphasis on the blue and green economy, this dependence on our seas will continue to increase in the years to come.

A better understanding of waves and how they affect our daily activities is critical to develop a sustainable maritime sector. This emphasizes the importance of data collection and modelling.

A study entitled "Wave Zonation around the Maltese Island", is very broad and wide ranging, with many potential areas of study and most importantly application. It was decided to focus this study on two main areas which are very critical:

 1. Safety at Sea - To improve safety at sea of vessels travelling through the Maltese territorial waters. This can be achieved by providing detailed wave forecasts based on a broader range of wave characteristics to the vessel masters such that informed decisions can be taken well ahead of the critical event itself.

 2. Climatology – In particular to carry out an Extreme Value Analyses (EVA) for Significant Wave Height H_s for the Maltese Islands. This EVA will produce estimates of the likelihood extreme waves will occur in the Maltese islands over the next N years, N being the next 50, 100 or 1000 years, to permit the safe and long-term sustainability of coastal and off-shore projects.

1.2 Research Problem and Motivation of the Study

1.2.1 Safety at Sea

The sea provides an efficient means of transport and source of subsistence and recreation. Transport by sea is a very important economic activity accounting for about 95% of international freight transport as compared to 0.5% by air (1). The increase in globalisation and the emergence of new economies is putting greater demand on sea transport increasing the number and capacity of vessels and the range of goods transported.

The Mediterranean Sea via the Suez Canal acts as a link connecting Europe to Asia, and the strait between Malta and Sicily is particularly busy due to the major shipping hubs in Malta, Barcelona and Hamburg (Figure 1).
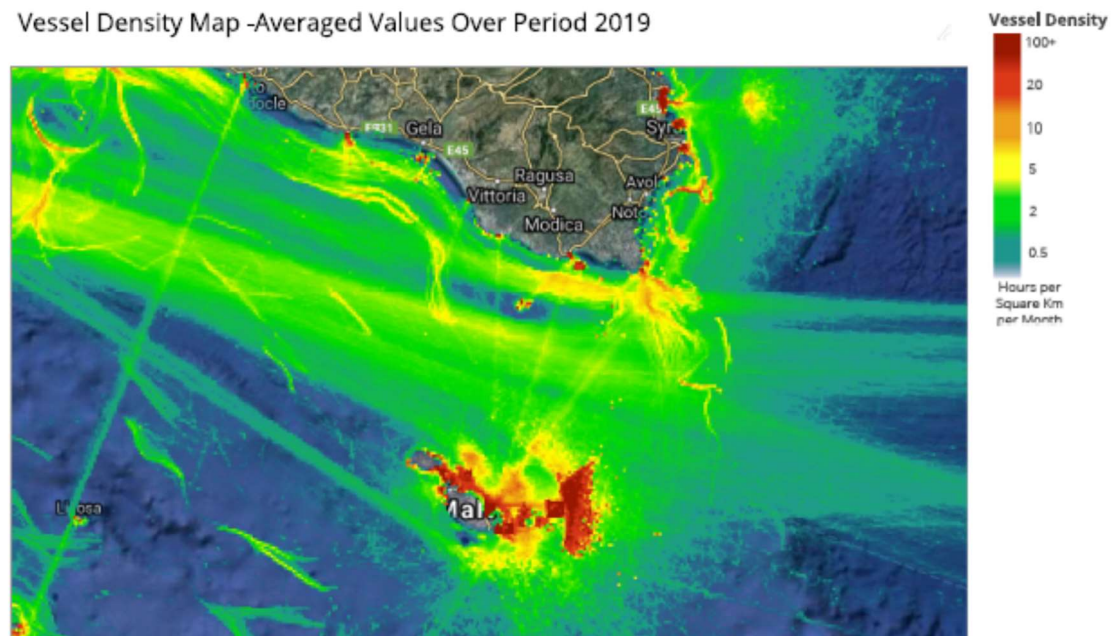
Figure 1 : Vessel Density Map around the Maltese Island in 2019; Source – EmodNET portal

Yachting, sailing and use of vessels as a means of recreation is also growing in importance. A closer look at the ports and harbours around the Maltese islands is indicative of how much this sector has grown in recent years. Fishing is also another aspect both in terms of recreation and as an economic activity.

Accidents at sea do happen. Reasons for this could be many ranging from human errors to poor maintenance of the vessels to crew experience. It is estimated that from 1980 to the year 2000, on average 0.30% of all commercial sea vessels were lost at sea each year (1). Also, over 1,000 lives are lost at sea each year. Weather conditions have a huge impact in this respect Approximately 30% of these ship losses were due to bad weather whilst another 25% due to 'other' unexplained causes (1).

Over a 10-year period between 2001 and 2010, the International Maritime Organisation (IMO) registered 3,648 Marine Casualties and Incidents Reports. *Zhang et al* (2) looked at each of these incidents to try and identify the primary causes. Out of these 3,648 Marine Casualties and Incidents Reports, 1,561 had enough information to be further investigated. Out of these 1,561 incidents, 755 (nearly 50%) accidents were weather related (2) .

Toffoli et. al. (3) carried out a similar exercise using the ship accident database within the Lloyds Marine Information Service. Between the period of January 1995 to April 1999, 650 vessels of gross tonnage above 100 tons reported incidents related to bad weather, the majority of which (66%) occurred in less than extreme wave conditions, with most incidents occurring at wave heights of less than 4 metres (3). Whilst it should be noted that extreme weather conditions would not be the only cause of the incident, it could be perceived as the instigator for such incidents to happen.

Unfortunately, a large number of incidents or near misses are not reported and these figures do not include incidents involving smaller vessels such as fishing and recreational vessels which are in larger quantities and more susceptible to weather induced incidents at sea.

Recent examples of weather induced incidents around the Maltese islands include the Togo registered Hephaestus which ran aground in Qawra (Figure 2) on the 10th February 2018 (4) and the British registered motor tanker Lady Sandra which broke in half after experiencing rough weather on the 26th March 2020 (Figure 3 and 4). A more tragic incident occurred in May 2013 when five French tourists lost their lives onboard a small boat off Gozo (5).

Figure 2: The Vessel Hephaestus ran aground in Qawra – Malta; Source: Transport Malta

Figure 3 and Figure 4: British registered motor tanker Lady Sandra broke in half off the coast of Malta when experiencing severe weather conditions; Source: Transport Malta

Waves not only compromise the structural integrity of a vessel through the impacts but also affect the vessel stability. To reduce the possibility of weather induced incidents, weather

forecasts for mariners are produced regularly all around the world. Meteorologists generally use third generation wave models to produce accurate sea state forecasts. The most common wave models are the WAve Model (WAM), Wave Watch III (WW3) and *SWAN* models (Simulating WAves Nearshore). Due to the complex nature of waves, wave models are generally phase averaging models, producing averaged values rather than individual values. The most common wave parameters forecasted are:

1. Significant Wave Height;
2. Peak or Mean Wave Period;
3. Mean Wave Direction;

Significant Wave height H_s, is defined as the average height of the highest one third of the waves in a period of time and it gives a practical indication of the sea state. Areas having high H_s values indicate areas of high risk.

It is interesting to note that IMO advises the ship master to constantly measure and monitor wave parameters such as wave length and wave period amongst others, during navigation (6). But while the physical monitoring of wave conditions by the vessel master is important, it could also mean that the vessel is already caught within an undesirable event before corrective action can be taken. There is no doubt that the best way to deal with extreme events is avoid them in the first place and that is why forecasts are so important.

In December 2000, the European research programme set up a 4.65 million Euro project called 'MaxWave' to provide timely and accurate forecasts of dangerous sea conditions. Whilst highlighting the fact that very large individual waves do occur, so far, no standardised warning system has yet been set up and forecasts are still based on the traditional values.

1.2.2 Wave Climatology around the Maltese islands

Wave climatology involves the analyses of data collected over a large number of years to produce statistics of the wave parameters for a specific area. An important aspect of wave climatology is the Extreme Value Analyses, that is, to use past data to predict the most extreme wave conditions in the coming years. These analyses are fundamental for the sustainability of any long term coastal or offshore projects. This information is also important to help identify areas at risk of flooding and also erosion.

A closer look at the Maltese coastline (Figure 5), indicates extensive development along the easterly coast which is predominantly at sea level, as compared to the westerly coast

characterised by high cliffs and higher altitude. These low-lying areas are prone to the effects of extreme waves.

Figure 5: Areas of development around the Maltese islands; Source: Sentinel June 2019

A recent example of the importance of such studies occurred on the 24th February 2019 when Malta experienced gale force winds of up to Force 11 from the Easterly direction. The damage caused was extensive running in the millions of Euros. Structures in the vicinity of the coast were extensively damaged (Figure 6) including also offshore structures such as fish farms.

Figure 6: Coastal structures damaged by the storm on the 24th February 2019. Source: Times of Malta

Extreme Value Analyses are used worldwide to help determine the *N*-year return period. The *N*-year return period is a statistical analysis of past data that gives an estimate of the extreme events based on the probability (*1/N* %) that they occur in any one year or alternatively once every *N*-years. Theoretically, the longer the past data set, the more accurate the prediction. *N* can be any value depending on the application of the data. In certain instances, an *N* value of 1,000 years are used to determine areas that will be affected by floods (7). Extreme Wave Analyses grew in importance in the 1990's and to this day it is still being studied to identify methods, distributions and procedures that produce the most accurate estimates.

Extreme Wave Analyses can be carried out on other parameters depending on the application of the data but Significant Wave Height H_s is generally used. Other parameters could include, Individual Wave Height *h* or Wave Crest elevation.

1.3 Aims and Objectives of the Study

As described above the project will focus on:

1. **Better safety at Sea in particular around the Maltese Islands**

Vessels have different basic functions, length and gross tonnage and also have different areas of operation. The IMO and the local maritime authority (Transport Malta) limit the range of operation of certain categories of vessels depending on the wave parameters and periods of the year. To this effect, in 2003 Transport Malta commissioned a report entitled "Malta Significant Wave Height Study" (8) to meet the EU directive 98/18/EC.

Yet despite of these regulations and despite the availability of weather and wave forecasts, accidents at sea still do happen. As previously discussed, the majority (66%) of weather induced incidents occurred in less than extreme wave conditions (3). This could mean that there are other factors including other wave parameters affecting the vessel stability which are currently not being closely monitored. Subsequently, this could mean that the forecasts being provided to mariners, based on traditional values of H_s and T_z are not sufficient.

This project aims to better understand the wave parameters affecting vessel stability and to provide a more detailed weather forecasting system to include these parameters. It should be noted that these forecasts are to be used as a reference and the final decision lies with the ship master.

2. Wave Climatology

While the blue and green economic sectors continue to increase, the long-term sustainability of these projects is based on the understanding of the environment they will be operating in. This project aims to carry out the wave climatology for the Maltese islands in particular to carry out an Extreme Value Analyses and generate the 100-year return value of Significant Wave Height and Individual Wave Height.

1.4 Study Area

The study focused on the sea around the Maltese Islands covering an area within the following coordinates:

Minimum Longitude :	$13.750°$ E	Maximum Longitude :	$15.450°$ E
Minimum Latitude :	$35.376°$ N	Maximum Latitude :	$37.124°$ N

that is from the southern coast of Sicily to a latitude south of the Maltese islands and having a resolution of $0.002° \approx 200$m. The study area along with the respective bathymetry is shown in figure (7) below.
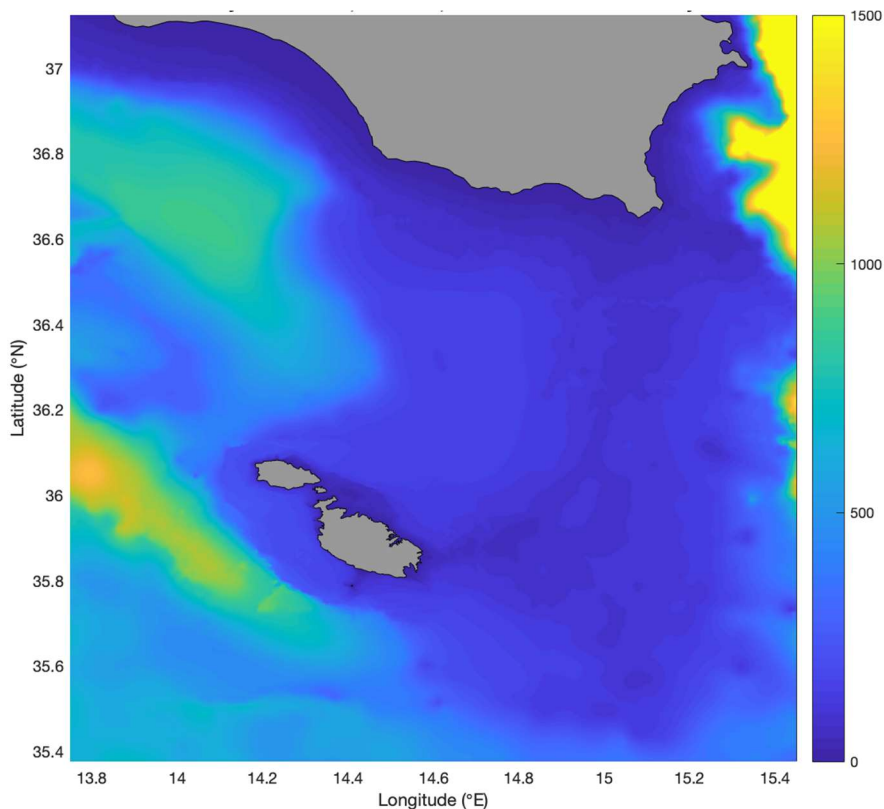
Figure 7: The Bathymetry of the study area; Reference: Physical Oceanography Research Group

1.5 Review of Dissertation

The second chapter will focus on the literature review covering aspects related to the stability of vessels. It will initially focus on reviewing the current legislations governing the areas of operation of specific vessels based on wave parameters. It will subsequently continue by identifying other wave parameters besides H_s affecting the stability of vessels at sea. This is an important aspect of the study because it will determine if the data currently being provided to mariners in the weather reports is comprehensive enough or not. Subsequently, the literature review covers the Linear Wave Theory to determine the mathematical formulae to help quantify these parameters using the data generated by the *SWAN* Wave model. These values will then be used to generate the required plots for the area around the Maltese islands having temporal and regional resolution similar to the data generated by the *SWAN* Wave model.

The third chapter will focus on the literature review covering the wave climatology and in particular the various methods and studies currently being used to generate the Extreme Value Analyses (EVA) in chapter 4. This will be important in identifying the methodology to be employed when conducting the study. To carry out an EVA, data covering a number of past years is required. This is required to analyse trends, fit mathematical distributions and hence produce the forecast. Data covering a minimum period of 5 years is required. Since data covering this period is not readily available, the *SWAN* Wave model was used to generate the required hindcasts.

In this project, data generated by the *SWAN* model was used extensively and so this data needs to be validated against actual in-situ data. The literature review will also cover this aspect.

Chapter 4 will cover the methodology that was employed to reach the objectives set out in this project, whilst Chapter 5 and Chapter 6 will present the results attained.

Chapter 2 : Literature Review - Vessel Stability

2.1 Introduction

The first part of this chapter will focus on understanding the various legislations and vessel classifications limiting the range of operation of vessels based on wave parameters. A more detailed explanation of the vessel classes together with the relative local codes of practice and EU legislations applicable can be found in Appendix A.

It will be followed up by an analysis determining which other wave parameters besides Significant Wave Height H_s affect vessel stability.

2.1.2 Terminology related to Waves

Figure (8) represents a perfectly sinusoidal wave, highlighting the most commonly used wave parameters and their corresponding terminology (9) (10):

Wavelength L, is the horizontal distance (in metres) between two successive troughs or two successive crests.

Wave period T, is the time taken (in seconds) for the successive passage of crests (or troughs) past a fixed point.

The frequency f, is the number of crests (or troughs) which pass a fixed point in 1 second (Hertz)

The wave height H, is the difference in height between the successive wave crest and wave trough.

The amplitude a, is the maximum height of the wave from the mean sea-level and is taken as half the wave height H.

The depth d, is the depth of the sea bed from the mean Sea Water Level (SWL).

The value η represents the vertical height of the wave at a particular location x and time t measured from the SWL.

Significant Wave Height H_s is defined as the average of the highest one third of the waves.
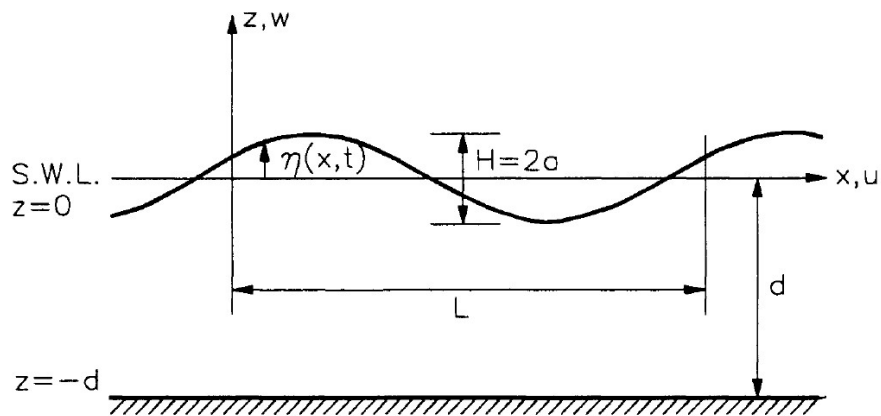
Figure 8: A perfectly Sinusoidal wave. Reference (10)

2.2 Codes of Practice and E.U. Directives

The two E.U. directives which base a vessel's area of operation on wave conditions are:

   1. "Recreational Craft Directive" - E.U. Council Directive 2013/53/EU (11);

   2. "Safety Rules and Standards for Passenger Ships Directive"- E.U. Council Directive 2009/45/EC (12);

The Recreational Craft Directive is adopted in the local codes of practice through the Non-Convention Vessels Code (13).

2.2.1 E.U. Council Directive 2013/53/ EU – Recreational Craft

The E.U. council directive 2013/53/ EU covers the essential requirements for recreational craft of minimum length 2.5 m and maximum length of 24 m (11). This directive excludes vessels with a crew or vessels designed to carry passengers for commercial purposes. Recreational boat design categories need to meet both wind force and significant wave height values. The recreational boat design categories are defined in Table (1).

Recreational Boat Design Category	*Wind Force / (Beaufort Scale)*	*Sig. Wave Height / m*
A (Boat for Ocean use)	Exceeding 8	Exceeding 4
B (Boat for Offshore use)	Up to, and including, 8	Up to, and including, 4
C (Boat for Inshore use)	Up to, and including, 6	Up to, and including, 2
D (Bor for use in Sheltered Waters)	Up to, and including, 4	Up to, and including, 0.5

Table 1: Categorisation of Recreational Boat Design as per EU Council Directive 2013/53/EU

The boat design category is marked on the builders' plate which needs to be permanently affixed to the vessel.

2.2.2 E.U. Council Directive 2009/45/EC – Passenger Ships

The E.U. Directive 98/18/EC, subsequently updated to Directive 2009/45/EC, establishes safety rules and standards for passenger ships constructed in steel or an equivalent material, operating in domestic voyages. A 'passenger ship' means a ship which carries more than 12 people (12). This directive applies to:

1. New passenger ships, all lengths;
2. Existing passenger ships of 24 m and above constructed in steel or equivalent material;
3. High speed passenger craft

regardless of their flag when engaged on domestic voyages. This Directive does not apply to (12)

1. ships not propelled by mechanical means,
2. vessels not constructed in steel or equivalent,
3. pleasure yachts unless they are crewed and carrying more than 12 passengers for commercial purposes,
4. ships exclusively engaged in port areas amongst others

Passenger ships are divided in classes depending on the conditions of the sea area in which they could operate, distance from shore and other criteria such as 'where ship wrecked persons can land' and 'distance to the place of refuge'.

2.2.3 E.U. Council Directive 2017/2108 – Passenger Ships

The E.U. Council Directive 2017/2108 is an update on the E.U. Directive 2009/45/EC. The Directive 2009/45/EC (14) proved difficult to adapt to new passenger vessels below 24 m in length and have been excluded from this updated Directive. Instead, vessels below 24 m length should be subjected to specific safety standards determined by Member States. Transport Malta adopted the Non-Convention Code and the Commercial Vessel Code as explained in Appendix A. Also, the criteria 'where ship wrecked persons can land' and 'distance to the place of refuge' have been removed.

2.2.3.1 Categorisation of sea areas and classes of passenger ships

Sea areas are divided into the following categories (13):

"Area A" means a sea area outside of areas B, C and D.

"Area B" means a sea area, whose geographical coordinates are at no point more than 20 miles from the line of coast, corresponding to the medium tide height, but which is outside of areas C and D.

"Area C" means a sea area, whose geographical coordinates are *at any point no more than 5 miles* from the line of coast, corresponding to the medium tide height, but outside of sea area D if any. Additionally, the probability of the *significant wave height exceeding 2.5 m shall be smaller than 10 %* for a period of one year for all-year-round operation, or for a specific period for seasonal operation, such as summer period operation (12).

"Area D" means a sea area, whose geographical coordinates are *at any point no more than 3 miles* from the line of coast, corresponding to the medium tide height. Additionally, the probability of the *significant wave height exceeding 1.5 m shall be smaller than 10 %* for a period of one year for all-year-round operation, or for a specific period for seasonal operation, such as summer period operation (12).

Passenger ships are divided into classes according to the sea area in which they may operate. The vessel classes and areas of navigation can be found in Table (2).

Vessel Class	Areas of Navigation
"Class A"	domestic voyages in Areas A, B, C and D.
"Class B"	domestic voyages in Areas B, C and D.
"Class C"	domestic voyages in Areas C and D
"Class D"	domestic voyages in Areas D

Table 2: Classes of Passenger Vessels and Areas of Navigation; Reference : (12)

To this effect, in 2003 Transport Malta commissioned a study on the Significant Wave Height around the Maltese Island (8) as set out in the E.U. Directive 98/18/EC. The study produced contoured 10 percent exceedance plots of significant wave height for all year period (Figure 8) and for the summer period, April to October both months included (Figure 9).

The Offshore Wave Data was obtained from the U.K. Meteorological Office (UKMO) wave model for a 14-year period extending from 1998 – 2002. The Weather Data was collected from the Malta International Airport and wind speeds from the UKMO global wave model. The

Mike Zero Nearshore Spectral Wave model package with a 100m x 100m grid size was used nearshore (8).

The significant wave height that is exceeded for less than 10% of the time was found to be around 2.1 m annually and 1.75 m for the summer period of operation.
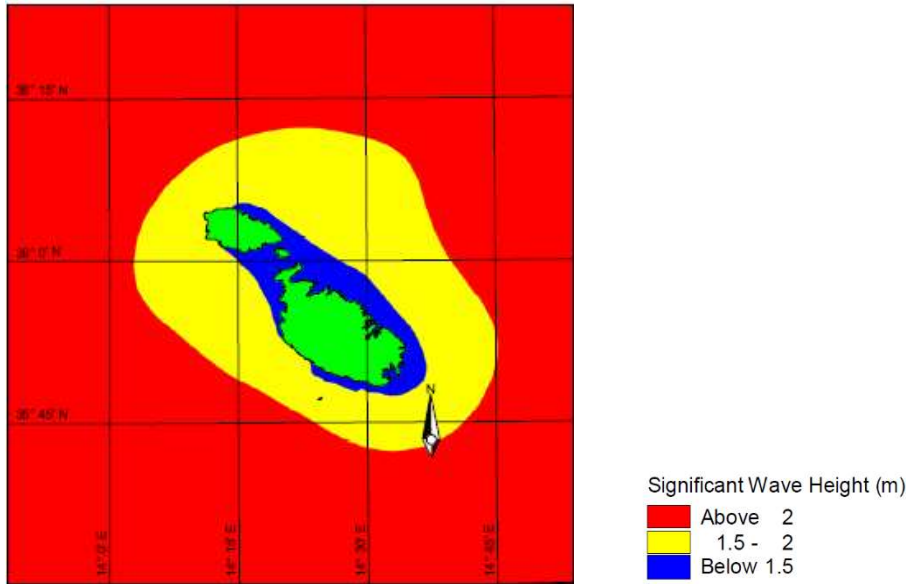
Figure 9: Zones for which specific values of Hs are exceeded 10% of the year ; Reference (8)
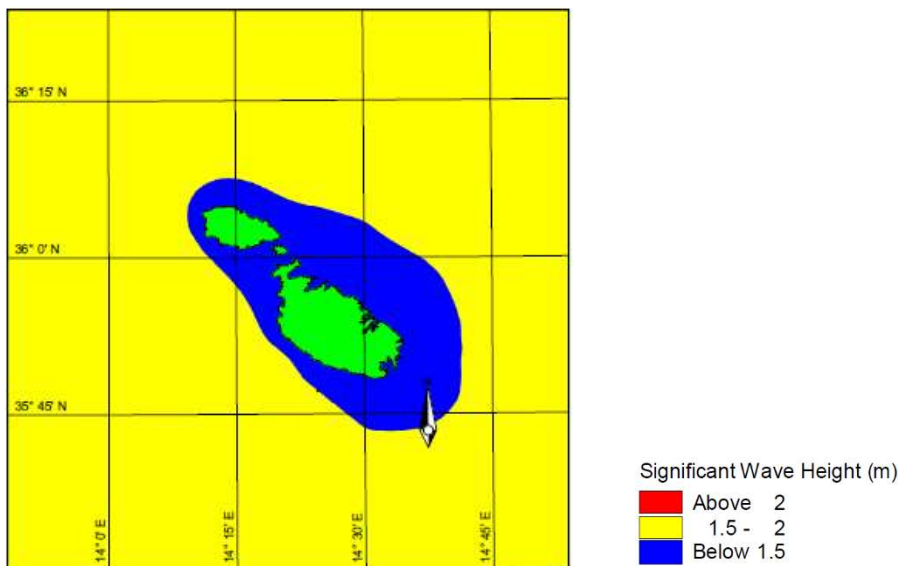
Figure 10:Zones for which specific values of Hs are exceeded in the summer period ; Reference (8)

2.3 Wave parameters that affect vessel stability

Despite all, wave and weather induced accidents still do occur and, in most cases, these occur in less than severe sea conditions, indicating that other parameters, wave related or otherwise, are also responsible.

2.3.1 Operational Guidelines

The International Maritime Organization (IMO) has set up operational guidelines for vessels facing adverse weather conditions (6). These guidelines are applicable for all types of vessels and are designed to be generic. Hence, these could be too conservative for certain categories of vessels and too favourable for others.

It is impossible to provide detailed guidelines for all vessels since the factors to be considered are too many including factors that are tangible such as hull design, vessel loading etc. but also intangible such as crew experience and maintenance of the vessel itself.

The final operational decision rests on the vessel master. However the IMO recommends the vessel master periodically monitors and takes note of the following parameters (6):

1. Natural Periodic Time of the ship at each condition of loading;
2. Wave Length - generally measured in comparison to ship length;
3. Wave Periodic Time, T_w - measured by a stop watch;
4. Significant Wave Height, H_s

It is interesting to note that *Toffoli et. al*. (3) , found that 66% of incidents involving merchant ships of size 100 tons and over around the world which were reported as being due to bad weather occurred at rather low significant wave heights, less than 4 m.

Figure (11) represent the terminology used to describe the direction of waves relative to the vessel whilst Figure (12) represents the terminology used to describe the 6-degrees of motion of the vessel.
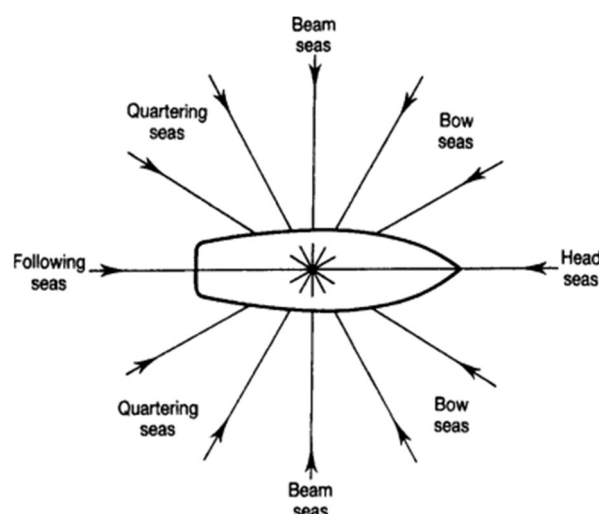
Figure 11: Terminology related to the direction of waves relative to the vessel
Reference: Spectral Analysis in Engineering, Concepts and Cases, Grant E. Hearn, 1995
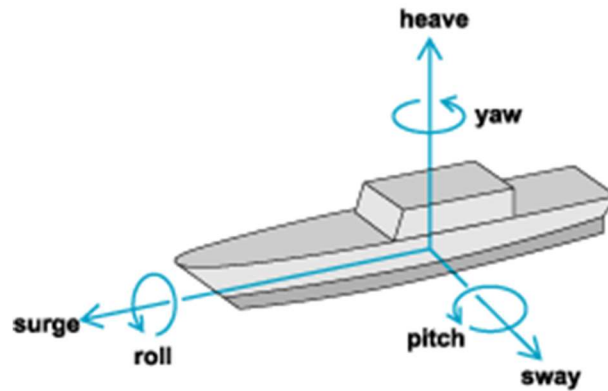
Figure 12: The 6- Degrees of freedom of a vessel
Reference: Evaluation of Human Discomfort Caused by Ship Vibrations, Constantin Picu,2021

The IMO has listed a number of dangerous phenomena when navigating through "following" and "quartering" rough seas (6). But vessels also encounter critical situations when facing "head" and "bow" seas (15). The combined list of dangerous phenomena includes:

1. Surf riding and Broaching;
2. Reduction of Vessel Stability;
3. Parametric Rolling;
4. Extreme Pitching;
5. Encounter wave grouping phenomena;
6. Extreme Individual Wave;
7. Breaking Waves.

2.3.2 Surfriding and Broaching

Surfriding and broaching are two interrelated phenomena that occur primarily in following seas with waves having a *wavelength comparable to ship length*. During broaching, the vessel can undergo a sudden, severe and uncontrolled change in direction leading to a considerable increase in tilting angle that can cause severe damage to the vessel. In certain instances, one wave would suffice to cause the vessel to capsize.

Broaching depends on factors such as the vessel position in the wave (crest or trough) , the ratio wave length / vessel length and Froude's number (16).

2.3.2.1 Types of Broaching

There are three different types of Broaching (16) :

1. Bow rooting – occurs when the vessel is travelling faster than the wave. The bow of the vessel will hit the back of the next crest, causing a sudden increase in resistance at the front. This impact will cause a loss in directional stability, which is termed "bow rooting".

2. Cumulative Yawing Motion – occurs when the ship speed is less than the wave speed. Each time a wave passes a vessel, the vessel is subjected to alternating forces; a force forward on the front part of the wave and a decelerating force on the back part of the wave. Unless the vessel is perfectly at right angles to the wave, these forces produce moments which will try to turn the vessel sideways.

This turning moment can only be counteracted by the vessels' rudder which in turn is dependent on factors such as the shape, size and amount of rudder immersed in the water at the time and the maximum rotation angle permitted.

If the moment applied by the waves is greater than that applied by the rudder, the ship will turn, increasing further the wave induced moment and causing it to turn even further. This will continue until the vessel turns to 90 degree to the wave direction exposing the vessel to beam seas leading to possible capsizing (15) . This is termed "cumulative yawing motion".

3. True Broach - As the relative velocity between wave and vessel gets smaller, vessels are exposed for longer periods to the accelerating force. As the gap continues to get smaller, this alternating force, will become continuous pushing the vessels to reach the wave speed, a phenomenon called "Surfriding". Again, the turning moment induced by this force is dependent on the direction of motion of the vessel relative to wave direction, with a zero moment when travelling at 90 degrees to the waves.

As will be discussed shortly, the stability of a vessel is at a minimum when the vessel is travelling on the crest of a wave. This phenomenon, combined with the high turning moment induced by broaching, can lead to vessels changing direction very suddenly. This in turn, will induce high centrifugal forces on the vessel causing it to tilt even further potentially causing the vessel to capsize. This is known as "true broach".

2.3.2.2 Factors affecting Broaching

Wave length, wave steepness and the relative velocity between vessel and waves play an important part in broaching, in particular the ratio wave length to vessel length. This ratio

together with wave steepness makes smaller vessels more prone to broaching than larger vessels. For smaller vessels, waves with smaller wave lengths and relatively small wave heights can be critical. The rate at which broaching occurs is inversely proportional to the square root of ship length, so it will occur at a much faster rate on smaller vessels (16). Studies on broaching have led to what is termed the broaching zone, that is the area where a vessel is prone to broaching (16).
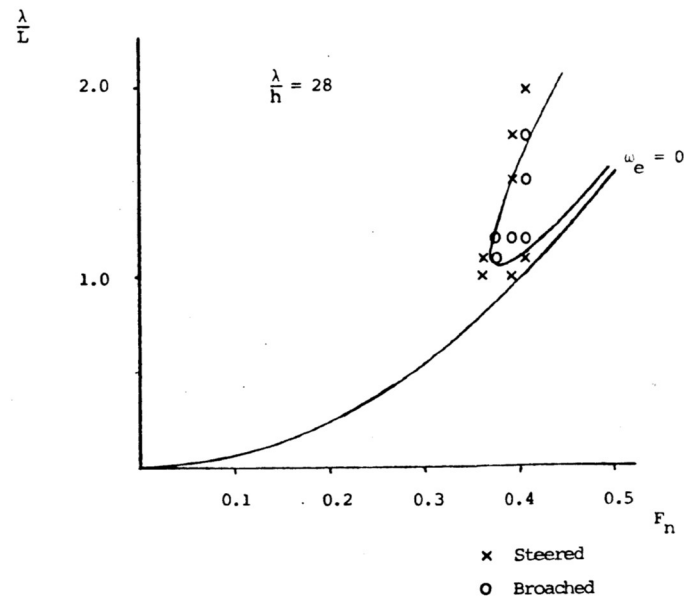
Figure 13: Effect of Froude's Number and ratio wave length / vessel length on Broaching; Ref. (16)

As can be noted in Figure (13) broaching occurred at ratios of Wave Length to Vessel Length greater or equal to 1 and Froude's Number greater than 0.30. The range of Froude's Number, F_n, was further elaborated based on vessel length as follows (17):

$F_n > 0.28$ for Vessel Length < 80m

$F_n > (0.0000181) \times Vessel\ Length + 0.282$ for Vessel Length > 80m

Experimental studies have found a relationship between the Froude's number and wave steepness that can induce broaching (17) :

$$Wave\ Steepness = 0.0310 \times F_n^{-3} + 0.282$$

The IMO (6) specifies that broaching occurs when the angle of encounter between the vessel and wave, α, is in the range (Figure 14):

Range of α : 135° < α < 225°

and ship speed in knots is higher than

$$Vessel\ Speed\ = 1.8\ \times\ (Vessel\ Length\ )^{0.5}\ /\ Cos(180 - \alpha)$$

The mergical zone is a zone in which a vessel experiences huge speed changes in following winds. The IMO defines this zone when both the angle of encounter and the vessel speed are in the following ranges (18):

$$Range\ of\ \alpha : \ 135° < \alpha\ < 225°$$

and speed in knots is higher than

$$Vessel\ Speed\ = 1.4\ \times\ (Vessel\ Length\ )^{0.5}\ /\ Cos(180 - \alpha)$$

To avoid both these phenomena, the vessel master is asked to adjust the vessel speed, angle of encounter 'α' or both to take the vessel outside of the mergical and critical zones (Figure 14)
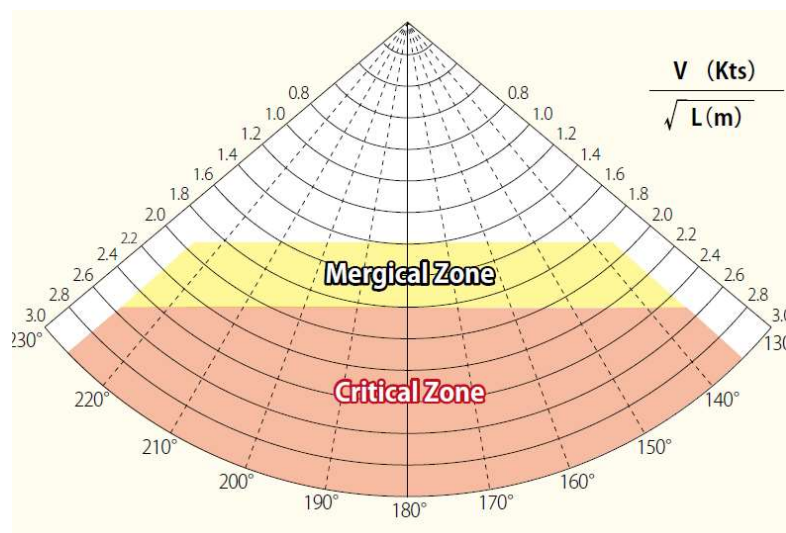
Figure 14: Definition of Mergical Zone and Critical Zone for Broaching ; Reference (15)

2.3.3 Reduction of Vessel Stability

Vessels experience a reduction in stability when passing over wave crests and this is particularly severe when the waves have *wavelengths of comparable length to the vessel length*. When passing over the wave crest, vessels experience a reduction in restoring moment leading to an increase in rolling angles. This reduction in restoring moment is mainly due to changes in the submerged form of the hull which effects the upthrust and the resistance to tilting of the vessel, leading to changes in position of the Metacentric Centre GM . The vessel would need larger tilting angles to attain the same levels of restoring moment (19).
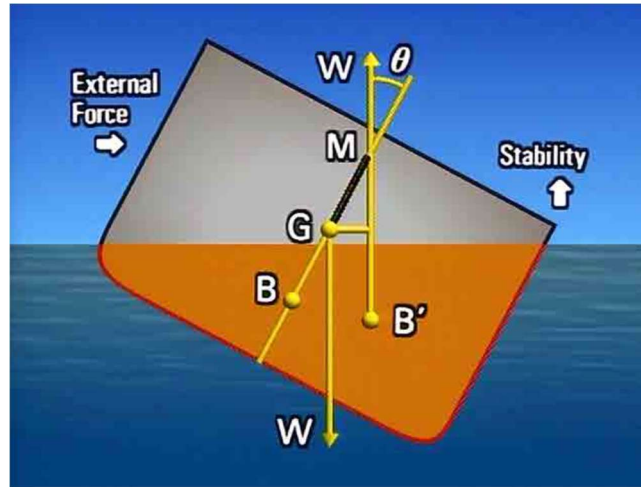
Figure 15: Forces acting on a vessel when tilting ; Reference (15)

The position and the forces acting on a vessel can be seen in figure (15) above ,

where G : centre of Gravity

B and B': centre of Buoyancy when vessel is flat and inclined at an angle 'Ø';

M : point of intersection between the centreline and the vertical line through B;

GM : governs ships stability and dynamic characteristics;

Stability of a vessel is determined by its righting moment, that is on the upthrust generated on the vessel mutiplied by $GM \sin Ø$ (also refered to as GZ) with Ø being the angle of inclination.

The general stabilty curve that needs to be met by new vessels as set up by the I.S. Code 2008 (19) is shown in figure (16). It clearly highlights how the vessel's righting moment changes with tilting angle. New vessels need to meet this stability curve in calm waters which unforunately does not reflect how the vessel behaves in rough seas in particular when the vessel is travelling over wave crests and troughs.
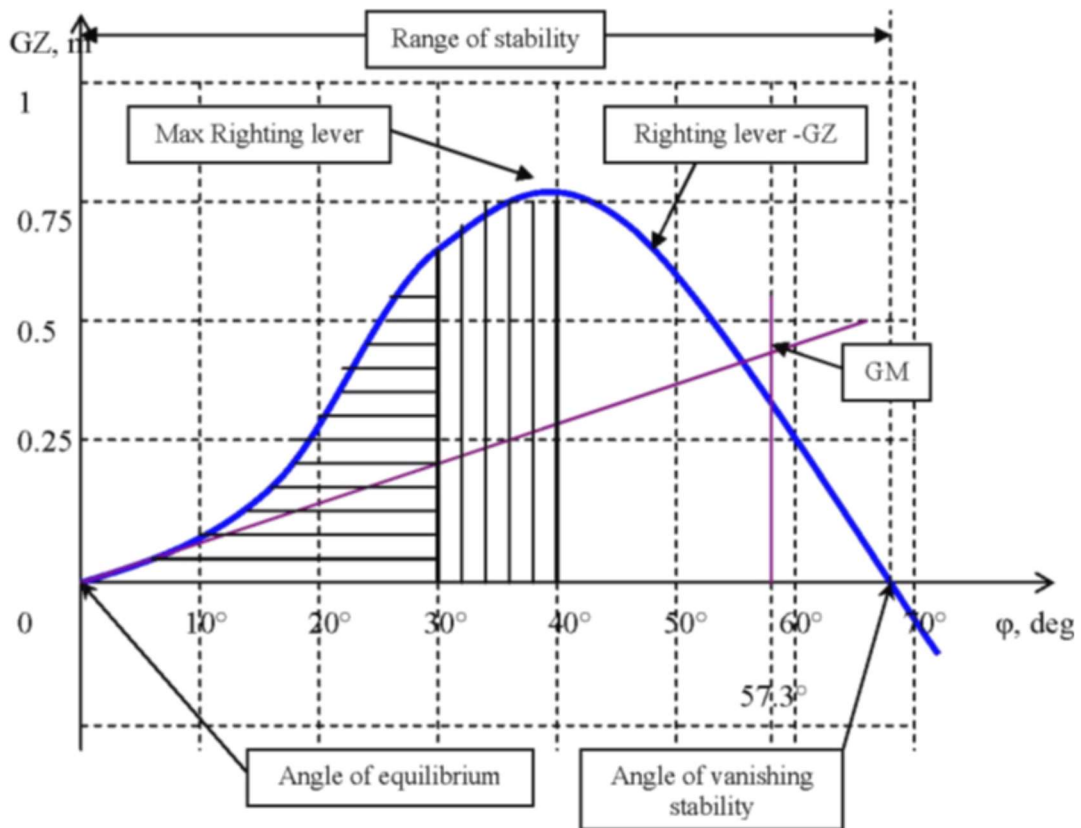
Figure 16: The general stabilty curve to be met by vessels as set by I.S. code 2008; Reference (19)

On the other hand, figure (17) shows how the stability curve shifts when the vessel is passing through the different phases of a wave as compared to calm water conditions. It is interesting to note that the stability curve can be negative when travelling on a crest and this can happen at relativly small tilting angle.
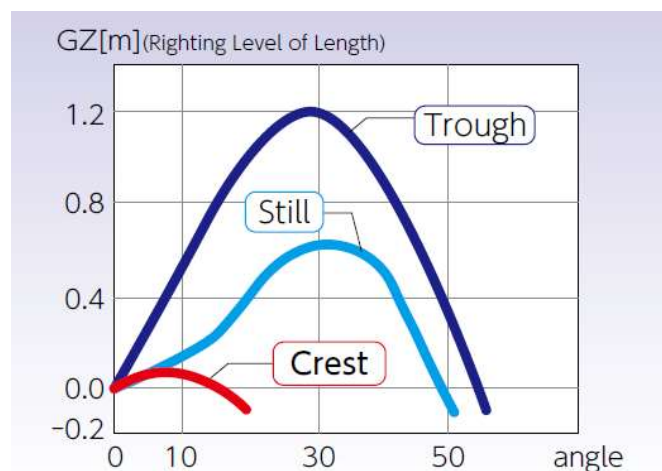
Figure 17: Variation in stability of a vessel travelling in different wave zones; Reference (15)

This reduction of righting moment leads to even larger angles of inclination when the vessel is passing over the crest. This situation is particularly dangerous in following and quartering seas, because of the increased time the vessel spends on the crest and hence with reduced stability.

Modern ships are designed to be more efficient when moving in water meaning more ships are being built with a narrow hull and large flare like container ships and fishing vessels. Whilst meeting the stability critera in calm water as discussed above, these are much less stable in rough water as compared to full hull ships such as tankers since the contact area is lower (19).

This stability reduction may become critical for wave lengths within the range of

$$0.6 \, VL < \text{Wave Length} < 2.3 \, VL$$

where *VL* is the ship's length in metres (6).

Within this range the amount of stability reduction is nearly proportional to the wave height (6). The steepness of the wave will also have an effect on the contact plane. The steeper the wave the less the submerged area of vessel (15) and hence the less the restoring moment for a particular angle of inclination.

2.3.4 Parametric Roll Motion *(Resonance)*

Resonance can cause the ship to roll to very large angles even in a moderate sea. In extreme cases this can cause the vessel to capsize. Parametric rolling occurs predominantly in following seas but can also occur in head seas (20).

Since the vessel is moving forward and generally at an angle to the waves , the actual wave period and wave length will be different from the values measured by a person on board. Time of Encounter (leading to Encounter Frequency) is the wave period as measured from on board. When the encounter frequency is close but not necessarily equal, to the natural roll frequency of a vessel, the roll angle of the vessel would gradually increase. After a number of successive waves this increase in roll angle can become critical. This event is called Parametric Rolling.

In following seas, the main causes of parametric rolling are the variation in stability of the vessel when passing through the various stages of the wave profile. In head seas, the frequent heavy heaving and/or pitching experienced when hitting the waves, may contribute to parametric roll motions, even with small waves (6).

Parametric rolling occurs when the *wave lengths are similar to or larger than vessel length* (19) and either

    a. Encounter Frequency ω_e is equal to the ship's Natural Rolling Frequency ω_o (6) . The stability is at a minimum only once per wave period, hence it will tilt more on one side than the other or

    b. Encounter frequency ω_e is equal to twice the ships' Natural Rolling Frequency ω_o (6)

$$\frac{2\,\omega_o}{\omega_e} = \pm n$$

where n = 1,2,3 etc.

It normally occurs when $n = 1$ and is called 2:1 resonance. The rolling motion will increase on each side. The IMO guideline 1228 provides a chart to help determine Time of Encounter, T_e, (6). Alternatively the Time of Encounter T_e can also be measured as the period of pitching by using stop watch or calculated by the formula:

$$T_e = \frac{3\,T_w^2}{(3\,T_w + V \cos \alpha)}$$

where V = Speed of the Vessel (knots);

       α = Angle between Vessel and Wave Direction ($\alpha = 0°$ means head sea);

       T_w = Wave Period;

2.3.5 Extreme Pitching

Pitching is the up and down movement of the front of the vessel relative to the back. The pitching motion is intensified when facing head seas. When the *wave length is almost equal to ship length*, the pitching motion increases significantly with large amplitude oscillations especially at the bow relative to the water level (15).

Greenwater is defined as unbroken waves coming over the bow, the side or the stern of a vessel (21). A number of incidents that occurred in the North Sea attributed to greenwater were studied. No relationship was found between the incidents and Significant Wave Height yet they occurred in wave conditions less severe than the vessel design maxima.

To understand better this phenomenon, a vessel was exposed to waves having different wave lengths but same wave heights (Figure 18 and 19). The bow's vertical motion peaked at wave lengths 20% longer than the vessel length, with values close to 2.75 times higher than those at lower wave lengths. Further studies of the incidents showed that a large number occurred when a wave length / ship length ratio of 1:1 existed (21) .

This confirms a previous study, stating that critical waves appeared when the wave lengths were of comparable size to the length of the ship (10-20% longer) (22).
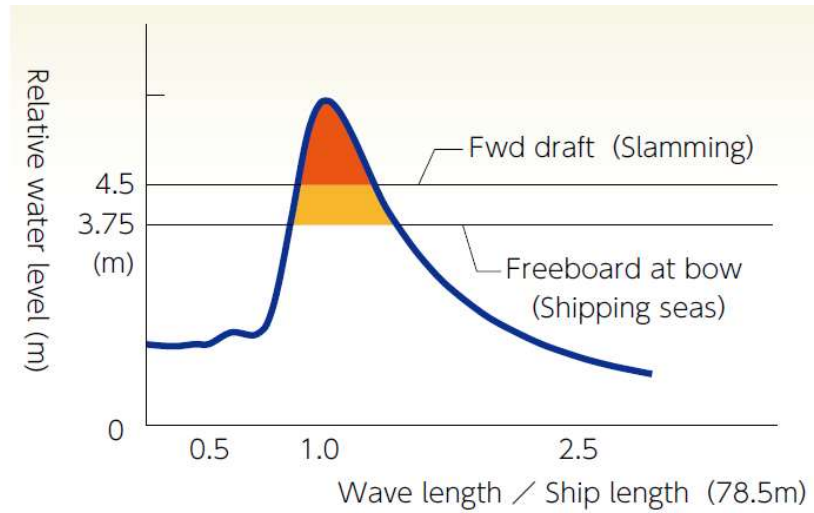
Figure 18: Pitching motion with Wave Length / Ship Length ratio ; Reference (15)
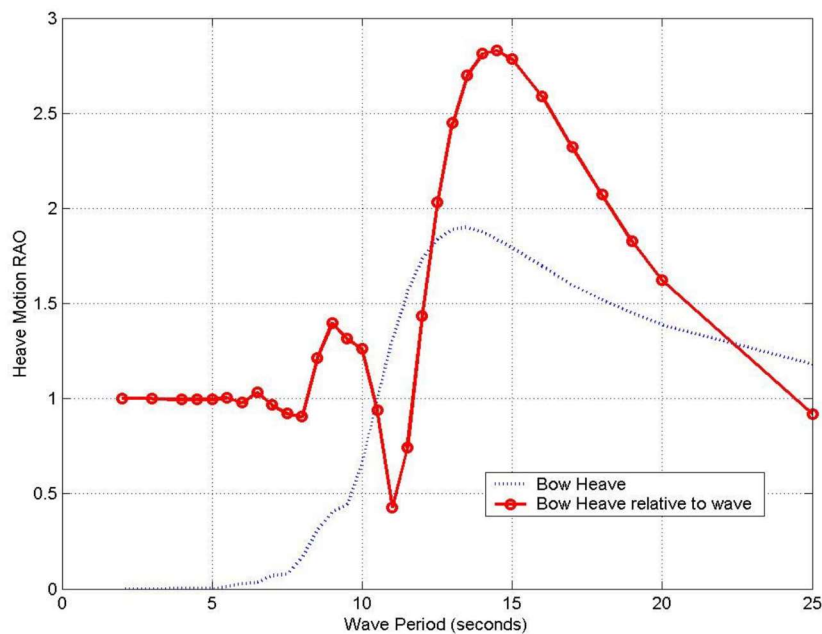
Figure 19: Variation in pitching motion with Wave Period; Reference (21)

2.3.6 Encounter Wave Grouping / Successive Wave Attack

This phenomenon is similar to extreme pitching but occurs when vessels encounter high waves which are faster than the speed of the ship itself. The vessel is now being hit repeatedly and severely from the back by a series of high waves, which can cause manoeuvrability issues.

High waves in this case are defined as waves having:

1. An average Wave Length larger than 0.8 Vessel Length and
2. Significant Wave Height H_s larger than 0.04 Vessel Length.

The IMO through Circular 1228 (6) specify the Dangerous Zone for successive high wave attack in following and quartering waves to be avoided by the ship master (Figure 20).
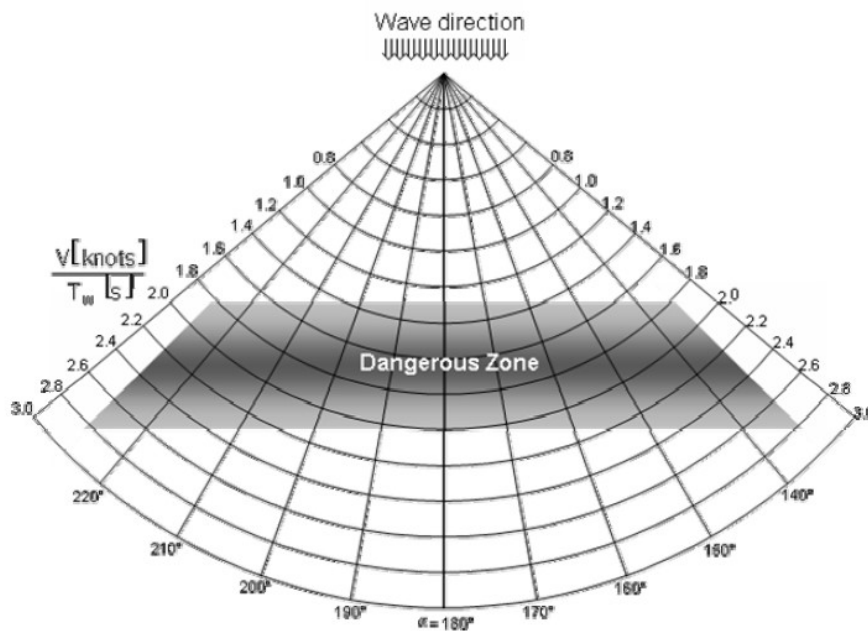
Figure 20: Dangerous zone of successive high wave attack in following and quartering seas; Reference (6)

2.3.7 Extreme Individual Wave

Significant Wave Height H_s is representative of a sea state and sea severity and is used in most weather forecasts for mariners. But H_s, is defined as the average height of the highest one third of the waves in a period of time. However average values are not representative of the extreme conditions. The most critical waves are those which are unexpected, that is , those single waves that are much higher than the normal wave or which come from an unexpected direction (23).

The Rayleigh distribution is used to calculate the probability distribution of the individual wave heights, *h,* and the probability individual wave heights exceed a given value (Probability or Percentage of Exceedance).

The probability individual waves reach heights greater than the significant wave height H_s, is around 0.135 or 13.5%. This coincides with the 1/7 rule observed by sailors.

2.3.8 Breaking Waves

It was observed that 50% of weather related incidents at sea occurred during sea states having total sea (wind sea and swell) steepness between 0.035 and 0.045 (3) . Wave steepness is defined as the ratio of wave height to wave length and has two effects on vessel stability. It will determine the period and the extent a vessel will experience a variation in stability when travelling over the crests and troughs. It is also an important factor in the formation of breaking waves.

Waves that become too steep will break and the impact of breaking waves on vessels are much larger than those of equivalent non-breaking waves (23). A vessel is particularly vulnerable to broad side breaking waves. Boats can capsize if impacted by a breaking-wave having a height equal to the breadth of the boat occurring at the broad side of a vessel (23). Hence smaller vessels such as fishing vessels, are much more exposed to such incidents.

Water particles at the crest of a wave move slightly in the direction of the wave, meaning the water particles at the top are faster. This difference in speed increases with higher waves. Waves in deep water move forward with a speed of L / T with L being the wave length and T the periodic time. Yet the forward speed of the crest can reach values of $7 H / T$ with H being the wave height. Should the wave height H exceed the value $L / 7$, the top part of the wave is moving faster than the actual wave and would break. According to Stokes Theory, waves cannot attain a height of more than $L / 7$ without breaking. The steepness of waves was found to rarely exceed 0.10 (9).

2.3.9 Co-occurring seas

Sea waves can be described as

1. Swell - when the waves are no longer under the influence of their generating winds. The profile is nearly regular and sinusoidal;

2. Wind sea – when the waves are still under the effects of the generating winds and are highly irregular and relatively short crested;

3. Co-occurring – when both wind sea and swell occur together.

Co-occurring seas are broadly defined as sea states having both swell and wind sea. Owing to the size of the Mediterranean basin, the swell element is usually small when compared to wind sea. Yet, the below information could still be relevant for other areas where swell has also an effect on the sea state.

2.3.9.1 Significant Wave Height, Co-occurring Seas

Almost half of the incidents reported as happening in co-occurring seas, had a combined value of Significant Wave Height, H_s (called Total Sea) between 0 and 3 m, with 24% of the incidents occurring with H_s between 2 and 3 m (2). These values are much less than what is normally considered extreme conditions.

2.3.9.2 Mean Wave Period, Co-occurring Seas

It was observed that nearly 66% of incidents in co-occurring seas had a difference between the mean period of swell and the mean period of wind sea (ΔT) less than 3 s, with nearly 30% occurring within ΔT of 2-3 s. Close mean periods between swell and wind sea in co-occurring seas (that is $\Delta T <3$ s) is an important factor leading to extreme events (2).

2.3.9.3 Mean Wave Direction, Crossing Seas

Crossing sea is defined as a sea state made up of more than one wave group travelling in different directions. When vessels are experiencing crossing seas, the vessel master would veer the ship in the direction perpendicular to the highest waves, leaving the vessel exposed to the impact of the lessor waves on the sides. Analyses of waves occurring in crossing sea states indicate that:

1. Crossing seas may induce high waves and generate dangerous sea conditions.
2. As the combined highest significant wave height is approached, the angle between the waves is reduced with the wave trains taking up similar directions

Zhang et. al. (2) found that 55 % of the incidents in crossing seas, happened when the difference between the mean wave direction of swell and wind sea, ΔD, was less than 30 degrees (2). When the values of ΔT (the difference between the mean periodic time of swell and wind sea) were analysed together with ΔD, the lowest value of ΔT (1.8 s) occurred when ΔD was between 30 and 40 degrees, indicating strong coupling between swell and wind sea waves.

This seems to confirm previous studies, that the rate of change in swell energy under the influence of wind sea energy reaches a maximum at approximately 40 degrees. This produces a steep swell, with a high wave energy and extreme wave height.

2.4 Weather forecasts

The weather forecast for mariners provided by most meteorological offices is based on the traditional data of Significant Wave Height, wave and wind direction and wave period of wind sea and swell, based on phase averaged wave models. Little or no information is provided on other wave parameters (Figure 21).

As noted above, there are other wave parameters that can cause high risk situations but as yet are not covered by most weather forecasts. A more detailed weather forecast could help to reduce the rate of accidents at sea.
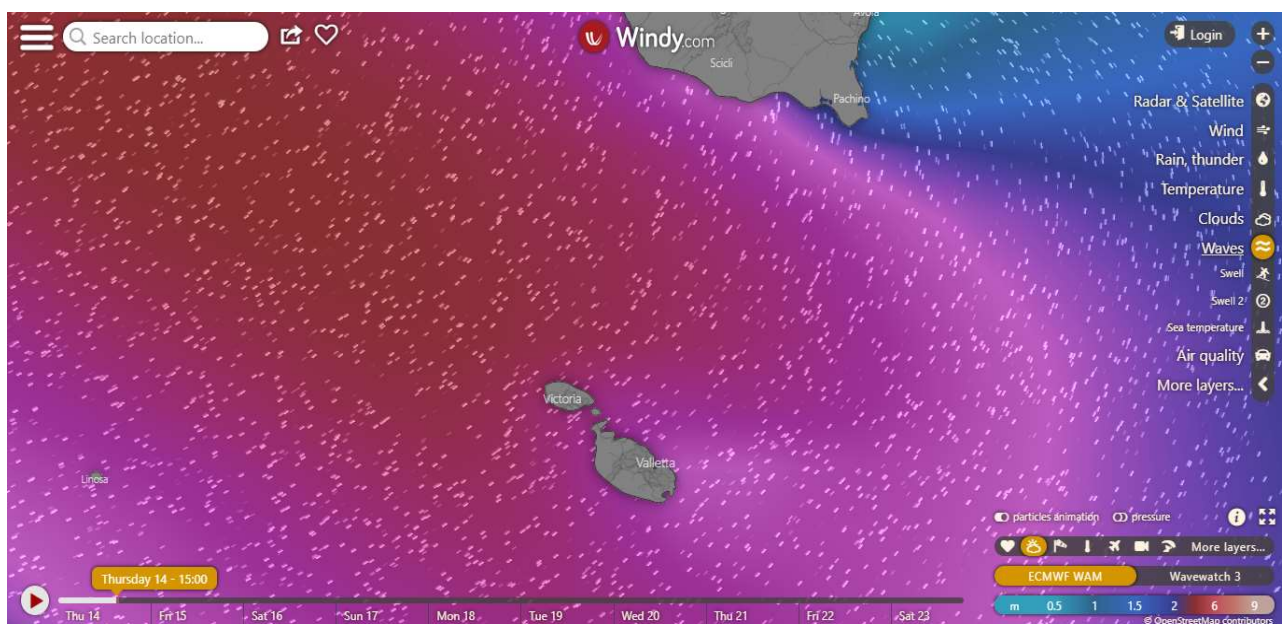
Figure 21: Typical on-line Weather forecasts; Source: windy.com

2.5 Linear Wave Theory

From the above literature review, the following wave parameters cause a reduction in stability of vessels:

1. Wave Length;
2. Single Extreme Wave Height;
3. Wave Steepness.

Using the forecasted data generated by the *SWAN* wave model and mathematical equations based on the Linear Wave theory, *2-D* plots of the above parameters can be produced. The

Linear Wave Theory is used to represent wave fields propagating *without any forcing* and is used extensively in wave studies. It is based on a number of assumptions (10), mainly:

1. Waves have a constant wave length and travel across seas of constant depth or areas where the depth varies slowly;

2. Wave Height is small compared to both wave length, *L,* and water depth, *d*. Hence the free surface can be taken as the still water level;

3. Wave motion is two dimensional and does not change with time;

4. Fluid is incompressible and effects of viscosity, turbulence and surface tension are neglected.

The wave steepness, very rarely exceeds the values 0.1 (9) and hence the small wave height in comparison to wave length assumption is often valid. Figure (7) shows the Bathymetry for the area being studied, which with the exception of the trench to the West of the Maltese islands, the changes in depth with distance are relatively gradual. These considerations justify the use of the Linear Wave Theory for this study.

Solving the conservation of momentum and the mass equation for the water particles at the boundary conditions, leads to the Airy Stokes equation:

$$\emptyset\,(x,z,t) = \frac{a\,g\,cosh\,[\,k(d+z)]\,.\,cos\,(\,kx - \omega\,t\,)]}{\omega\,cosh(kd)} \tag{1}$$

where $\emptyset\,(x,z,t)$ = velocity potential of water particles at coordinates *x* and *z* at time *t*

a	= wave amplitude = *h* /2 with *h* = wave height;
k	= wave number = $2\,\pi\,/\,L$;
L	= wave length;
ω	= wave angular frequency = $2\,\pi\,f = 2\,\pi\,/\,T$;
T	= wave periodic time;
d	= depth;

from which the basic properties of linear water waves can be derived mainly:

a. *Wave profile at the surface:*

$$\eta(x,t) = a\,sin(k\,x - \omega\,t) \tag{2}$$

where $\eta(x,t)$ = wave position at the free surface position *x* and time *t*.

b. *Dispersion equation* (10)

$$\omega^2 = \;g\,k\,tanh\,kd \tag{3}$$

indicating there is a unique relationship between *ω*, *k* and *d* (hence *T, L* and *d*)

c. Wave Length, L

$$L = \frac{g\,T^2}{2\,\pi}\ tanh\left(\frac{2\pi\,d}{L}\right)$$

$$L = \frac{g\,T^2}{2\pi}\,tanh\,(k\,d)\qquad\qquad(4)$$

for *deep water wave conditions*, as product $(k\,d)$ approaches infinity, $tanh(\,kd\,)$ approaches 1, hence wave length L is approximated by:

$$L = \left(g\,T^2 / 2\pi\right)\qquad\qquad(5)$$

for *shallow water wave conditions,* as product $(k\,d)$ approaches 0, $tanh(kd\,)$ approaches value $(k\,d)$, hence wavelength L is approximated by:

$$L = T\,\sqrt{g\,d}\qquad\qquad(6)$$

d. Wave Steepness

is defined as the ratio between wave height, h and wave length, L, hence for deep water using equation (5)

$$Wave\ Steepness = \frac{2\,\pi\,h}{g\,T^2}\qquad\qquad(7)$$

2.5.1 Classification of Waves

In waves, the water particles at the water surface move in vertical circles. The diameters of these circles decrease exponentially with depth from the surface. At a depth of about half a wave length, the diameter of the circular motion decreases to less than 4 per cent of those at the surface.

Hence, the effects of surface waves on water particles at depths greater than half the wave length are considered to be negligible. Similarly it is considered that the sea bed has a negligible effect on the surface waves at a depth greater than half the wave length (9). Yet, at depths less than half the wave length, the bottom starts to affect the waves at the surface and the wave characteristics start to change. Eventually it's only the wave period, T that remains the same, with values of both the wave speed and the wave length decreasing.

When the water depth becomes less than half the wavelength (ratio $d\,/L$ <0.5), there is an initial slight decrease in wave height. Yet when the water depth to wavelength $(d\,/\,L)$ ratio is about 0.06, the wave height increases back to the original value and continues to increase rapidly until the waves break. The water depth to wavelength ratio $(d\,/L)$ of about 0.06 seems to trigger

a change in properties indicating a change from what is termed intermediate to shallow water (9) . Hence the water depth, d to wave length ratio (d/L) acts as a means of classifying waves and provides valuable indications of the effects of the sea bed on the wave properties (Table 3)

Classification of waves	Ratio Depth d / Wavelength L
Deep water	$d > L/2$
Transitional depth	$L/20 < d < L/2$
Shallow water	$d < L/20$

Table 3: Classification of waves based on the Depth to Wavelength (D/L) Ratio

2.5.2 Spectral Representation of Waves

The complex wave profiles can be best represented through the use of a Spectral (or Fourier) model (10), were the water surface elevation is approximated by the summation of a high number of individual sinusoidal waves which satisfy the Linear Wave Theory and have profiles following equation (2). Hence, the surface elevation is given by

$$\eta(t) = \sum_{i=1}^{i=n} a_i \sin(\omega_i t + \emptyset_i) \qquad (8)$$

Where $\eta(t)$ = water surface elevation at time t;

      a_i = wave amplitude of wave component i;

      ω_i = wave angular frequency of wave component i;

      $\emptyset_i$ = phase of wave component i

The Fourier series model can be extended to include also the possibility of the individual waves coming from different directions leading to:

$$\eta(x, y, t) = \sum_{i=0}^{i=n} a_i \sin[k_i(x_i \cos \theta_i + y \sin \theta_i)] - \omega_i t + \emptyset_i) \qquad (9)$$

Where k_i = wave number;

      θ_i = angle between the x axis and direction of propagation of wave component i;

2.5.3 Finite Amplitude Wave Theory

Linear Wave Theory assumes wave heights to be small and so the free surface can be taken as the still water level ($z = 0$) rather than the disturbed wave surface ($z = \eta$).

The Finite Amplitude Wave Theory on the other hand takes into consideration the influence of the wave itself on its properties. Hence the free surface is now taken to be the amplitude of the wave. The Linear Wave Theory predicts the crest and the troughs are of same height equally

divided by the still water level. The Finite Amplitude Theory describes waves as having flat troughs and peaked crests, with the crests higher than the troughs above the still water line.

2.5.4 Statistical Description of Wave Records

The wave profile is extremely variable with rapid and continuous changes. The sea state is best described through the use of averages and statistical analyses of the measured instantaneous data. The two main parameters used to define a sea state are (9):

1. Significant Wave Height, H_s and
2. The Mean Zero up-crossing (or down-crossing) Wave Period, T_z

As discussed above, Significant Wave Height H_s is defined as the average wave height of the highest 1/3 of waves. The wave period T_z is the time period between two consecutive down-crossing or up-crossing of the wave from the free surface. There is no preference as to which to use whether up-crossing or down-crossing since generally, if the data is sufficiently long, the mean values of both would be very similar.

A sufficient number of individual measurements should be taken to allow representative mean values to be calculated. Generally, 200 waves are measured and the means calculated.

2.5.5 Rayleigh Distribution

The Probability Distribution Function of individual waves having wave heights h is thought to follow a Rayleigh Distribution (10). Hence, the probability a wave height is equal to value h, *P(h),* is given as follows:

$$P(h) = (h/ b^2)\; exp\,(- h^2 / 2\, b^2) \qquad\qquad (10)$$

Where b $\quad$ = scale parameter with $b^{\,2} = 4\,\sigma^2$

$\quad\;\; \sigma^2$ $\quad$ = the variance of wave function

Substituting for the scale parameter

$$P(h) = (h/4\,\sigma^2)\; exp\,(- h^2 /8\,\sigma^2)$$

From analyses carried out by Goda (10), the significant wave height H_s , is equal to *4 σ.* Substituting into above equation, the Probability Distribution Function, *P(h),* becomes:

$$P(h) = (4h/H_s^2)\; exp\,(- 2\, h^2 / Hs^2) \qquad\qquad (11)$$

The Probability of Exceedance $Q(h)$ is the probability wave heights are higher than a given wave height h. This is equal to the summation of $P(h)$ from wave height h to infinity (9).

$$Q(h) = \int_h^\infty P(h)\, dh$$

$$Q(h) = \int_h^\infty (4h/H_s^2) \exp(-2h^2/Hs^2)\, dh$$

$$Q(h) = exp\ (-2h^2/Hs^2) \qquad\qquad (12)$$

The Cumulative Distribution Function, CDF(h), on the other hand which is the summation of the probability distribution function from wave height 0 to wave height h is given by

$$CDF(h) = \int_0^h P(h)\, dh$$

$$CDF(h) = 1 - [exp\ (-2h^2/Hs^2)] \qquad\qquad (13)$$

Based on the above, if we have an actual time series with a known number of waves N and a known value of H_s, it is possible to determine the number of waves n that exceed wave height h as follows

$$n = Q(h) \times N$$

But the probability of exceedance $Q(h)$ can be found from equation (12)

$$h = H_s\ \sqrt[2]{0.5\ ln\left(\frac{N}{n}\right)}$$

Hence the highest wave h_{max} that occurs during a period with N waves, $n = 1$ and

$$h_{max} = H_s\ \sqrt{0.5\ (ln\ N)} \qquad\qquad (14)$$

Another parameter of interest is the highest wave that can occur during a 3-hour interval. Based on the above, and assuming H_s and T_z are the values attained over a 3-hour period

$$H_{max\ 3\ hour} = H_{s\ 3hr}\ \sqrt{0.5\ (ln\ N_{3hr})} \qquad\qquad (15)$$

Where $H_{s\ 3hr}$ = Average H_s over the three-hour period

$\quad\quad N_{3h}$ = number of waves in three hours, = (3 x 60 x 60 / $T_{z\ 3\ hr}$]

$\quad\quad Tz_{3hr}$ = up cross period over the 3 hours

In certain instances, such as in off-shore structures, the crest height of the waves is also relevant. Extreme waves tend to be asymmetrical and are not perfectly sinusoidal in shape. Typical values of crest heights are about 0.60 of the total wave heights.

Chapter 3 : Literature Review- Wave Climatology

A study of the Significant Wave Height both off-shore and for areas closer to the shore was commissioned in 2003 by the Maltese national maritime agency, Transport Malta (8). It was concluded that the off-shore Significant Wave Height that is exceeded for less than 10% of the time annually was 2.1 m and for the summer period 1.75 m, with the predominant wave and wave direction being North West. The off-shore wave directions were very similar to the wind direction, confirming that the wave climate around Malta is mainly based on sea waves rather than swell. Sea waves are characterised by being highly irregular, steep and having relatively short-wave periods when compared to swell.

A wave climate analysis generally includes the following (9):

    a) Extreme Wave Analyses with the long-term return (e.g.,100 years) values of Significant Wave Height;

    b) Percentage frequency of wave heights or wave periods;

    c) Joint distribution of Significant Wave Height H_s and Wave Period T_z;

    d) Time series plots of wave heights and wave periods;

    e) Relationships between Significant Wave Height H_s and Individual Wave Height h.

3.1 Extreme Value Analyses

Extreme weather leads to extreme wave conditions. These extreme conditions need to be taken into consideration during the planning and design stages of off-shore, near-shore and coastal structures to ensure their long-term sustainability. Predicting the extreme conditions including estimating the extreme wave values that occur once every *N years* (the *N-year* return value) are thus needed. This is particularly relevant for island states like Malta.

Extreme Value Analyses involves the statistical analyses of long-term data collected either from direct measurements and / or hindcasts. These statistical analyses are used to generate forecasts of the extreme wave values that have a probability of occurrence of once every *N years*, termed the *N*-year return period. These return periods are typically set at 50 or 100 years, but sometimes also return periods up to 1,000 years are required as used in the mapping of flood risk areas in the United Kingdom (7).

The extreme value analyses can be carried out for various wave parameters such as Significant Wave Height H_s, Individual Wave Height h or also for a combination of parameters such as Wave Height or Mean Zero Up Cross period T_z. These can be evaluated by either carrying out

Extreme Value Analyses for each variable independently (i.e. Marginal modelling), or jointly (i.e. Conditional modelling) (24). The necessity to study other parameters besides H_s is becoming more important.

In Extreme Value Analyses (EVA) extreme peaks need to be selected from all the data. The chosen extreme peaks need to be *independent* events and ideally part of the same extreme distribution phenomena. In a large sample, these independent extreme values have an approximate distribution that is independent of the distribution of each wave variable (7) .

The first attempts to introduce a standard procedure to carry out the extreme value analyses of waves was done by a working group set up by the Maritime Hydraulics Section of the International Association for Hydraulic Research (IAHR) in 1990 (25). Prior to that the accepted procedure was to use all available data and fit a log normal distribution to the observed data.

The working group recommended the following procedure to carry out the EVA of H_s (25):

1. Selection of data for analysis. This involves data harmonisation, filling up any missing data by means of hindcasts and also the identification of outliers;
2. This data is subsequently sorted to choose the peak values that are independent yet form part of the same extreme distribution. This is a very important step for the subsequent accuracy of the analyses;
3. Fitting an appropriate distribution to the selected independent extreme data;
4. Assessment of the goodness of fit of the distribution;
5. Compute the extreme return values from the fitted distribution.
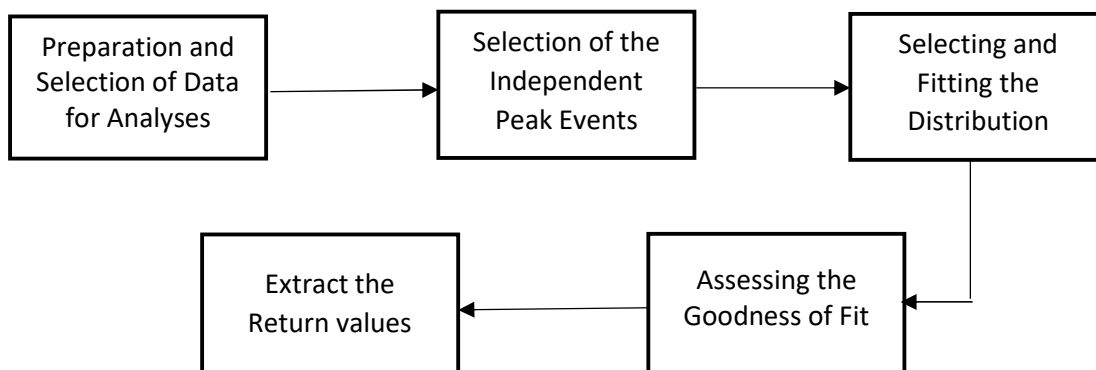
Figure 22 : Flow chart representing the main steps used in carrying out the EVA

3.1.1 Data and Sample data selection

As in the case of most statistical analyses, proper data selection is very important. The sea state data can be obtained either through direct measurements or by using models and generating hindcasts. The data should be complete as much as possible, especially since periods of non-data from in-situ measurements could potentially indicate periods of extreme bad weather (26).

Long term forecasts carry a lot of uncertainties, as the data available for analysis is much smaller by comparison. Hence the availability of longer data sets reduces the statistical uncertainties of the extreme predictions and errors. Data sets of 9-year, 17-year and 44-year periods were used in EVA studies. It was observed that the estimated 100-year return value of H_s (i.e. H_{S100}) decreased by approximately 10% with the longer datasets (27).

Outliers (i.e. values significantly larger than the other ones) should be looked into in more detail, since they could be extreme values belonging to a distribution with a longer return period than the current period of data analyses. Thus they could potentially provide valuable information (27). On the other hand, they could also be due to some error during measurement.

3.1.2 Data Harmonisation

Wave profiles continually change with time. A wave group consists of waves having similar properties and parameters. The time series of waves can be seen to be made up of waves belonging to different wave groups assimilated together. In EVA, the focus is on the wave group having extremely high values. In this analysis the extreme values of Significant Wave Height H_s will be taken. There is no standard approach to the selection of the peak data and it remains a subject of continuous debate and research (24).

During data harmonisation, the time series is analysed to select the independent extreme waves ideally belonging to the same extreme distribution (27). This is an important step since otherwise, storms from very different wave groups and hence with different properties will be studied together, affecting the analyses and results (27).

Various parameters can be used to select independent peak waves. Peak waves will be first sorted based on their H_s values. But since a storm can consist of a large number of peaks, to guarantee independent events, storms need to be distinguished from one and another. To this effect parameters related to the duration and time lapse between one storm and another can be used. Other parameters such as the time period of the year and wave direction could also be used (28).

3.1.3 Sorting data with time

As can be observed from any H_s time series, the H_s values vary considerably depending on the period of the year with higher values reached in the winter months and much lower values during the summer periods. This seasonal variation in peak H_s values indicates that the waves do not belong to the same wave distribution (28) . In a recent study conducted by *Soares et al,* (28), the H_s time series was sorted on a monthly bases and the 100-year return period of H_s (i.e. H_{S100}) was determined for each month. The predicted H_{S100} for the month of January was 12.17 m while that for July was 5.26 m. The months with such low peak wave heights have a significant effect on the EVA should they be included in the analyses.

Hence, only months having more homogenous wave data should be used. There is no general yearly period applicable for all areas and it is very much location dependant (28).

Since we are dealing with very long-term analyses, theoretically human induced effects such as global warming, should also be considered. This is a further complication and source of inaccuracy which was not considered in this study.

3.1.4 Wind/Wave direction

Theoretically wind or wave direction would be very useful in selecting the extreme events forming part of the same distribution. The peak significant wave heights would be binned based on the wave direction and the *N-year* return period for each wave direction calculated together with the *N-year* return period of the whole dataset (24).

However, this is very challenging since the wind and wave direction generally change considerably during storms. If the storm event is associated with just the one wind/wave direction registered during the peak H_s value, this could lead to an inaccurate analysis. The mean wave direction over the duration of the storm would give a better indication but still is not very accurate.

3.2 Data Selection of Peak Values

There are three main methods in which the peak data can be selected for EVA (29).

1. Initial Distribution (ID) Method

All the peak data from the time series are used for analyses with no sorting or sampling. It has the advantage that it generates a large amount of data over a limited time period.

But this method does not meet the requirements of Extreme Value Theory in that:

   a. the peak data are NOT independent events

   b. the data is unsorted so data belonging to very different wave distributions are studied together (28)

   c. periods of low wave conditions are included which affect the extreme wave estimates.

A 3- parameter Weibull distribution is normally used with this data (29).

2. *Annual Maximum (AM) Method*

Only the Annual Extreme wave height is used and all the other peaks are discarded. The chosen peak data are independent events but the data is too limited to carry out an accurate statistical study. Generally, a Gumbell distribution is used with this data set (29).

3. *The Peak Over Threshold (POT) Method*

An upper threshold value of H_s is set and only peaks that exceed this threshold are considered. Since within the same storm, H_s tends to vary a lot both above and below this threshold, controls are put in place to ensure that the peaks belonging to the same storm are grouped and that the different storms are clearly identified and are independent of one other. Only the highest wave in each storm cluster is chosen for further statistical analyses. Normally, a Generalised Pareto Distribution or Weibull distribution (2 or 3 parameter and truncated) are used with the POT method (29).

A comparative study between the above three methods of data selection, using different distributions and fitting techniques was carried out by *Van Vledder et al* (30). The maximum difference in the estimated 100-year return wave heights attained by the three methods was of around 10%.

Another study compared the results of the EVA at 4 open sea sites around the Mediterranean using the both the Annual Maximum and the POT method. The difference in the 100-year return values of H_s using the two methods was of 0.8 m, less than 10% (31).

The working group set up by the Hydraulic Research (IAHR) strongly recommend using the POT method of data selection with a Weibull distribution (25).

3.2.1 The Peak Over Threshold Method

The POT method consists of fitting a distribution to the *peak excess values of independent* storm events. The peak excess value is calculated by choosing the peak value within a storm cluster and subtract the upper threshold value. A distribution is fitted to these peak excess values allowing the return values to be calculated (7).

3.2.1.1 Distinguishing between Storm Periods / Independence of Peaks

As discussed above, it's important that peak events used in the EVA are independent. The following properties as indicated in figure (23) can be used to select independent peaks from an H_s time series (32):

1. Minimum storm duration, D : Since wave heights fluctuate above and below the set threshold during a storm, a minimum storm duration D will ensure that peaks within a short period of time are categorised together as part of same storm. Brief fluctuations below the threshold can be considered as part of the same storm.

2. Interarrival Time / Meteorological Independence Criterion, I : sets a minimum time interval between peaks for them to be considered independent. It also ensures that brief crossings below the storm threshold are not categorised as separate extreme event. The interarrival time is normally taken as 48 hours.

3. Mean wind / mean wave direction which is very difficult due to the variation in wind direction during a single storm.
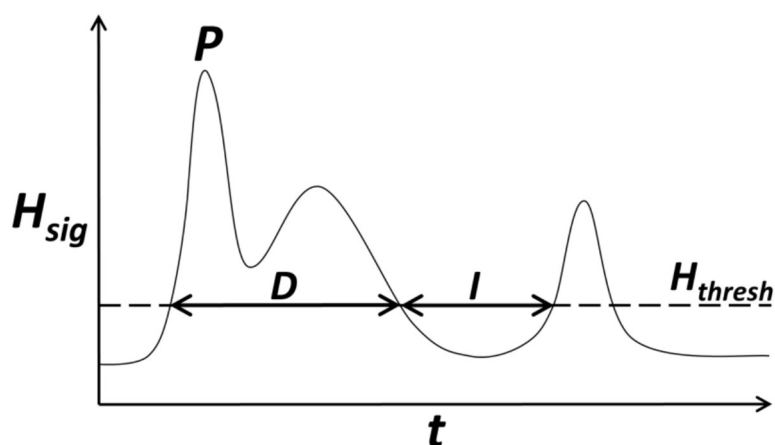
Figure 23: Definition of storm properties ; Reference (32)

There are no standard values or procedures to set the upper threshold value or the storm parameters since they are very site dependant (32). Yet the selection of the upper threshold value and the interarrival time are very important and changes to these values could lead to variations in the estimate of the return period (29).

In the studies carried out by *Van Vledder et. al.* (30), changes to the interarrival time did not affect the estimated return period by much. When the interarrival time was changed from 18 hours to 72 hours, the number of peaks selected went down from 228 to 144, but it had little or negligible effect on the estimated return wave height. Yet changing the upper threshold value affected considerably the estimated return wave height. Changing the threshold from 3.0 m to 9.0 m, the number of peaks reduced from 342 to 30 and H_{S100} changed from 16.0 m to 14.5 m (approxiamtely10%) , the higher the value of the threshold, the lower the value of H_{S100} (30).

Various studies have adopted different values for the storm parameters. The interarrival time is normally taken as 48 hours (7) but various studies have utilised a range of time between 6 and 96 hours.

Kapelonis et. al. (31) carried out EVA using POT method in various open sea areas around the Mediterranean (Figure 24). The interarrival time was set at 96 hours with different threshold values for the different areas (4.5 m for the Algerian and Tyrrhenian, 4.2 m for the Ionian and 3.7 m for the Levantine sub-basins)**.**
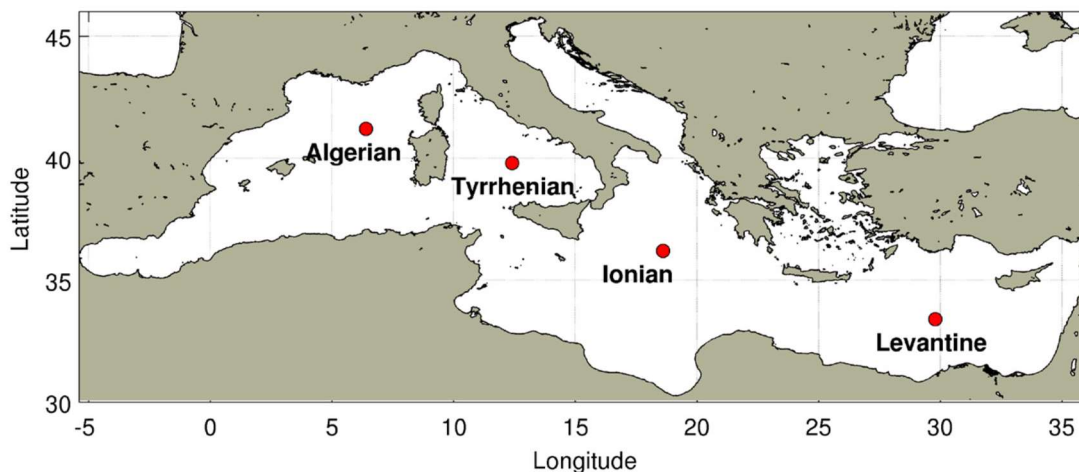
Figure 24: Map showing the areas analysed by Kapelonis et al.; Reference (31)

In certain studies, the 95-percentile value of the H_s timeseries is used as the upper threshold value, with a minimum storm duration of 9 hours and an interarrival time set at 18 hours (33).

3.2.1.2 Selecting the Upper Threshold Value

There is no clear procedure for setting the upper threshold value since it is very site-specific, depending on the bathymetry, the local characteristics and the exposure to the wind (33). The upper threshold value should be set to give the best compromise between bias and variance. Lowering the threshold too much will include waves belonging to other distributions, while setting the threshold too high will generate fewer peak excesses leading to statistical inaccuracies (29).

Fitting a Generalised Pareto Distribution to the excess values is becoming more accepted and popular. An advantage of using this distribution is that it helps in determining the upper threshold value depending on the changes in the distribution shape parameter when changing the threshold value. An optimal range of H_s upper threshold value is attained when changes to the upper threshold value leads to minimal change to the distribution shape parameter (7).

When fitting other distributions, more practical approaches are used for determining the H_s upper threshold value. These include using the 95-percentile value of the data [(24) (33)] or by determining a number of annual peaks or total number of peaks from the storm data (25). There is no ideal number of independent peaks, but it was suggested that a suitable number would be in the range of 50 -75 to be chosen from a wave data series that is as long as possible (30). *Katalinic et. al.* (29) set the threshold value giving an average of 4.3 peaks per year .

3.2.2 Comparing Distributions

In the early 1990, the IAH working group recommended using the POT method along with a Weibull Distribution fitted using the Maximum Likelihood Method. Since then, the GPD-Poisson model has grown in importance and is now the recommended and mostly used method (27).

There are 3 types of Weibull distribution - Three parameter, two parameter and the truncated. The three parameter seems to provide an acceptable fit for H_s for most conditions (25). Yet some studies have shown that the difference obtained in the *N-year* return value when using the three and the two parameter Weibull distribution is minimal (28).

The GPD-Poisson method sets the threshold value at very high values, leading to only a small number of peaks to be used for statistical analyses [(7),(24)]. This affects the accuracy of the results and hence the GPD does not necessarily yield the best fit as compared to other

distributions. In certain instances, the Weibull distribution gives a better fit to the selected data as compared to the GDP (27).

Another difference between GPD and the Weibull distribution is the tail end of the distribution. While the Weibull distribution is unbounded on the right-hand side, a GPD with a negative shape parameter is bounded, implying a finite upper value to H_s. Having a bounded upper value is reasonable because there is a physical limit to the size wave height can reach. In most recent studies (but not in all) the shape parameter of the GPD is negative (24). A finite upper bound of H_s does not imply an upper bound on individual waves within a storm, though, an upper bound on individual waves is to be expected too.

Fitting a Weibull and GP distributions are expected to give similar estimates for return values close to the length of available data. The differences in the estimated values will get larger for return values much longer than the data period (24).

As discussed above, an important advantage of using the GPD distribution is what is called the 'threshold stability property' which helps in determining a reasonable H_s upper threshold value. The threshold stability property implies that on changing the upper threshold value, the GPD's shape parameter should remain the same and the scale parameters are related (7).

If the shape and scale parameter of the distributions are changed, whilst not affecting the short term estimates , they could produce large differences when estimating the longer return periods (24).

3.2.2.2 Fitting the Distribution to the data

The Probability of non-exceedance is calculated for the selected peak H_s values. The chosen distribution/s are than fitted to the probability of non-exceedance plots (29).

The non-exceedance probability values can be calculated as follows:

1. For Initial Distribution Method of analyses

the *non-exceedance probability,* that is the probability wave height H_i is less than H, for actual data is given by:

$$Pr\,(H_i\,<\,H)\,=\,n_i\,/\,(N+1) \tag{16}$$

Where N = number of values

        i = rank of H; with i equal to 1 being the highest.

        n_i = number of observations *lower* than H which is equal to (N- i)

and the *exceedance probability*, that is the probability wave height H_i is greater than H, that is $Q\ (H_i > H)$, is given by:

$$Q\ (H_i > H)\ =\ i\ /(N + 1) \tag{17}$$

2. For Annual Maxima and POT method of analyses

Gringorten and Goda (29) recommend using the following non-exceedance probability equation :

$$P\ (H_i\ <\ H) = 1 - \frac{(i - 0.44\ )}{(N + 0.12)} \tag{18}$$

The Probability of Exceedance, that is the probability wave length H_i is greater than H

$$Q\ (H_i > H) = \frac{(i - 0.44\ )}{(N + 0.12)} \tag{19}$$

Where N = Number of values

      i = rank number with $i = 1$ being the highest.

Distributions are than fitted to the extreme data points using one of the below methods (25):

1. Least Square Method (LSM)
2. Maximum Likelihood Method (MLM) – least affected by outliers and can be used on truncated and complete datasets. Used with large datasets.
3. Method of Moments (MOM) - moments of the model and moments of the observed data are compared, the moments being the mean, standard deviation, skewness etc. It cannot be used on truncated data and is the preferred method when having small to moderate samples.

It is interesting to note that when fitting the same distribution to the same data yet using different fitting mechanisms, the N return values could vary considerably. Whilst the MLM and the LSM analyses had similar values , MOM gave the worst results (28). Based on the above, the Maximum Likelihood Method was chosen.

3.2.2.3 Assessment of Goodness of Fit of Distribution

After fitting the distributions to the actual data, the distribution that gives the closest results needs to be chosen. In certain cases, a visual inspection would be enough even though this is subjective test. When using a large number of points or many distributions, a quantitative test is preferred. The following types of quantitative tests (30) can be used:

1. Kolmogorov – Smirnov;
2. Anderson – Darling;
3. Chi -square;

4. Correlation coefficient, r;

5. The Coefficient of Determination, R^2;

6. Residue of Correlation Coefficient (REC);

amongst others. For this study the Correlation Coefficient, r , will be calculated. To evaluate the Correlation Coefficient, r , the observed values of Hs need to be compared to the H_s values derived from the distributions which have the *same probability of exceedance*. The higher the value of r the better the fit.

3.2.3 The *N*-year Return Period

Once the best fitting distribution is chosen, the *N-year* return period can be determined. The *N-year* return period can be described as the significant wave height that has a probability of occurrence of once every *N* years. Normally the *N* is set as either 50 or 100 giving a probability of occurrence of once every 50- or 100-year period. But the *N* value could be set any value or duration depending on the application of the data.

For POT analyses, the return period T_{Return} expressed in years for a wave height H_s having a probability of non-exceedance $P(H_s)$ is given by (30):

$$T_{Return} = \frac{1}{[n\,(1-P(H_s))]}$$

or

$$T_{Return} = \frac{1}{[n \times Q(H_s)]} \tag{20}$$

Where T_{Return} = Return Period (years)

$\quad\quad n \quad\quad$ = mean number of extreme events per year

$\quad\quad P(H_s)$ = probability of non- exceedance for wave height H

$\quad\quad Q(H_s)$ = probability of exceedance

3.2.3.1 Estimating the associated value of T_z

A rough estimate of the value of the mean wave period T_z at the time when significant wave height H_s reaches the extreme value can be calculated by assuming a wave steepness of 1/18 for open ocean or 1/14 in areas that are fetch limited [(9) (34)] . The Pierson Moskowitz spectrum for a fully developed sea takes a constant significant wave steepness of 1/19.7. The equation for wave steepness is given by:

$$Wave\ Steepness\ =\ 2\,\pi\,H_s/g\,T_z^2 \tag{21}$$

3.3 Return Value of Individual Wave height *h*

Besides determining the *N-year* return value of H_s, it would also be very important to determine the highest individual wave height *h* in that same time period. There are various methods in which this could be done (34)

Method 1: By calculating the highest wave h_{max} that will occur during the same storm as H_{sn} where H_{sn} is the *N* return value of H_s. This can be done by using the Rayleigh distribution equation (11) for the 3-hour period and using the calculated H_s found above.

$$P(h) = \frac{4\,h}{H_S^2}\ exp\left[-\frac{2\,h^2}{H_S^2}\right]$$
(22)

The most likely highest wave occurring during a 3-hour storm, $h_{\max 3hr}$, is given by:

$$h_{max\ hr} = K\,H_{sn}\,[(log\,N_n)/2]^{1/2}$$
(23)

Where N_n = total number of waves occurring in the 3 hours of *N-return* period;

              = $(3\ x\ 60\ x\ 60)\,/\,T_{zn}$

    H_{sn} = N-year return value of H_s;

    K = constant equal to 0.90 *(34)*;

    T_{zn} = mean zero crossing period storm containing the *N-year* return period;

Method 2: Estimating *h* for every selected H_s peak. The estimated value of the highest wave *h* is calculated for the 3-hour period related to each of the peak significant wave height H_s chosen from the data selection process. The same EVA procedure is than followed using these highest values of *h* instead of the peak H_s. The return value will then be for *h*.

Both the above methods assume the highest wave *h* will occur during the same storm as the highest H_s. This is not always the case and in a number of instances it was found that then highest *h* occurred during a storm having lower H_s (35).

Method 3: Battjes gives an alternative method for determining the highest wave height *h*. Using this method, the value of *h* attained is normally higher by a factor of 1.12 as compared to the value given by using the *N-year* return period of H_s. The Battjes theory in principle uses all the individual wave heights within a specific time period such as one year and from this data, to determine the wave height having a probability of occurrence once every *N-years* or percentage exceedance of 1 / *N*. This involves a huge amount of wave height data which in most cases is not available.

Hence, Battjes suggested using the H_s / T_z plots since H_s is calculated from h over fixed time periods (generally 3-hour periods) and H_s data is generally available. The Rayleigh distribution is used for each H_s / T_z box to determine the percentage exceedance of individual wave heights h for each H_s in each box.

Knowing the mean wave period T_z, over 3 hours, the approximate number of individual waves occurring in that H_s / T_z box can be evaluated. Multiplying the probability value obtained by the Rayleigh distribution by the number of waves in that time frame, gives an approximate value of the number of individual waves with wave height h in that H_s / T_z box.

Adding this value for each H_s / T_z box and dividing by the total number of waves over that period one attains the PDF and CDF from which the *N-return* period can be determined. As can be imagined there are a lot of assumptions and approximations. One drawback of this system is that it involves also calculating the number of waves in periods of low waves. In theory, as done previously in the POT method, these should not be included.

The Number of Waves occurring in the 3-hour period, N, is given by

$$N = (3 \times 60 \times 60)/ T_z = 10{,}800/ T_z$$

Where T_z = Mean Wave period

Hence the probable number of waves with wave height greater than h in 3 hours for a given H_s is given by:

$$P(Nr) = P(h\ //H_s) \times 10{,}800/ T_z \tag{24}$$

Where $P(h\ //H_s)$ = probability individual wave height is greater than h for a value of H_s

$$P(h\ //H_s) = \exp\ (-\ 2\,h^2 / H_s{}^2)$$

Adding over all values of H_s and T_z, the expected number of waves exceeding wave height h in 1 year, *M(h)*,

$$M(h) = \sum\sum \left[N(H_s,\ T_z) \times P(h\ //H_s) \times 10{,}800\ /T_z \right] \tag{25}$$

Where *M(h)* = expected number of waves exceeding wave height h

$N(H_s,\ T_z)$ = number of events falling in box centred H_s and T_z in 1 year

For a 100-year return period, *M(h)* would be 1 /100. Hence the value of h which gives an *M(h)* value of 0.01 is the highest individual wave having a return period of 100 years.

3.4 Data Validation

Model data will be used extensively both for forecasting and also for the climatology. To this effect, the model data will be compared to actual data to determine the level of correlation between the two sets of data.

3.4.1 Real Time Data

The in-situ data was collected using a Datawell Waverider Buoy. The buoy was deployed at a longitude of 14.16888 deg. and latitude of 36.08152 deg. on the 14th September 2011 (Figure 25 and 26). It was operational from the 20th September 2011 with the last transmission was on the 28th May 2014 at 09.17 am. It transmitted a wide array of data every 30 minutes including Significant Wave Height H_s (in cm) and the Mean Wave period T_z in seconds.

Figure 25: Aerial image of Location of the Buoy

Figure 26: Buoy on site

3.4.2 The *SWAN* Model

SWAN is a third-generation wave model for obtaining realistic estimates of wave parameters in coastal areas, lakes and estuaries from given wind, bathymetry and current conditions. *SWAN* can be used on any scale relevant for wind-generated surface gravity waves. The model is based on the wave action balance equation with sources and sinks. *SWAN* can be used on any scale relevant for wind generated surface gravity waves. However, *SWAN* is specifically designed for coastal applications that should actually not require such flexibility in scale.

The Physical Oceanographic Research Group has set up and runs the *SWAN* model on the Malta shelf domain as follows:

- • Minimum Longitude = 13.75º East;
- • Maximum Longitude = 15.45º East;
- • Minimum Latitude = 35.376º North;
- • Maximum Latitude = 37.124º North;

- Resolution: $0.002° \approx 200m$;
- No. of grid cells in the x-direction = 331;
- No. of grid cells in the y-direction = 271;

This domain extends from the southern part of Sicily to a latitude south of the Maltese Islands and is shown along with the bathymetry in Figure (27).
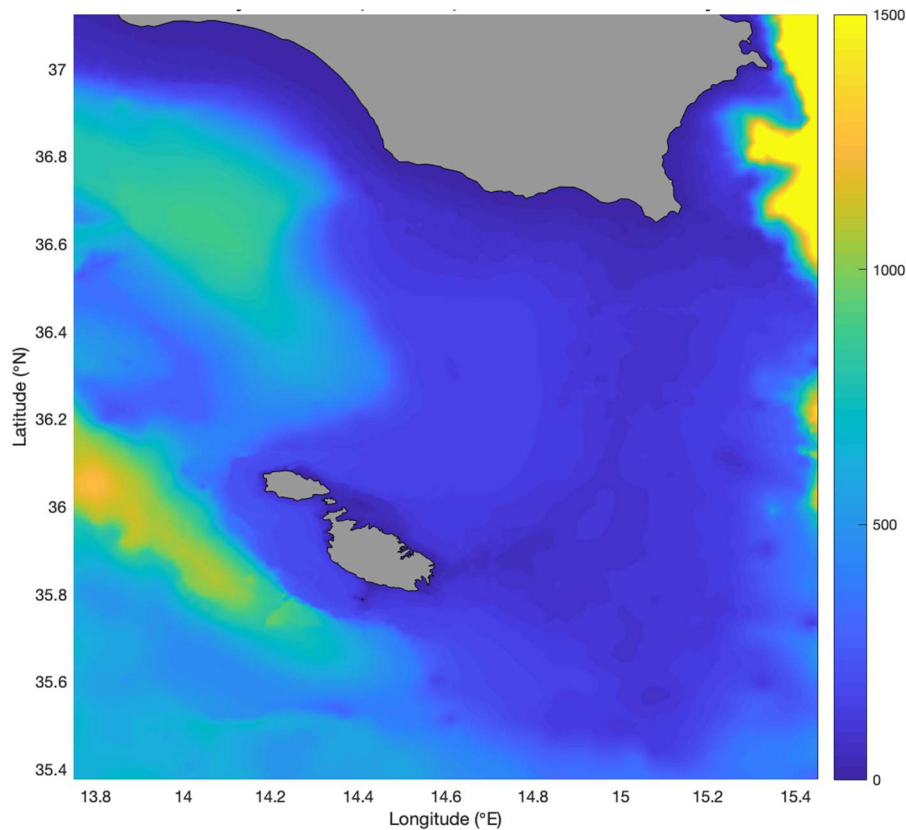
Figure 27: The ROSARIO-SWAN Domain showing the bathymetry of the area

The *SWAN* model requires the following data inputs

1.   Spectral information from the MARIA WAM model
ROSARIO-SWAN is nested within the *MARIA-WAM* system and spectral data at the points corresponding to the *ROSARIO-SWAN* boundary are provided by *MARIA-WAM* at a spatial resolution of $0.125°$ ($\approx 14km$) and a temporal resolution of 3 hours.

2.   Atmospheric forcing
The *ROSARIO-SWAN* model can be run from two wind data sources, *SKIRON* and *MARIA ETA* wind data

a. The *SKIRON* Wind data is provided by the Skiron Mediterranean Atmospheric forecast that is run by the University of Athens. The SKIRON 0.05 data is run daily at 10 metres height above ground with a spatial resolution of $0.005°$ ($\approx 600m$) and with a temporal resolution of 1 hour. The *SWAN* model data generated using the *SKIRON* wind data is termed *ROSARIO SKIRON SWAN*.

b. *MARIA ETA* Atmospheric model, The *SWAN* model data generated using the *MARIA ETA* wind data is termed *ROSARIO MARIA SWAN*.

3.   Current Data

The *ROSARIO SWAN* also utilises sea currents hourly data derived from the *ROSARIO -6420* hydrodynamic model. This model is an extension of the *ROSARIO 6420* Malta Shelf hydrodynamic model that was developed by the Physical Oceanography Research Group within the University of Malta. The *ROSARIO -6420* model is run daily to produce hourly forecasts of hydrodynamic parameters with a spatial resolution of $0.015625°$ ($\approx 1.7km$).

ROSARIO SWAN model has been running operationally since 8th December 2019. Generating hindcast data prior to this date presented a number of challenges mainly:
1. Re-generation of spectral data from *MARIA-ETA-WAM*.
2. The archive for *SKIRON* 0.05 data starts from 2016. Prior to 2016, only data from *SKIRON* 0.1 is available which consists of data with a resolution of $\approx 1.1km$ rather than $\approx 600m$
3. *ROSARIO 6420*-hourly started operation during 2013, and yearly complete data is only available from 2014 onwards. However, prior to 2014 data at the same exact domain but with a temporal frequency of 3 hours instead of 1 hour is available.

The gridded data files generated by the *SWAN* model consist of
1. Significant Wave Height, m;
2. Peak Wave Period, s;
3. Mean Absolute Wave Period, s;
4. Mean Wave Direction (in degrees using the Nautical convention – i.e., 0 signifies that waves are coming from the NORTH);
5. Depth, m;
6. Longitude and Latitude;
7. Date/Time;

Hindcast data from the *ROSARIO MARIA SWAN* model was extracted from 20.09.11 to 12.06.2019. Hindcast data from the *ROSARIO SKIRON SWAN* is available for the period 20.09.11 at 00.00 till the 23 December 2012 at 12.00.

3.4.3 Fitting of Linear Regression Line

In order to determine the relationship between the model data and actual data, linear regression lines were fitted to the corresponding scatter points using the Least Square Method. Linear regressions were used in previous comparative analyses between Model and Actual wave data conducted by the Physical Oceanography Group (36).

The linear regression equation between Model Data X and Actual Data Y is given by:

$$Y = m\ X + C \tag{26}$$

where m = slope coefficient and

C = intercept to the y axis

Slope coefficient or Regression Coefficient (m) was calculated using the following equation:

$$m = \frac{\Sigma\,[\,(X - X_{mean})\times(Y - Y_{mean})]}{\Sigma\,[\,(X - X_{mean})\,]^2} \tag{27}$$

Where X = value generated by model

Y = corresponding value measured by the buoy

Slope coefficient m indicates equality of X and Y and in an ideal case, m is equal to 1. Slope coefficient's value less than 1 shows overestimation of observations values Y, or underestimation of observation values when slope coefficient is greater than 1.

The Regression constant or Intercept to the Y-Axis was evaluated using equation

$$C = Y_{mean} - (m\,X_{mean}) \tag{28}$$

Theoretically, for a perfect match between the observed and the model data, the regression line would be linear with an intercept equal 0 and slope of 1.

3.4.4 Determining Closeness of Fit

To determine the closeness of fit of the linear regression line to the scatter points, the following parameters are used:

1. The Linear correlation coefficient, r
2. The Coefficient of Determination, R^2
3. Standard Error of Estimate or Root Mean Square Deviation (RMSD), s

The Linear correlation coefficient, r

The linear correlation coefficient is used to measure how well the regression line fits the scatter points derived from the model and actual data. A value $r = 1$ means a perfect positive correlation and the value $r = -1$ means a perfect negative correlation. In the current analysis, r is expected to be a positive value. Values of r greater than 0.8, indicate strong linear relationship between model and observations.

The Correlation Coefficient r was determined using the equation:

$$r = \frac{S_{xy}}{S_x \, S_y} \tag{29}$$

Where $S_{xy} = \dfrac{\Sigma(X - X\,mean) * (Y - Y\,mean)}{n}$ (30)

$$S_x^2 = \frac{\Sigma(X - X\,mean)^2}{n} \tag{31}$$

$$S_y^2 = \frac{\Sigma(Y - Y\,mean)^2}{n} \tag{32}$$

The Coefficient of Determination, R^2

This coefficient indicates how close the actual data (the data from the buoy) is to the linear regression line. Values of R^2 range from 0 to 1.0, with higher values of R^2 indicating that the regression line perfectly fits the data. Coefficient of determination was calculated from the following equation (33):

$$R^2 = 1 - \frac{\Sigma \, (Y - Y_{est})^2}{\Sigma \, (Y - Y_{mean})^2} \tag{33}$$

where $\Sigma \, (Y - Y_{est})^2$ = sum of squares of residuals,

Y_{est} = Y values calculated from the linear regression equation

Standard Error of Estimate or Root Mean Square Deviation (RMSD), S

Represents the average value the actual data differ from the fitted regression line. The Standard Error of Regression gives this difference using the units of the actual data. Smaller values are better because it indicates that the actual data are closer to the fitted regression line.

$$S_{yx} = \sqrt{\frac{\Sigma(Y - Y_{est})^2}{n}} \tag{34}$$

Chapter 4 : Methodology

4.1 Wave Forecasts

Besides H_s, the following wave parameters were also found to affect vessel stability:

1. Wave Length, L;
2. Wave Steepness;
3. Individual Wave Height, h;

Using the forecasted data generated by the *ROSARIO MARIA SWAN* model and using *MATLAB* for the calculations, 2-D Plots for the above wave parameters were generated for the areas surrounding the Maltese Islands.

4.1.1 Wave Length

Wave length kept emerging as a main cause of reduction in vessel stability, in particular the ratio wave length to vessel length. While vessel length is subjective to the end user, wave length can be determined based on the Linear Wave Theory *equation (4):*

$$L = \frac{g\,T_z^2}{2\,\pi}\ tanh\left(\frac{2\pi\,d}{L}\right) \tag{4}$$

Where L = wave length;

        T_z = mean zero up-cross period;

        d = depth;

For *deep water wave conditions*, when the ratio $d\,/\,L > 0.5$, wave length L is approximated by *equation (5):*

$$L = (g\,T_z^2\,/\,2\pi) \tag{5}$$

Whilst for *shallow water wave conditions,* when ratio $d\,/\,L < 0.05$, wave length L is approximated by following *equation (6):*

$$L\ = T_z\sqrt{g\,d} \tag{6}$$

For each of the hourly forecasts, all grid points are categorised as either deep water waves or intermediate / shallow water waves by evaluating the ratio $d\,/\,L_o$ with L_o being taken as the wave length based on the deep-water conditions that is $(g\,T_z^2\,/\,2\pi)$.

Subsequently, the wave length is calculated using either *equation (5)* or *equation (6)* depending on the water wave condition of each grid point at the time.

Solving for *L* in *equation (4)* for intermediate / shallow water conditions required setting up an iterative loop in *MATLAB*, substituting successive results of *L* back into the equation. The loop will continue to run until the difference between two consecutive values of wave length *L* that is *(L_N – L_{N-1})* is less than 0.1. Hence L_N is considered to satisfy *equation (4)*.

In this way 2-D plots of the wave length values based on the forecasted data generated by the *SWAN* Model for the entire area surrounding the Maltese islands was created. The time period and interval of the forecasts are the same as the data generated by the model. Contour lines representing the wave length to depth ratio were superimposed on these plots.

4.1.2 Wave Length Time series

In certain cases, more location specific information is required as in the case of mooring sites, or for off-shore structures. For such applications, the program also generates a predicted time series of wave length for any chosen coordinate. The coordinates are chosen manually by the user using the cursor. The program will choose the closest model grid point to the coordinates chosen and extract the forecasted time series for that nearest grid point. The data resolution is very high with the distance between each grid point of 200 m.

The data generated must be easy to follow and to comprehend. To facilitate interpretation of the values, the wave length values can be grouped together in bins. In this exercise, the wave length was grouped in 10 metre bins for the range 0 to 120 m.

Various other data can be superimposed depending on the application. In the case of broaching, where wave direction is also very important, wave direction can be superimposed over the wave length 2-D plots. The wave direction for each grid point is represented by the arrows, the length of which is proportional to the magnitude of H_s. This was done by breaking down H_s in the longitudinal and latitudinal components using the mean wave direction.

The mean wave direction *Ø* generated by the *SWAN* model is given in degrees using the Nautical Convention that is *Ø = 0* degrees indicates waves coming from the North.

The longitudinal value of H_s; *U*, is given by

$$U = \cos Ø \times H_s$$

Whilst the latitudinal value of H_s; *V*, is given by

$$V = \sin Ø \times H_s$$

4.1.3 Wave Steepness

Wave steepness is defined as the ratio of wave height and wave length. Having already established the wave length for each grid point and using the *ROSARIO MARIA SWAN* model for the forecasted data of H_s, the wave steepness for each grid point can be found.

Breaking waves are said to occur in areas having wave steepness ratios greater than a threshold value. There is no agreement on this threshold value, with two values given prominence (9)

Either wave length *L / 7*   or   wave length *L / 10*

Taking the smallest threshold value i.e., $L/10$ as the limit for breaking waves, the areas with a possibility of breaking waves can also be mapped. Wave direction is very important especially considering breaking beam seas and wave direction can also be superimposed.

4.1.4 Rayleigh Distribution

Significant Wave Height H_s is the defined as the average of the highest 1/3 individual waves occurring over a period of time, meaning that individual waves higher than H_s will occur during that time period. The significant wave height H_s is provided by the model forecast, but the individual wave height 'h' needs to be calculated.

The Rayleigh distribution is used to evaluate individual wave heights *h* for a particular significant wave height H_s. The probability an individual wave has a height *h* in a time period having significant wave height H_s is given by *equation (11):*

$$P(h) = \left(\frac{4\,h}{H_S^2}\right) exp\left(\frac{-2\,h^2}{H_S^2}\right)$$

(11)

Whilst the probability of exceedance, *Q(h)*, that is the probability that wave heights have values exceeding a value *h*, is given by *equation (12):*

$$Q(h) = exp\left(\frac{-2\,h^2}{H_S^2}\right)$$

(12)

There are various options for how this information can be presented. Three different methods were developed and the most practical remains to be determined depending also on the end use of such data.

Method 1 – To evaluate the highest wave that can occur in the next 3 hours, H $_{max\ 3\ hour}$

Reference is made to *equation (15)* with $H_{max\ 3\ hr}$ being defined by:

$$H_{max\ 3\ hr} = H_s \ (0.5 \ln N)^{1/2}$$

Where H_s = Average H_s over the 3-hour period;

N = Number of waves occurring in the 3-hour period = $(3\ x\ 60\ x\ 60) / T_z$;

T_z = Average Mean Wave period T_z over the 3-hour period;

Since the *ROSARIO MARIA SWAN* model provides hourly data of H_s and T_z, the mean of the hourly values of H_s and T_z over the 3 hours period are used.

Method 2: Percentage Exceedance

The operator specifies the wave height, *h* of interest depending on the vessel size or as specified on the vessel plate. Using Rayleigh's Distribution, the probability a wave will exceed the given height *h* is evaluated for every grid point and timeframe of the model. These result in 2-D plots of the percentage exceedance for the specified wave height *h* for each grid point around the Maltese islands. The higher the percentage exceedance, the greater the probability of encountering such waves, that is a 20 % exceedance means that a wave in every five exceeds the specified value. Knowing the H_s values for each grid point, *equation (12)* was used to determine the Percentage Exceedance, *Q(h),* for the wave height specified by the user.

Method 3: Individual Wave Height

Alternatively, the operator can specify the percentage exceedance of interest i.e., for wave energy forecasts we could be interested in the wave height *h* that say occurs 50% of the time. The output would be a plot of the minimum individual wave height *h* that meets that percentage exceedance. Again, this would be carried out for each grid point and time frame as the *ROSARIO MARIA SWAN* model.

This time, the Rayleigh distribution is used to determine the percentage exceedance for a range of *h* (in this case from 0 to 10 metres) based on the H_s value of each grid point. For each grid point, the range of wave height *h* is plotted against percentage exceedance. Using these graphs, the value of *h* that has the exceedance percentage equal to that specified by the operator is extracted. This method requires the most computational power of the three.

4.2 Climatology

The *ROSARIO MARIA SWAN* model was used to generate the hindcast data for the period between September 2011 and June 2019. The climatology was carried out for the grid point

closest to the location of the in-situ buoy that is for coordinates 14.1680 deg. longitude and 36.0812 deg. latitude. As described above, the model data generated consisted of:

1. Date and Time;
2. Significant Wave Height H_s, m;
3. Wave Peak Period $\underline{T}_p$ and Mean Zero-Up Cross Period T_z, s;
4. Wave direction;

The hindcast values of H_s and T_z obtained from the *ROSARIO MARIA SWAN* model were adjusted using the correction factors obtained from the model data validation exercise.

4.2.1 Extreme Value Analyses

Different methods or combination of methods can be used to carry out an Extreme Wave Analyses. These methodologies are continually being studied and evolved to find the most accurate and reliable method of analyses.

The objective of this part of the study is to provide a reliable estimate of H_s and h that have a probability of occurrence of once every 100 years. *MATLAB* was used for all calculations and plots. The relevant programs can be found in Appendix I.

4.2.2 Percentile Values

Percentile values are used extensively in statistics and can be described as the probability that a value H from a timeseries is not exceeded. The percentile values are found by first ranking the values and determining the probability of exceedance $Q\ (H_i > H)$ by using the equation:

$$Q\ (H_i > H) = i\ /\ (N+1) \qquad\qquad (35)$$

Where $Q\ (H_i > H)$ = probability H_i is greater than value H;

N = total number of observations;

i = ranking number of H with rank $i = 1$ giving the highest value of H;

The percentile values are found using $1 - Q\ (H_i > H)$. By plotting the percentile values against H, the value of H for a given percentile value can be found.

4.2.3 Peak Over Threshold Method (POT)

The extreme peak events were chosen using the Peak Over Threshold (POT) method. The POT method consists of selecting the highest H_s value *from each independent storm*. The highest value of each storm event is thought to belong to the same extreme distribution. The POT

method was chosen since it guarantees independence of the extreme peak values, yet it generates sufficient data for accurate statistical analyses should a sufficiently long data set be available

4.2.4 The Peripheral Months

The months with very low values of significant wave height, H_s, affect the outcome of the EVA particularly if percentile values are used to set the threshold values. As discussed in the literature review, it was decided to remove these peripheral months from the original time series. The peripheral months were determined as follows:

1. The 85, 90 and 95-percentile values of the H_s full time series were calculated. These values were successively set as threshold values;

2. The peaks exceeding the threshold value whilst allowing a minimum time period of 1 day in between successive peaks were selected;

3. These peaks were sorted by month. The months with a low number of peaks were considered as the peripheral months and omitted from the EVA.

4.2.5 Determining the Upper Threshold Value for the EVA

As previously discussed, there are no fixed guidelines or values for setting the upper threshold values in EVA and its value is very much site dependant. The threshold value should be ideally set to find a best compromise between bias and variance. Setting the threshold value too low will include peaks belonging to other distributions, whilst setting the threshold value too high will result in fewer peaks which gives rise to statistical inaccuracies.

The option of using the Generalised Pareto Distribution and setting the upper threshold in the range where there is no change in the distribution shape parameter was discarded in favour of a more practical approach that uses a combination of percentile values and a target number of peaks. A target *of between 50 and 75 peaks over the length of the time series* was set. The longer the time series, the more accurate this process would be since this would set the upper threshold at higher values. Considering that the data available covers a period of 8 years, peaks from a wider range of distributions, not only the extreme distribution, could also be included.

Various values were assigned to the upper threshold value to help understand how this parameter affects the *100-year H_s* forecast. Since the 95- percentile value was 3.5 m, the upper threshold value was given values within the range 2.5 m and 4.5 m.

4.2.6 Selecting the Independent Extreme Events

During storms, wave conditions change rapidly. Yet it is important that peaks belonging to the same storm are grouped together and that different storms are clearly identified and are separated from one another. To this effect, a number of conditions were applied to these peaks. Only the highest significant wave height from each storm cluster was chosen and subsequently used in the EVA.

Besides exceeding the upper threshold value, the following conditions were used to distinguish one storm event from another.

1. Interarrival Time - Peaks occurring within 24 hours of each other are said to belong to the same storm cluster. This is on the lower end with most papers taking an interarrival time of around 2 days;

2. Lower Threshold Value - H_s values must go below a minimum threshold value in between two extreme events for the events to be classified as separate and independent. This minimum threshold value was set at the 60-percentile H_s value of the time series so as to be slightly higher than the mean value of H_s (50-percentile) to allow for higher-than-average waves due to any oncoming swell. To find the points within the H_s timeseries that fall under this lower threshold, the H_s time series was inverted about the horizontal 60-percentile line. The peaks of this inverted time series would indicate the lower values of the time series. Two peaks were considered to be independent if a peak of the inverted time series was found in between the two peaks.

Using an H_s upper threshold value, interarrival time and the peaks of the inverted time series the extreme events were grouped in independent storms. Only the highest peak within each extreme storm is selected. Other data related to these peaks such as wave length, wave direction, wave steepness and zero up-cross period were also calculated, collected and analysed

4.2.7 Fitting Distributions to Extreme data

The peak values of these independent extreme storms, are thought be part of the same distribution. Fitting distributions to these peaks will enable the *N-return* period to be determined.

The Gringorten Equation, equation *(18)*, was used to determine the Probability of Exceedance, $Q(H_i > H)$, of the actual hindcast data

$$Q(H_i > H) = \frac{(i - 0.44)}{(N + 0.12)}$$

Where N = total number of observations;

     i = ranking number of H, with rank $i = 1$ being the highest;

Whilst the probability of non-exceedance $Pr(H_i < H)$ can also be found as $1 - Q(H_i > H)$. The Return Period are given by *equation (20)*:

$$T_{Return} = \frac{1}{n\ Q(H_i > H)}$$

Where T_{Return} = return period in years for wave heights H_i exceeding H;

     $Q(H_i > H)$ = probability of wave heights H_i exceeding H;

     n = average number or peaks per year;

As discussed in Section 3.2 distributions should be fitted to the Cumulative Distribution values by using the Maximum Likelihood Method (MLM). Weibull and the Generalised Pareto Distribution (GPD) are normally used when carrying out POT analyses with the GPD becoming increasingly popular. Gumbell distributions are normally used with Annual Maxima method of analyses. Since there is no fixed rule determining which distribution should be used, the three distributions, were fitted to the hindcasted peak data. The distribution which fits most closely to the actual data was subsequently determined.

Gumbell Distribution:

The following equations relate to the Gumbell Distribution:

$$PDF = \frac{1}{\alpha}\ exp\left[-\left(\frac{X_i - u}{\alpha}\right) - exp\ \left(-\frac{X_i - u}{\alpha}\right)\right] \qquad (36)$$

$$CDF = exp\left[- exp\ \left(-\frac{X_i - u}{\alpha}\right)\right] \qquad (37)$$

Where $S_x^2 = \frac{1}{(n-1)} \sum_{i=0}^{n}(X_i - X_{mean})^2$;

     $u = X_{mean} - 0.5772\ \alpha$;

     $\alpha = \frac{S_x\ \sqrt{(6)}}{\pi}$;

Using the actual data at hand, the values of S_x, u and α can be determined mathematically. Hence, the Cumulative Distribution and Probability Distribution plots can be made. The Probability of Exceedance, Q, and subsequently the *N-year* return period can also be found.

Weibull Distribution:

The following equations relate to the Weibull Distribution. The Exceedance Value (X_E) is obtained by subtracting the upper threshold value from the peak values of X.

$$PDF = (\beta/\alpha) \, (X_E/\alpha)^{\beta-1} \, exp[-(\frac{X_E}{\alpha})]^{\beta} \qquad (38)$$

$$CDF(X) = 1 - exp[-(X_E)/\alpha]^{\beta} \qquad (39)$$

Where α   = shape parameter;

     β = scale parameter;

     X_E = (X – Upper Threshold value);

MATLAB was used to fit a Weibull Distribution to the Exceedance Values X_E by using the Maximum Likelihood Method. Hence the values of the Shape and Scale parameters were found. Having found the values for α and β, the Probability Distribution Values, the Percentage Exceedance and subsequently the Return Periods can be calculated.

Generalised Pareto Distribution:

A similar procedure to the Weibull distribution was adopted for fitting the Generalised Pareto Distribution which has a PDF given by

$$PDF\,(X) = \left(\frac{1}{\alpha}\right)\left(1 - \beta\,\frac{X_E}{\alpha}\right)^{\frac{1}{\beta-1}} \qquad (40)$$

Where α   = shape parameter;

     β = scale parameter;

     X_E = (X – Upper Threshold value);

4.2.8 Determining the Best Fitting Distribution

The Cumulative Distribution plots for the three distributions and the actual hindcasted data were plotted together. Though visual comparisons give an indication of the best fitting distribution, a quantitative test was also sought. The Linear Correlation Coefficient *r* was chosen to compare and determine the best fitting distribution.

The Linear Correlation Coefficient, *r* was determined by comparing H_s values having the same probability of exceedance. To find these values of H_s, the probability of exceedance Q of the three distributions and of the hindcasted data were plotted separately against H_s. These plots were than used to find the values of H_s for probability of exceedance values between 0.05 to 0.95 increasing in steps of 0.05m.

Scatter plots of the actual H_s against the corresponding values of H_s derived from each distribution having *the same probability of exceedance Q*, were made. Using the least square method, linear regression lines were fitted to these scatter points. The correlation coefficient r was subsequently found.

4.2.9 The Return Period

The Return Periods are evaluated by using *Equation (20):*

$$T_{Return} = \frac{1}{n\, Q(H_i > H)}$$

Significant wave height H_s was plotted against the Return Period T_{Return}. Using the distribution equations and parameters, the plots were extended to cover the 100-year return period.

4.2.10 The Return Value of Individual Wave Height h

The forecast of the highest individual wave that will occur in the next *N-years, h_{max}*, can also be calculated. One method assumes that h_{max} will occur in the same extreme event as the highest H_s. The Rayleigh distribution is than used during the 3-hour period of the extreme event hosting the highest H_s. Hence, h_{max} is given by *equation (23):*

$$h_{max} = K\, H_{s\,n}\, [(\log N_n)/2]^{1/2}$$

Where N_n = total number of waves occurring in the 3 hours of n return period;

           = $(3\ x\ 60\ x\ 60)\ /\ T_{z\,n}$;

   $H_{s\,n}$ = N-year return value of H_s;

   K = constant equal to 0.90;

   $T_{z\,n}$ = Mean zero up- crossing period storm containing the n return period;

The second uses the method developed by Battjes that is *equation (25):*

$$M(h) = \sum\sum \frac{N\left(H_s,\, T_z\right) \times P(h\,//H_s)\ \times\ 10{,}800}{T_z}$$

Where $M(h)$ = expected total number of waves exceeding value h in 1 year;

   $N\left(H_s,\, T_z\right)$ = total number of counts in matrix $H_s\ /\ T_z$;

   $P\,(h\,//H_s)$ = Pr. of exceedance of h for a value of H_s based on Rayleigh Dist.;

                = $exp\,(-\ 2\,h^2\ /\ Hs^2)$;

   T_z = mean wave period;

It was decided to use this equation to tentatively evaluate the expected value of *h* for the *N-year* return period since it can be mathematically calculated using *MATLAB*. The methodology adopted was as follows:

1. The data for a full year was collected and the H_s and T_z values for 1 year were grouped together in packets. The H_s / T_z table, $N\left(H_{s,}\,T_z\right)$, was produced

$$N\left(H_{s,}\,T_z\right) = \begin{pmatrix} \left(H_{s,}\,T_z\right)_{1,1} & \cdots & \left(H_{s,}\,T_z\right)_{1,j} \\ \vdots & \ddots & \vdots \\ \left(H_{s,}\,T_z\right)_{i,1} & \cdots & \left(H_{s,}\,T_z\right)_{i,j} \end{pmatrix}$$

Where the rows represent the values of H_s in a pre-set range incrementing in groups of 0.50m, with the first row $H_{s\,1}$ being the lowest and the $H_{s\,i}$ being the highest.

The columns represent the values of T_z in a pre-set range incrementing in groups of 0.50 s, with the first row (T_{z1}) being the lowest and the (T_{zj}) being the highest .

$\left(H_{s,}\,T_z\right)_{i,1}$ can either be the actual number of counts or the probability of occurrence in each entry. In this case, the total number of counts in 1 year was taken.

2. The estimated number of waves that occur in 3-hour periods for each T_z was calculated using: $(\,3\,\times 60\,\times 60\,)\,/\,T_z$

$$\begin{pmatrix} \dfrac{10,800}{T_{z\,1}} \\ \vdots \\ \dfrac{10,800}{T_{z\,m}} \end{pmatrix}$$

The individual elements of matrix $N\left(H_{s,}\,T_z\right)$ were multiplied by the corresponding number of waves occurring in the 3-hour period. This now gives the total number of waves falling within period ($H_s\,T_z$):

$$Nr\left(H_{s,}\,T_z\right) = \begin{pmatrix} \left(H_{s,}\,T_z\right)_{1,1}\,x\,\dfrac{10,800}{T_{z\,1}} & \cdots & \left(H_{s,}\,T_z\right)_{1,n}\,x\,\dfrac{10,800}{T_{z\,1}} \\ \vdots & \ddots & \vdots \\ \left(H_{s,}\,T_z\right)_{m,1}\,x\,\dfrac{10,800}{T_{z\,m}} & \cdots & \left(H_{s,}\,T_z\right)_{m,n}\,x\,\dfrac{10,800}{T_{z\,m}} \end{pmatrix}$$

Adding all the elements within the same row would give the total number of waves falling within specific H_s groups for the chosen year:

$$\text{Total nr. of waves within Hs groups} = \begin{pmatrix} \sum_1^n (H_s, T_z)_{1,j} \, x \, \frac{10,800}{T_{z\,1}} \\ \vdots \\ \sum_1^n (H_s, T_z)_{m,1} \, x \, \frac{10,800}{T_{z\,m}} \end{pmatrix} = \begin{pmatrix} (T)_{H_s 1} \\ \vdots \\ (T)_{H_s m} \end{pmatrix}$$

Rayleigh distribution was used to determine the probability of exceedance for h grouped in 0.5 m groups for the range of 0 to 10 m, $P(h\,//H_s)$, for each of the H_s groups:

$$P(h\,//H_s) = \begin{pmatrix} P(h_1/H_{s\,1}) & \cdots & P(h_n/H_{s\,1}) \\ \vdots & \ddots & \vdots \\ P(h_1/H_{s\,m}) & \cdots & P(h_n/H_{s\,m}) \end{pmatrix}$$

Knowing the number of waves falling in each H_s group, multiplying with the probability of exceedance and adding together gives:

$$M(h) = \begin{pmatrix} (T)_{H_s 1} \\ \vdots \\ (T)_{H_s m} \end{pmatrix} \begin{pmatrix} P(h_1/H_{s\,1}) & \cdots & P(h_n/H_{s\,1}) \\ \vdots & \ddots & \vdots \\ P(h_1/H_{s\,m}) & \cdots & P(h_n/H_{s\,m}) \end{pmatrix}$$

$$M(h) = \begin{pmatrix} \sum_1^m P(h_1/H_{s\,i}) \, x \, (T)_{H_s i} & \cdots \cdots \cdots & \sum_1^m P(h_n/H_{s\,i})) \, x \, (T)_{H_s i} \end{pmatrix}$$

Where $M(h) =$ Expected total number of waves exceeding value h in 1 year.

So far, we evaluated *M(h)* which is the expected total number of waves exceeding value h in 1 year for a range of h from (0 to 10 in 0.5m groups). Values of *M(h)* were plotted against h and *MATLAB* was used to determine the best fitting curve and relative equation.

Subsequently using this equation, the wave height h that has a probability of exceedance of once every *N-years* is determined by finding the h with the probability of occurrence *M (h)* of *1 / n.*

4.3 Data Validation

The extreme value analysis was carried out using the hindcast data generated by the *ROSARIO MARIA SWAN* model. An analysis was done to assess the level of correlation between actual and model values for:

Significant Wave Height, H_s - Full range;

Significant Wave Height, H_s - Values in excess of 2.3 m;

Mean Wave period, T_z;

The *ROSARIO MARIA SWAN* model data was compared to the in-situ measurements obtained from the Datawell buoy. The buoy was deployed at a Longitude 14.16888 ⁰ East and a Latitude of 36.08152 ⁰ North over a period of 2.5 years between the 20 September 2011 and 28 May 2014. The model grid point closest to the buoy is at Longitude of 14.168 ⁰ East and Latitude of 36.0812 ⁰ North. The hindcast model data for this grid point was extracted for this period.

The models generated data every hour while the buoy transmitted data back every 30 minutes. Also, the buoy data series is not continuous. To compare the model with the actual data, the model data was sorted to be within 20 minutes of actual data. The rest of the data was not used. Correlations between the model and observations were done by fitting a linear regression line using the least square method for the full data and for each of the eight wind direction sectors. A linear regression line fitted using the least square method, represents the Y- values having the least deviation from all the data set. The linear regression equation between model data (X) and observations (Y) is given by:

$$Y = m\ X + C$$

where m = slope coefficient;

            C = intercept to the y axis;

Theoretically, for a perfect match between the observed and the model data, the linear regression line would have an intercept equal 0 and slope of 1. The following parameters were used to generate the regression line and to determine the closeness of fit of both models.:

     1. Slope coefficient or Regression Coefficient, m;

     2. Regression constant or Intercept, c;

     3. Linear correlation coefficient, r;

     4. Coefficient of Determination, R^2;

     5. Standard Error of Estimate or Root Mean Square Deviation (RMSD), S;

These were evaluated for the full time series and for each of the wave directions as indicated in *Table 4*. The Wave direction is given in degrees using the Nautical Convention that is the direction from where waves are coming.

Degrees	Centred at	Direction
337.5 – 22.5	360 (0)	N
22.5 - 67.5	45	NE
67.5 – 112.5	90	E
112.5 – 157.5	135	SE
157.5 – 202.5	180	S
202.5 – 247.5	225	SW
247.5 – 292.5	270	W
292.5 – 337.5	315	NW

Table 4: Wave direction sectors using the Nautical Convention

1. Slope coefficient or Regression Coefficient, m

The Slope coefficient m was calculated using *equation (27)*:

$$m = \frac{\Sigma[(X - X\ mean) * (Y - Ymean)]}{\Sigma[(X - X\ mean)]^2}$$

Where X = value generated by model

Y = corresponding value measured by the buoy

Slope coefficient, *m*, indicates the relationship between the model and actual data and in an ideal case is equal to 1. If the slope coefficient has a value less than 1, it means the model is over-estimating the actual value. Similarly, if the slope coefficient has a value greater than 1, it means the model is under-estimating the actual value.

2. Regression constant or Intercept, c

The Regression constant 'c' was evaluated using *equation (28);*

$$c = Y_{mean} - (m \times X_{mean})$$

3.The Linear correlation coefficient, r

The Linear Correlation Coefficient 'r' was determined using *equations (29-32);*

$$r = \frac{S_{xy}}{S_x\ S_y}$$

Where n = number of terms

$$S_{xy} = \frac{\Sigma(X - X\ mean) \times (Y - Y\ mean)}{n}$$

$$S_x^2 = \frac{\Sigma(X - X\ mean)^2}{n} \qquad S_y^2 = \frac{\Sigma(Y - Y\ mean)^2}{n}$$

4.The Coefficient of Determination, R^2

Values of R^2 range from 0 to 1, with higher values of R^2 indicating that the regression line perfectly fits the scatter points. Coefficient of determination was calculated from equation (33);

$$R^2 = 1 - \frac{\sum (Y - Y_{est})^2}{\sum (Y - Y_{mean})^2}$$

where $\sum (Y - Y_{est})^2$ = sum of squares of residuals;

$\qquad Y_{est}$ = Y values calculated from the linear regression equation;

5.Standard Error of Estimate or Root Mean Square Deviation (RMSD), s

Indicates the average difference between the regression line and the scatter points. The smaller the 's' value, the closer the data points are to the regression line.

$$S_{yx} = \sqrt{\frac{\sum (Y - Y_{est})^2}{n}}$$

In Extreme Value Analyses, only model data that exceed an upper threshold value is used meaning that potentially a better fitting regression line can be found for these higher values of H_s as compared to the full time series. The validation exercise was repeated for H_s values greater than 2.3 m to provide more accurate values in EVA.

The correction factor differed slightly from that of the full time series, and this corrected H_s values will be used in the EVA.

Chapter 5 : Results - Plots of Forecasts

In this chapter, the H_s, H_s >2.3 m and T_z data generated by the *SWAN* models will be validated against the in-situ buoy data. Any correction factors required, will be determined and applied to the data as applicable.

The 2-D plots for the forecasted data of wave length, wave steepness and individual wave height for the Maltese territorial islands will also be presented.

5.1 Data Validation

An analysis was carried out to compare the values of Significant Wave Height, H_s, obtained by *ROSARIO MARIA SWAN* model with the in-situ values obtained by the Datawell buoy. A similar analysis was carried out for the Mean Wave Period ,T_z, using the *SWAN* model running both on the *MARIA* and the *SKIRON* wind dataset. Figures (28 and 29) represents the H_s and T_z time series respectively for the data obtained from the in-situ Datawell buoy.
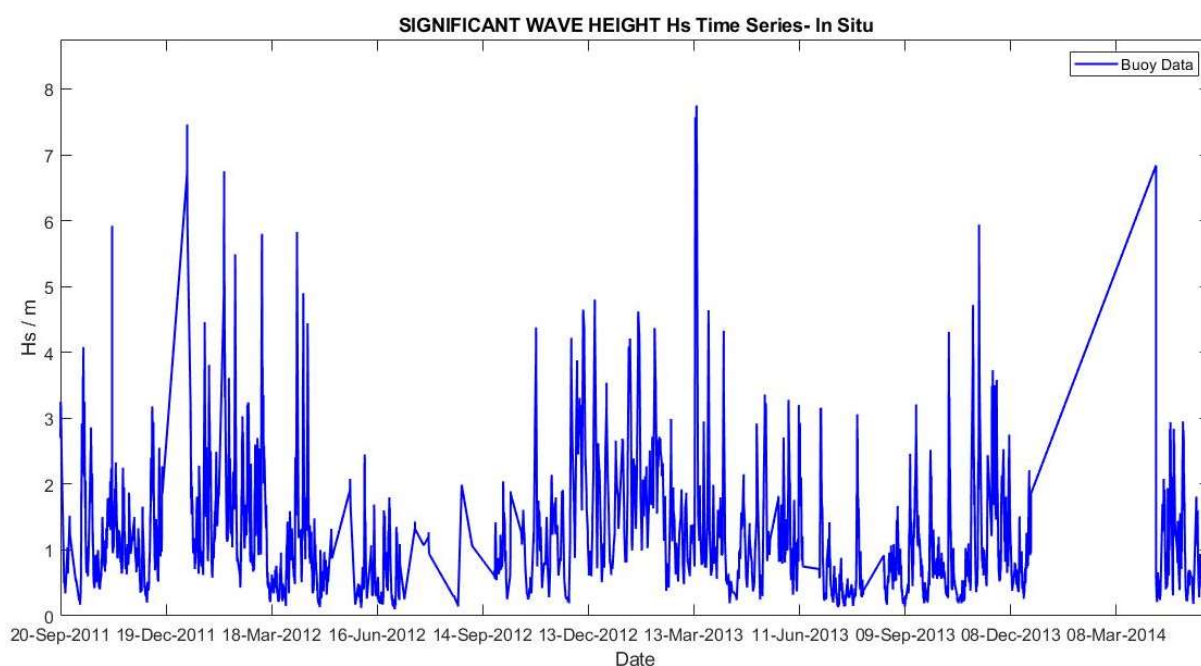
Figure 28: Significant Wave Height H_s with time as recorded by the in-situ Datawell Buoy
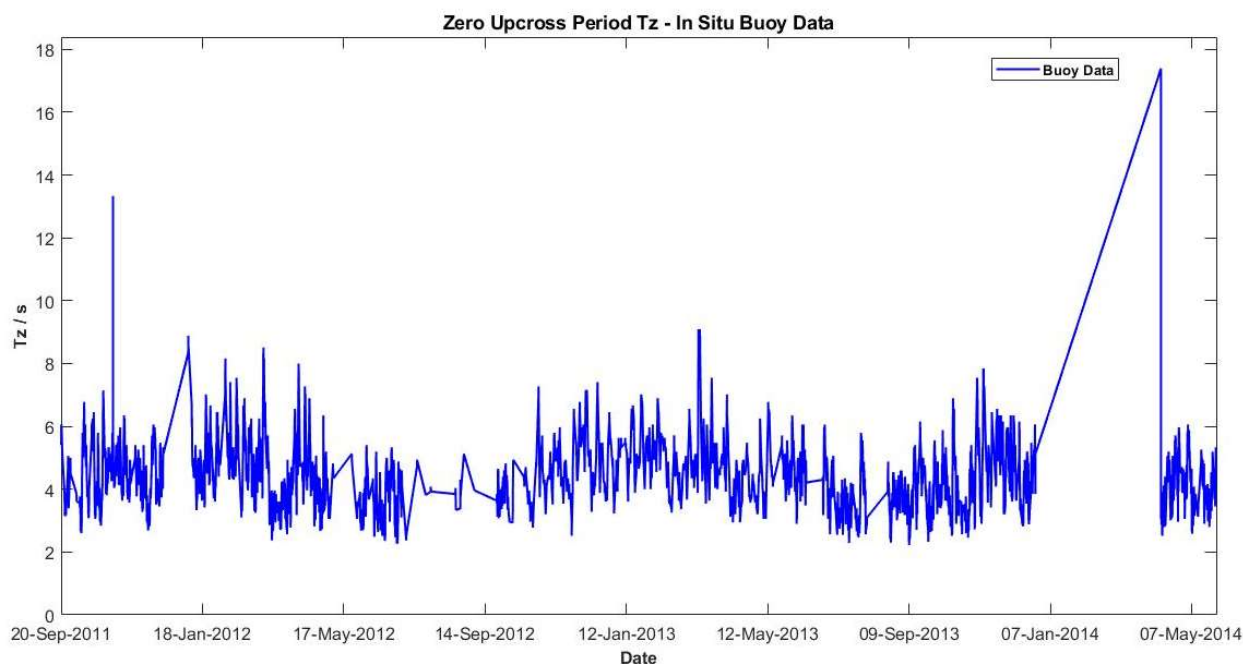
Figure 29: Mean Wave Period T_z with time as recorded by the in-situ Datawell Buoy

Data from the *ROSARIO MARIA SWAN* model covered the period 2011 – 2019. The time series of H_s and T_z using the *ROSARIO MARIA SWAN* model were plotted together with the observed data for the period 2011- 2014 (Appendix C).

The data was sorted such that model data and observed data were within 20 minutes of each other. The data was subsequently subdivided according to wave direction using the Nautical convention, that indicates the direction from where waves are coming. Data having unpractical values were removed from the observed data.

Degrees	Centred at	Direction
337.5 – 22.5	360 (0)	N
22.5 - 67.5	45	NE
67.5 – 112.5	90	E
112.5 – 157.5	135	SE
157.5 – 202.5	180	S
202.5 – 247.5	225	SW
247.5 – 292.5	270	W
292.5 – 337.5	315	NW

Table 5: Wave Direction using the Nautical convention

There are gaps in the data collected and transmitted by the buoy. Gaps of 5 days or longer occurred on the following dates. A noticeable gap occurred in the period between 25 December 2013 and 11 April 2014. This gap is clearly visible in the time series.

Year	Month	Day	Hour	Minute		Year	Month	Day	Hour	Minute
2011	9	28	11	26		2012	9	5	12	10
2011	12	16	7	57		2012	10	8	10	41
2012	5	8	19	27		2013	5	17	8	41
2012	7	9	15	26		2013	6	14	6	40
2012	7	18	17	26		2013	8	4	21	10
2012	7	30	17	26		2013	12	25	21	11
2012	8	27	10	40		2014	5	28	9	17

Table 6: Gaps of 5 days or longer in Data Transmission of the Buoy.

5.1.1 Significant Wave Height, H_s

Linear Correlation H_s (Full Time Series): ROSARIO MARIA SWAN / Actual Data

Scatter plots were used to determine the relationship between the model and the actual data, with model data on the X-Axis and actual data on the Y-Axis. The model data was taken for the grid point closest to the buoy location for the same moment in time.
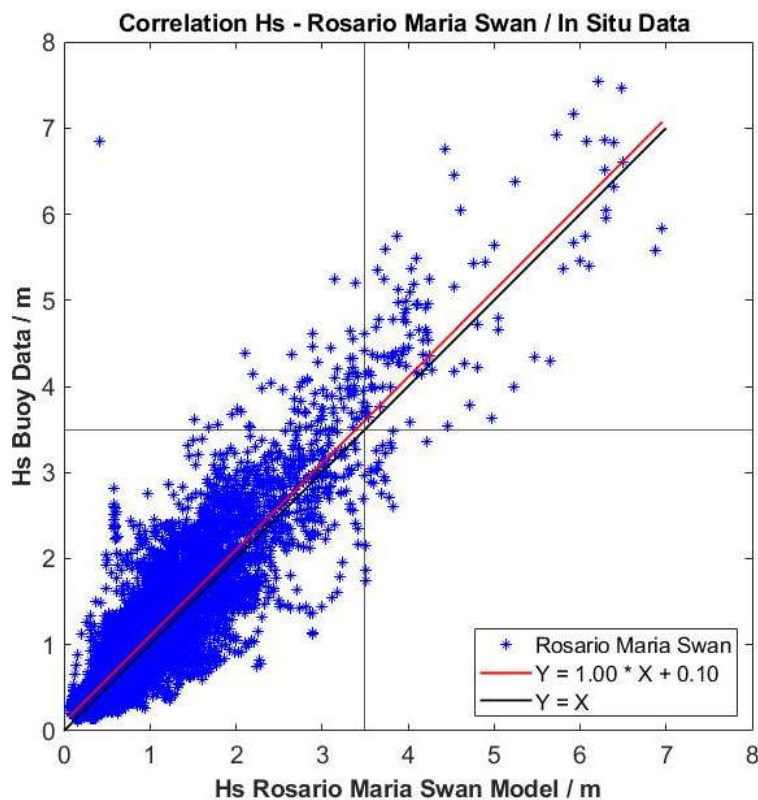
Figure 30: Correlation H_s Buoy / ROSARIO MARIA SWAN Model

The scatter plot of the *ROSARIO MARIA SWAN* H_s (X-Axis), against the corresponding in-situ H_s values (Y-axis) is shown in Figure (30). The scatter points are marked in blue, the fitted correlation line was plotted in red while the ideal regression line, meaning the model data replicates exactly the actual data, is drawn in black.

The values of the Regression Coefficient, *m*, Regression constant, *c*, Linear correlation coefficient, *r*, Coefficient of Determination, R^2, Root Mean Square Deviation (RMSD) were calculated and are represented in Table (7). *MATLAB* was used to carry out the calculations of these coefficients. Also, the *MATLAB* internal functions were used to confirm the values of the coefficients attained through calculations.

ROSARIO MARIA SWAN	r	M	R^2	Intercept	RMSD
Hs Full Time Series	0.890	1.00	0.791	0.1	0.387

Table 7: Values derived from Correlation all data Hs – ROSARIO MARIA SWAN / Buoy data

The fitted correlation line has a slope of 1 with an intercept value of 0.1. This indicates that the model data replicates reliably the actual buoy data *when considering the full H_s time series.*

Linear Correlation H_s (Reduced Time Series): ROSARIO MARIA SWAN / Actual Data

When carrying out the Extreme Value Analyses of H_s, only H_s values greater than an upper threshold value will be used. It is to be determined whether the model data replicates reliably the actual data of H_s for H_s values greater than the upper threshold value.

The upper threshold value has a significant effect on the end results of the EV Analyses. For this purpose, to better understand the effect of this parameter on the end EVA, the upper threshold value was assigned various values in the range 2.5 m to 4.5 m.

Hence, since the lowest upper threshold of H_s was set at 2.5m, an H_s cut off value of 2.3m was set to determine the level of correlation between the model and actual values. The model and actual H_s values less than 2.3m were removed from the data series and a linear regression line was fitted to the remaining scatter points (Figure 31).

ROSARIO MARIA SWAN	r	M	R^2	Intercept	RMSD
H_s All Data	0.889	1.504	0.791	0.097	0.387
H_s > 2.3	0.814	0.926	0.663	0.51	0.5951

Table 8: Values derived from Correlation of all H_s values and for H_s values > 2.3m
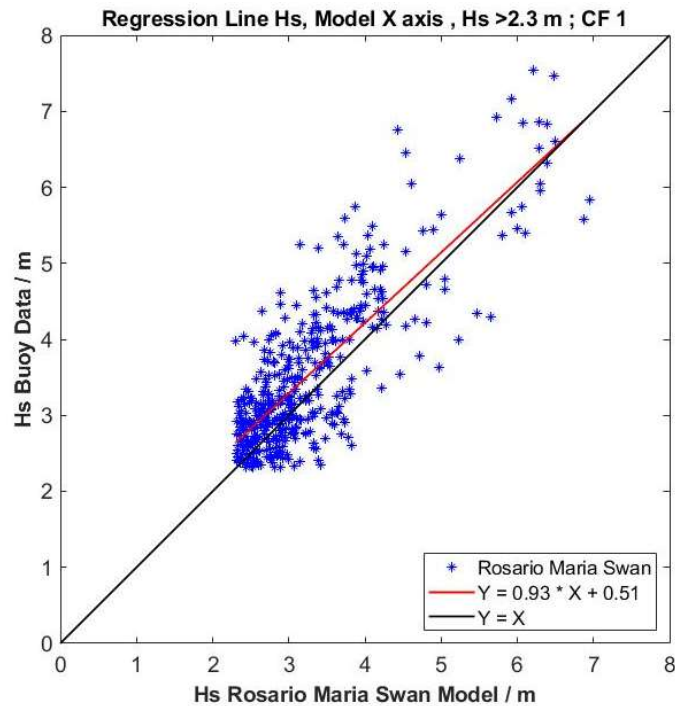
Figure 31: Fitted Correlation Line; ROSARIO MARIA SWAN for H_s Values > 2.3 m

The fitted regression line for the data series H_s > 2.3 m exhibits a slight difference as compared to the full H_s time series. With a slope of 0.93 and an intercept 0.51 m , the gap between model and actual data is wide at H_s values close 2.5 m getting very close at the higher H_s values of around 7 m.

Various correction factors were applied to the reduced H_s time series and the regression lines fitted. The values attained are presented in Table (9) whilst the corresponding plots of the regression lines with various correction factors are presented in Appendix (C.3).

ROSARIO MARIA SWAN – H_S greater than 2.3m

Correction Factor	r	M	R²	Intercept	Error Estimate
1	0.814	0.926	0.663	0.51	0.5951
1.05	0.8168	0.8647	0.667	0.58	0.583
1.10	0.8188	0.8051	0.670	0.67	0.5702
1.15	0.8251	0.7556	0.6807	0.7379	0.5512

Table 9: Values derived from Correlation of H_s > 2.3m and different Correction Factors

From these plots, it was decided to apply a correction factor of 1.05 to the model H_s data with values greater than 2.3 m. When comparing the regression line of the model H_s data >2.3m (Figure 31) and model data H_s >2.3m with a correction factor of 1.05 (Figure 32), the variation

between the ideal and the actual regression line is more uniformly distributed along the range of H_s values.

Also, when applying the 1.05 correction factor the model and actual regression line are in close proximity at H_s values of 4.5 m (Figure 32) rather than at 7 m as obtained when no correction factor is applied. As will be seen later, most of the H_s peaks used in the EVA (28.5%) occur at H_s values of 4.5 m (Figure 74). Hence it is very important that the model values used replicate closely the actual values at the 4.5 m mark.

For these reasons the 1.05 correction factor is applied when carrying out the EVA when only H_s values greater than 2.3 m will be used.
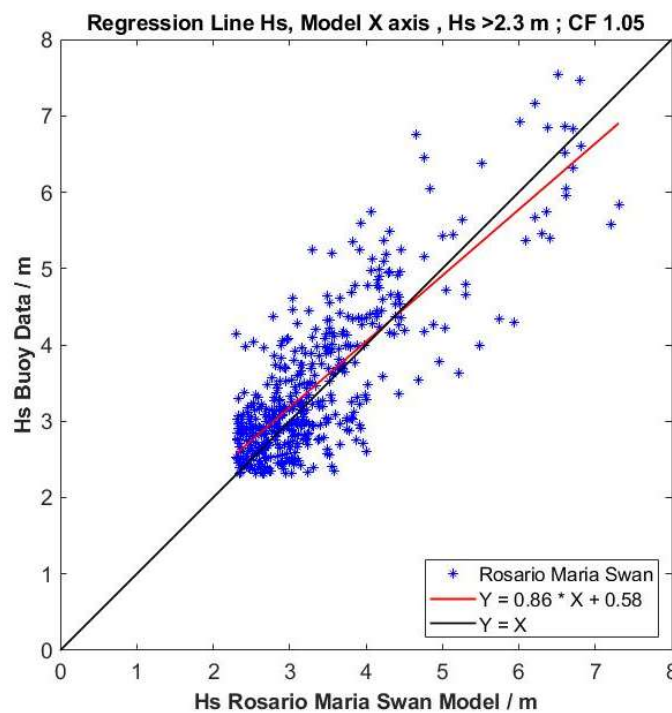
Figure 32: Fitted Correlation Line ROSARIO MARIA SWAN for H_s Values > 2.3 m. Correction Factor 1.05

5.1.2 Mean Wave Period, T_z

The time series of the Mean Wave Period T_z using the *ROSARIO SKIRON SWAN* Model and *ROSARIO MARIA SWAN* were plotted in Appendix (B.1) and Appendix (C.1). The linear correlation plots of T_z using the *ROSARIO SKIRON SWAN* and the *ROSARIO MARIA SWAN* models against the actual in-situ values for the period 2011 and 2012 were produced (Figure 33).
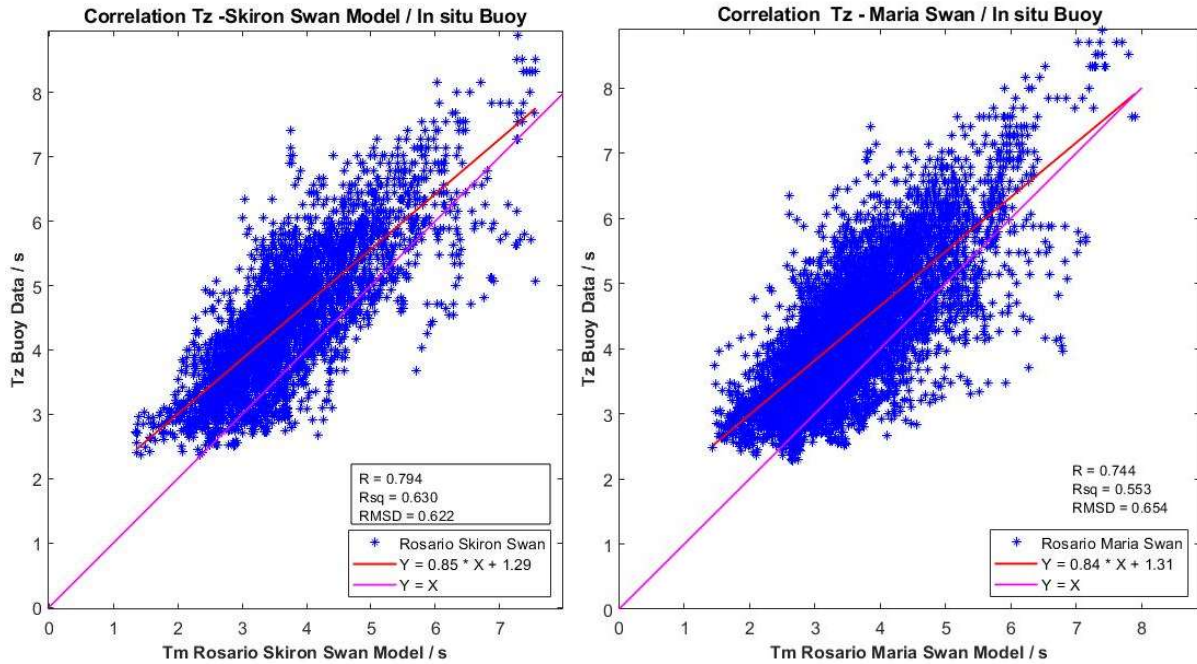
Figure 33: Correlation of T_z derived from Observation against ROSARIO SKIRON and ROSARIO MAIA SWAN Model

T_z	Period	r	M	R^2	Intercept	RMSD
ROSARIO SKIRON SWAN	2011-2012	0.794	0.854	0.630	1.295	0.622
ROSARIO MARIA SWAN	2011-2012	0.744	0.835	0.553	1.066	0.623

Table 10: Values derived from Correlation of T_z – SKIRON and MARIA SWAN against Buoy data

From the above plots is emerges that at higher T_z values, the regression line gets closer to the ideal regression line. The linear correlation plots for the data sorted by wave direction can be found in Appendix (B.2), Appendix (C.2) respectively. The values are summarised in Table (11) and Table (12) hereunder.

ROSARIO SKIRON SWAN:

Direction	r	M	R^2	Intercept	RMSD
N	0.80	0.96	0.64	1.15	0.48
NE	0.75	0.75	0.56	1.70	0.75
E	0.70	0.64	0.49	1.87	0.64
SE	0.82	0.86	0.67	0.88	0.42
S	0.70	0.81	0.49	1.01	0.55
SW	0.79	0.81	0.63	1.41	0.46
W	0.88	0.99	0.78	1.04	0.52
NW	0.89	0.98	0.79	1.04	0.44

Table 11: Values derived from Correlation of T_z data sorted by direction – ROSARIO SKIRON SWAN

ROSARIO MARIA SWAN:

Direction	r	M	R²	Intercept	RMSD
N	0.760	0.942	0.578	1.066	0.623
NE	0.744	0.776	0.554	1.463	0.689
E	0.687	0.635	0.472	1.781	0.616
SE	0.742	0.757	0.551	1.216	0.457
S	0.563	0.634	0.316	1.539	0.566
SW	0.679	0.658	0.462	1.898	0.619
W	0.839	0.975	0.704	1.093	0.576
NW	0.865	0.939	0.748	1.198	0.464

Table 12: Values derived from Correlation of T_z data sorted by direction – ROSARIO MARIA SWAN

5.1.3 T_z Comparison *ROSARIO MARIA SWAN / SKIRON SWAN*
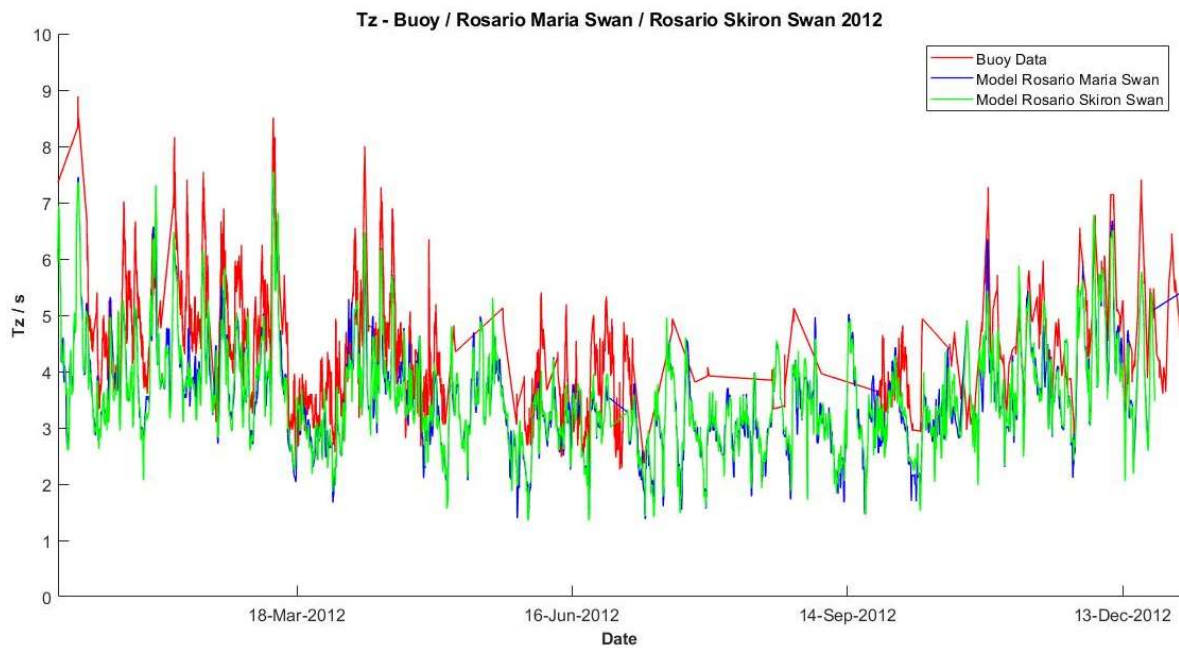
Figure 34: Time series T_z – Actual Data, ROSARIO MARIA SWAN Data, ROSARIO SKIRON SWAN Data

Figure (34) shows the time series for the actual T_z data, and the T_z data generated by the *ROSARIO SKIRON SWAN* and the *ROSARIO MARIA SWAN* models. The linear regression lines obtained by the two models were plotted together and compared to the ideal regression line in Figure (35). As can be observed, both models are very close, whilst having a noticeable difference from the ideal Regression line.
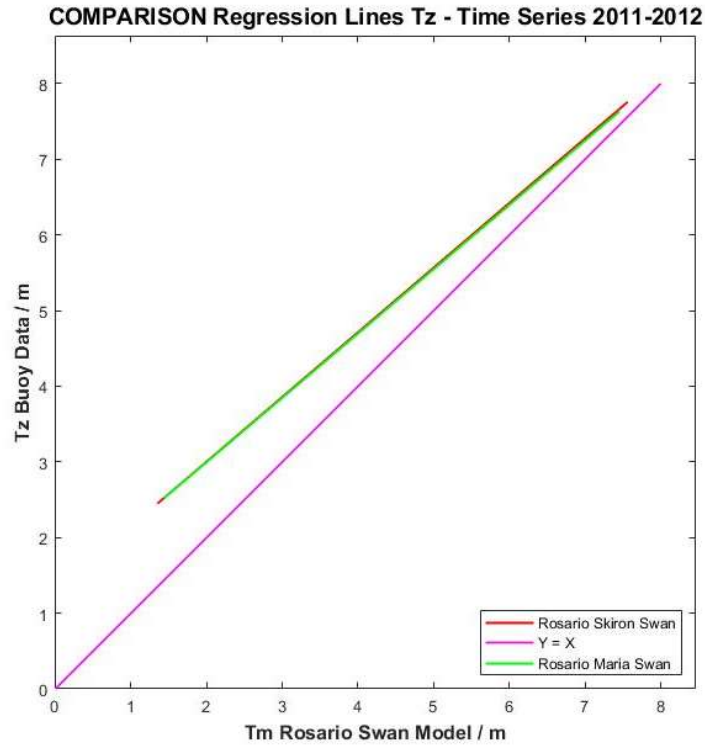
Figure 35: Plots Regression Lines Tz Models against ideal regression line

All Data – T_z	r	M	R^2	Intercept	RMSD
ROSARIO SKIRON SWAN	0.794	0.854	0.630	1.295	0.622
ROSARIO MARIA SWAN	0.744	0.835	0.553	1.312	0.654

Table 13: Values derived from Correlation of T_z ROSARIO SKIRON SWAN and ROSARIO MARIA SWAN

5.1.4 Correction Factor – Mean Wave Period T_z

The purpose of these comparative plots is to understand how the model data is replicating the actual data. If necessary, a correction factor can be applied with the model data such that the model replicate better the in-situ data.

All Data - T_Z	r	M	R^2	Intercept	RMSD
ROSARIO SKIRON SWAN	0.744	0.835	0.553	1.312	0.654
ROSARIO SKIRON SWAN - Corrected	0.744	1.04	0.553	0	0.654

Table 14: Values Correlation lines T_z and Corrected T_z - ROSARIO MARIA SWAN

The data obtained from the *ROSARIO MARIA SWAN* model was corrected as follows:

$$T_Z \ = \ 0.8 \, T_{Z\,MODEL} \ + \ 1.25$$
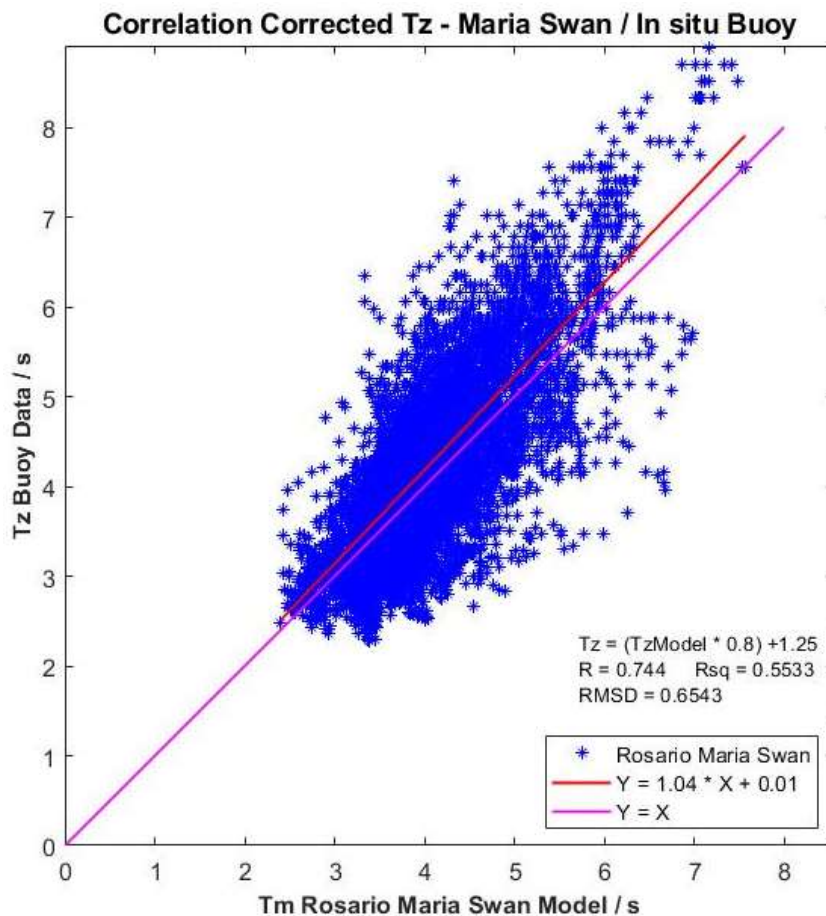
Figure 36: Correlation line based on corrected T_z values

5.1.5 Wind Strength and Direction

The 3-hourly wind data obtained from *ROSARIO MARIA* which serves as data input to the *ROSARIO MARIA SWAN* model consists of:

1. Date and hour;
2. Wind strength resolved in the U direction, m/s;
3. Wind strength resolved in the V direction, m/s;
4. Wind strength, m/s;
5. Wind direction;

The data available covered the period 2011 – 2014. The wind direction was changed to follow the Nautical Convention that is the direction from where the wind is coming to be in accordance with the convention followed by the wave direction. Data was sorted to eliminate incorrect data.
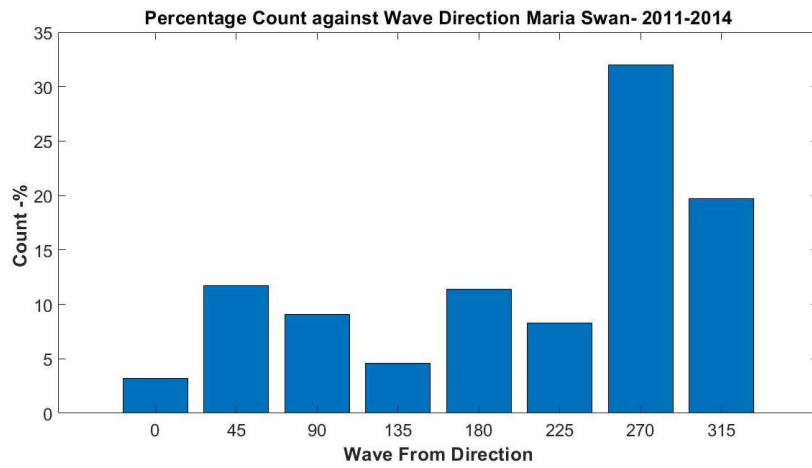
Figure 37: Percentage Occurrence Wave direction
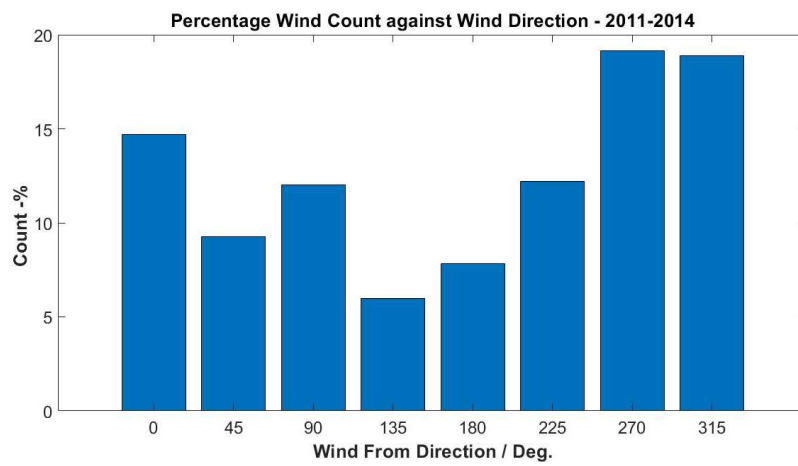
Figure 38: Percentage Occurrence Wind direction

Dir.	Wave Dir. %	Wind Dir. %
N	3.5	14.7
NE	13.2	9.2
E	11	12
SE	4.5	6
S	10.2	7.8
SW	8.1	12.2
W	29.3	19.1
NW	20.1	18.9

Table 15: Number of Occurrences Wind / Wave Direction

5.2 Plots of Forecasted values of Critical Wave Parameters

As previously discussed, wave length L, individual wave height h, and wave steepness have a significant impact on vessel stability. It is interesting to note that IMO advises the ship master to constantly measure and monitor wave parameters such as wave length and wave period amongst others during navigation. But while the physical monitoring of wave conditions by the vessel master is important, it could also mean that the vessel is already caught within an undesirable event before corrective action can be taken. There is no doubt that the best way to deal with extreme events is avoid them in the first place and that is why forecasts are so important. Whilst such information should assist in safer navigation, so far, no weather forecast provides such information. The forecasted data generated by the *ROSARIO MARIA SWAN* model on the 06th December 2020 was used to generate the 2-D forecast plots for the wave parameters:

Wave length, L;

Individual wave height, h;

Wave steepness;

The data generated was hourly and the forecasted period was of 37 hours. The below plots were produced. Further plots taken over the 37-hour forecasted period taken at a 5-hour interval can be found in Appendix (D).

	Atmospheric conditions and temperature °C	feels like °C	Atmospheric pressure mmHg	wind speed m/s	Humidity
Night	+18°	+18°	746	▼ SE 10.3	88%
Morning	+18°	+18°	743	▲ S 10.9	87%
Day	+15°	+15°	746	► W 8.7	75%
Evening	+14°	+14°	748	► W 8.7	56%

Table 16: Weather forecast 06 December 2020

5.2.1 Forecasted Wave Length, *L*
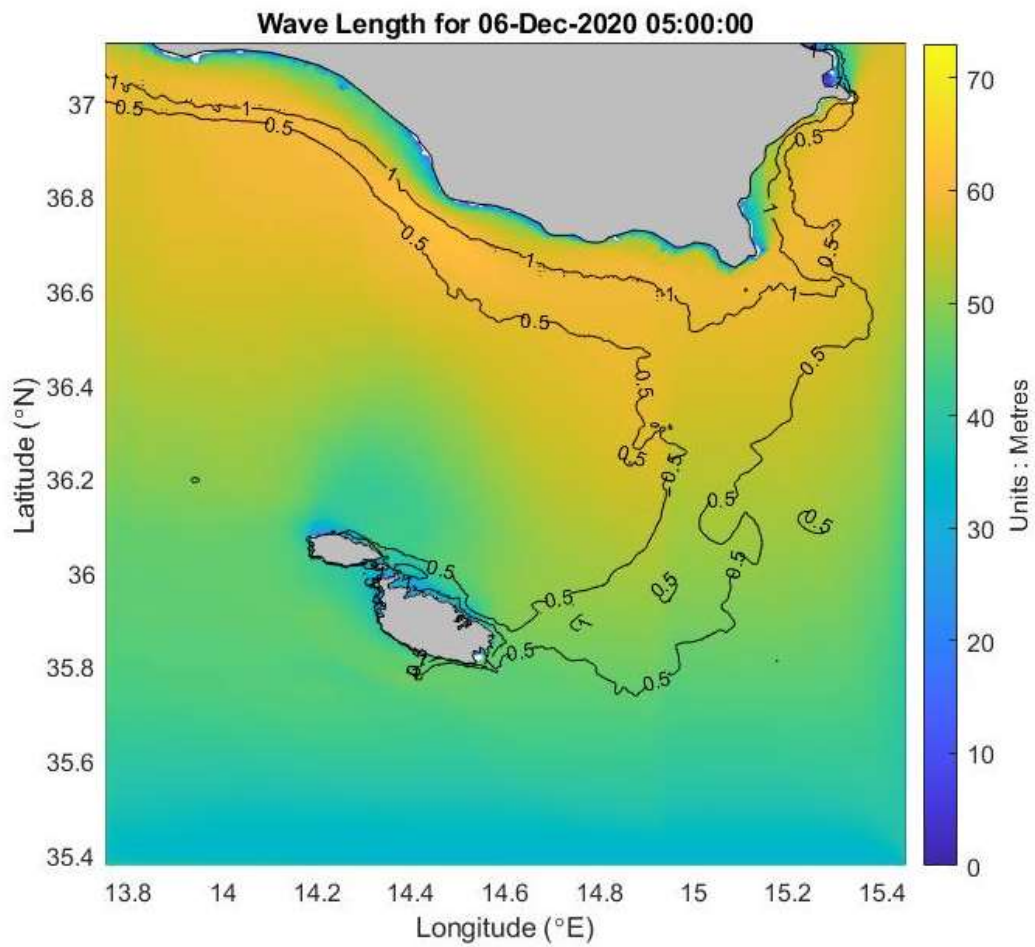
Figure 39: Forecasted wave length , 06 December 2020 at 05:00 with contour lines wavelength / depth

Figure 39 shows the forecasted wavelength for the 06th December 2020. The contour lines represent the ratio depth *d* with wave length *L*.

5.2.2 Forecasted Wave Length *L*, Values Grouped
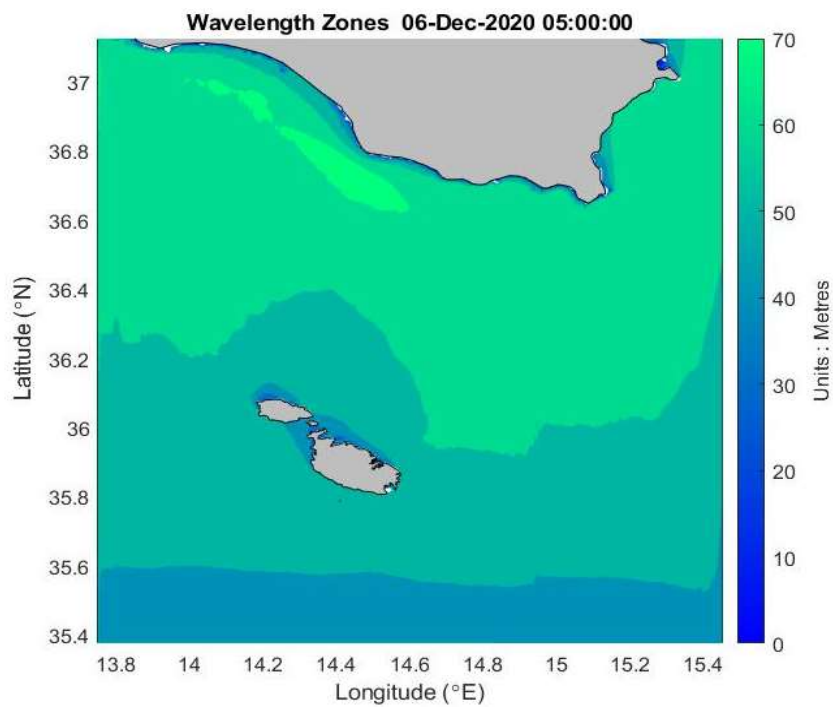
Figure 40: Forecasted wave length, 06 December 2020 at 05:00, wave length Values binned

5.2.3 Forecasted Wave Length, *L*, with Wave Direction
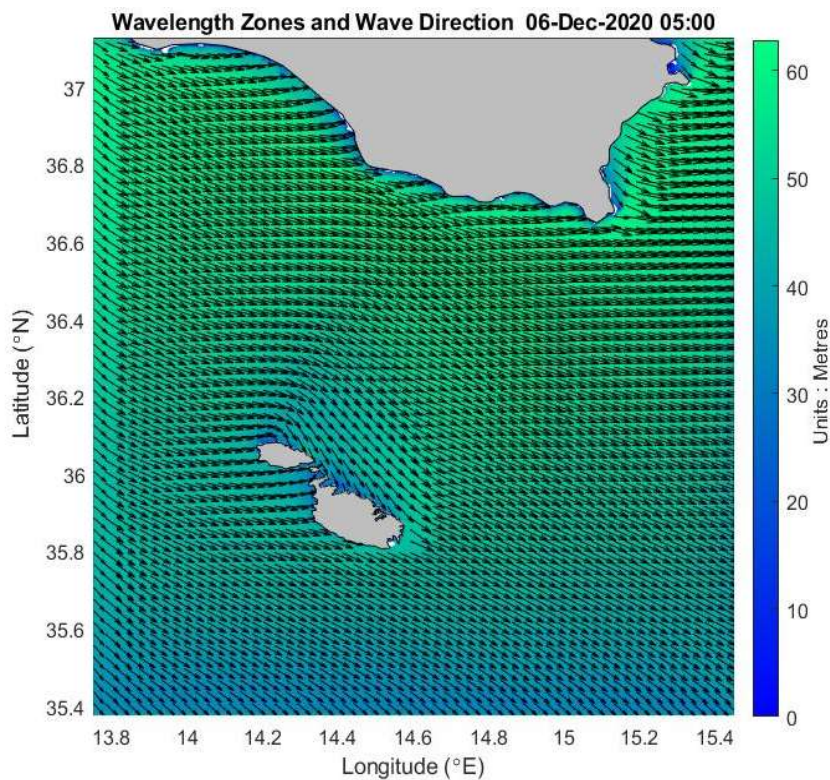
Figure 41: Zoned Wave Length with Wave Direction

5.2.3 Forecasted Highest Wave, h, occuring in the following 3 hours; $h_{max\,3h}$
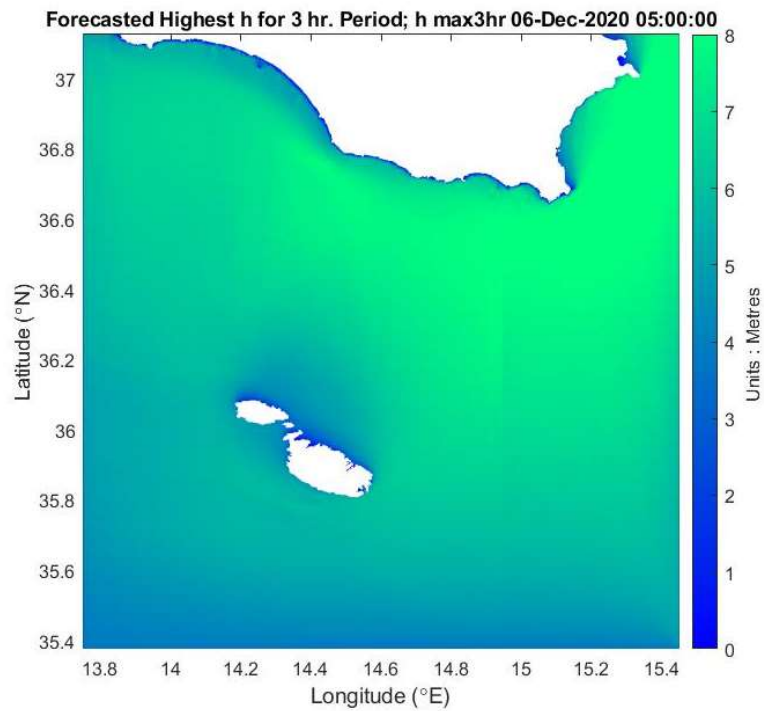
Figure 42: Forecasted Highest Wave in the following 3 hr period h_{max} 06 December 2020, 05:00

5.2.4 Percentage Exceedance for an operator specified Individual Wave Height h for 3-hour time periods
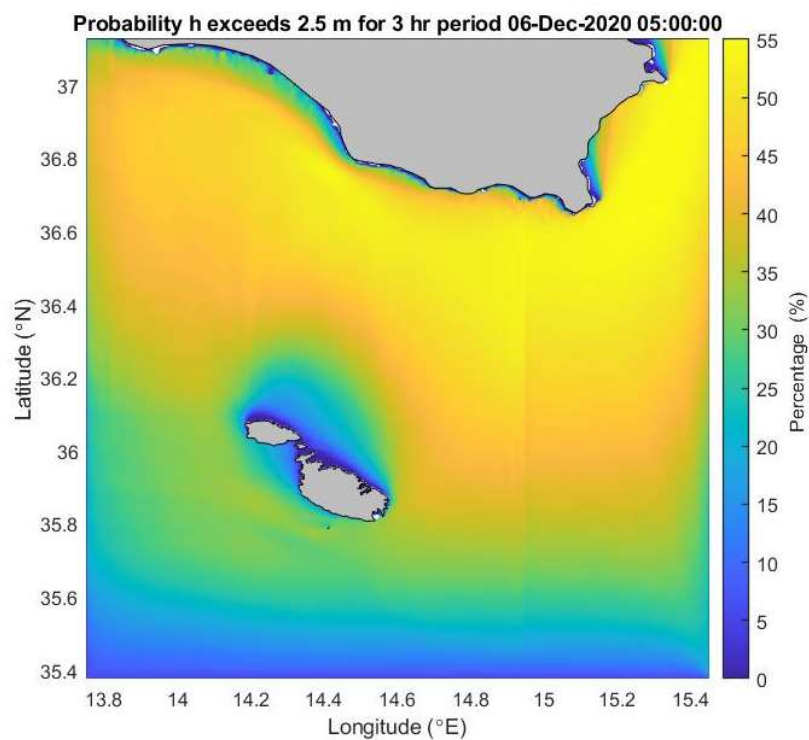
Figure 43: Probability 'h' exceeds operator specified wave height, 2.5m on 06 December 2020, at 05:00 valid for 3 hour time period

5.2.5 Forecasted Individual Wave Height *h*, for a specified Percentage Exceedance
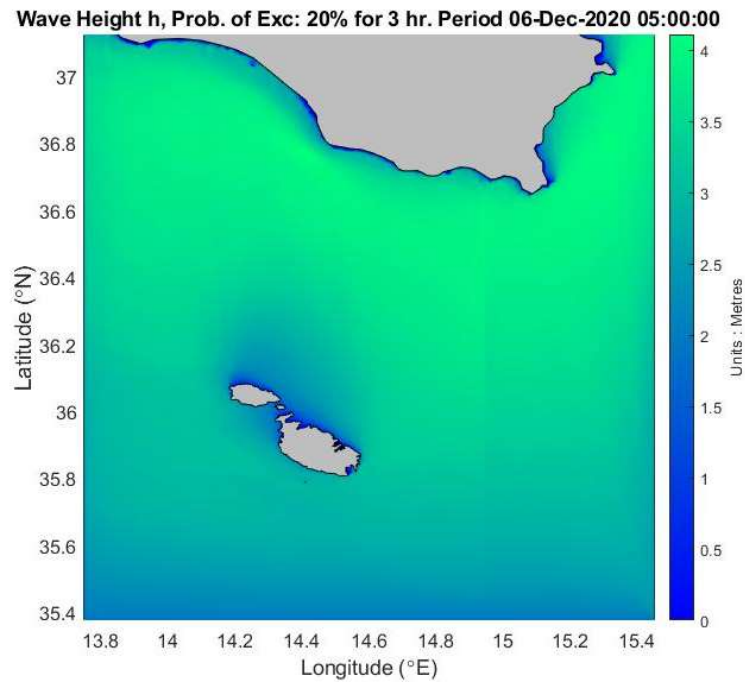
Figure 44: individual Wave height, h, having Percentage exceedance greater than operator input (20%); 06 Dec. at 05:00 valid for 3 hour time period

5.2.6 Forecasted Wave Steepness
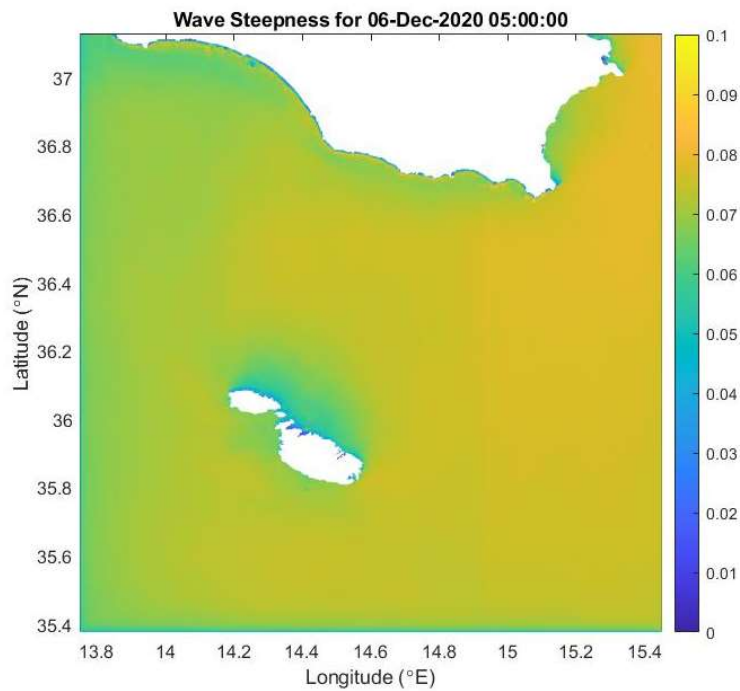
Figure 45: Plot of forecasted Wave Steepness; 06 December 05:00 am

Having determined the forecasted values of wave steepness and wave length for each grid point, the areas were breaking waves can occur can be found assuming breaking waves happen when wave heights are greater than the ratio (Wavelength L / 10).

5.2.7 Forecasted Wave Length time series for operator specified grid point
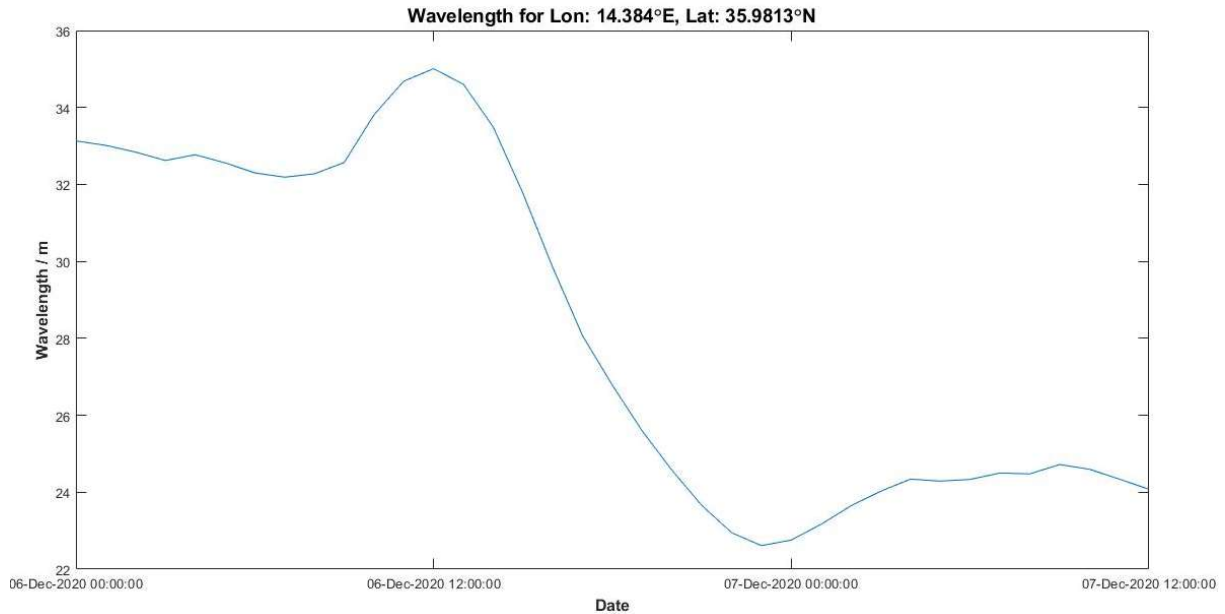
Figure 46: Forecasted Wave Length at operator specified grid point

5.2.8 Comparison between h_{max3hr} and H_s Values

Whilst the difference in definition between Individual and Significant Wave Height is understood, it is interesting to quantify the potential difference between these two values. The plots of the forecasted H_s (Figure 47) and h_{max3hr} (Figure 48) for the 6th of December 2020 were compared. The difference in values were also plotted (Figure 49).
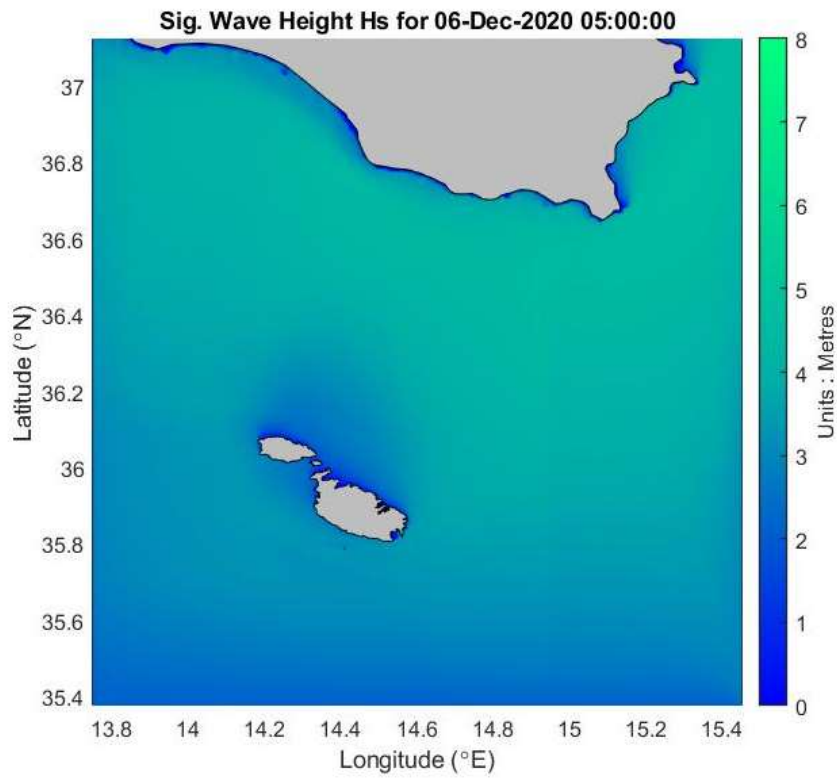
Figure 47: Significant Wave Height H$_s$ for 06 December 2020; 05:00 am
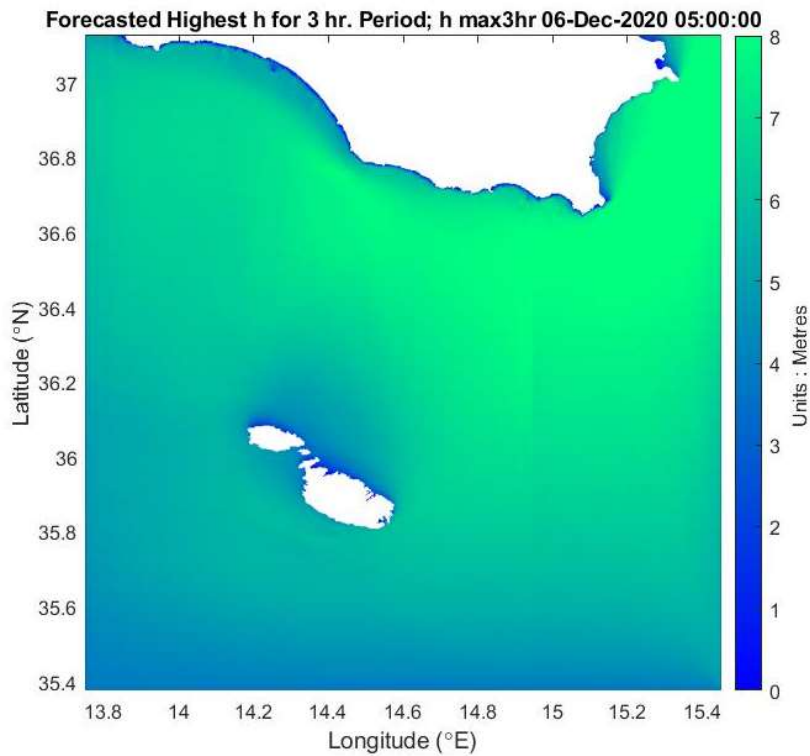
Figure 48: Forecasted Highest Individual Wave Height h $_{max\ 3hr}$ for 06 December 2020; 05:00 am
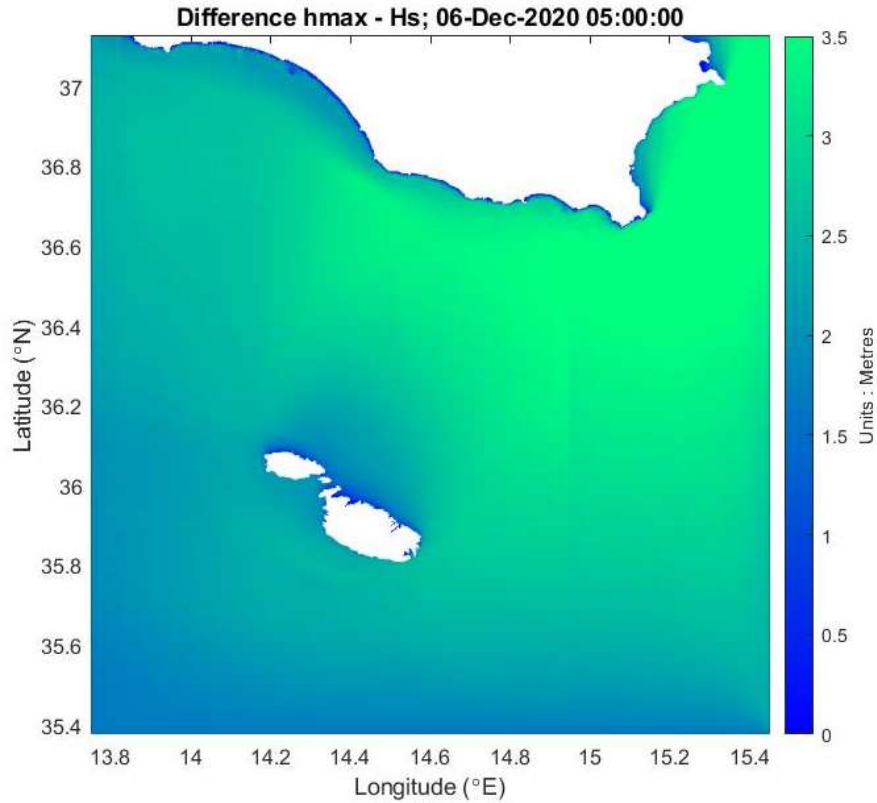
Figure 49: Difference between H_s and $h_{max\,3hr}$ for 06 December 2020; 05:00 am

Whilst the highest forecasted value of H_s was of 4.5 m occurring at the south / south east coast of Siciliy, the highest predicted value of $h_{max\,3hr}$ was around 8.0 m. The maximum gap between both sets of values being around 3.5 m that is $h_{max\,3hr}$ being approximately 70% higher than the H_s value. Generally, weather forecasts only provide information on H_s. It should be noted , that this large diffence in values between H_s and $h_{max\,3hr}$ in the region of 3.5 m will not always occur.

Chapter 6 : Results Extreme Value Analyses

In this chapter, the EVA for the grid point off the coast of Dwejra was carried out. The forecasted 100-year return values of H_s, T_z and h were found. The following procedure was adopted when carrying out the Peak Over Threshold (POT) method of peak data selection and Extreme Value Analyses (EVA):

1. Peripheral months with low number of peak H_s values were removed from H_s time series;

2. The H_s Upper Threshold Value: A range of values between 2.5 m and 4.5 m were used as the Upper Threshold Value. The highest H_s value within each independent extreme event and the number of extreme events were found. Other data belonging to these peaks were also determined;

3. Three distributions (Weibull, GPD and Gumbell) were fitted to these peak H_s values;

4. The *N-year* return periods and the forecasted 100-year return H_s values were determined for each upper threshold value and for each distribution. From the data attained, the upper threshold value and the distribution giving the most realistic values were determined;

5. Having found the forecasted 100-year H_s value and using the Rayleigh distribution, the 100-year return h value was forecasted. Also, the *100-year* return h value was forecasted using the Battjes method. The values attained with both methods where compared.

6.1 Determining the Peripheral Months

As discussed in the literature review, peripheral months with very low values of H_s impact the EVA and should be removed. The peripheral months were determined based on the number of peak H_s values exceeding a threshold per month. Months exhibiting low counts were removed.

The percentile values were calculated using the plot of probability of exceedance against H_s (Figure 47). The values are presented in Table 17. The 95-percentile value was subsequently used as the threshold value together with a two-day interarrival time.

Percentile Value	H_s /m
95	2.85
90	2.23
85	1.87

Table 17: Percentile Values- H_s Full Time Series
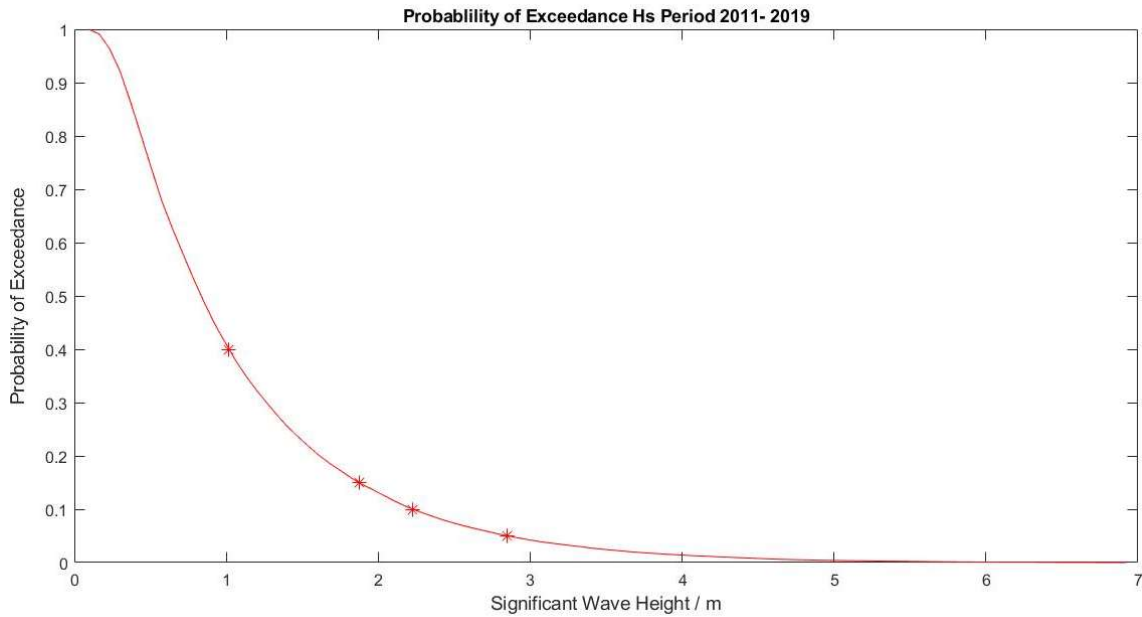
Figure 50: Probability of Exceedance / Full H_s time series to determine percentile values

Whilst some months had an extremely low number of peaks making it easy to remove, others were not so clear (Figure 51). The threshold value was hence increased from 3.0 m to 4.0 m and again the peaks were sorted by month (Figure 52).
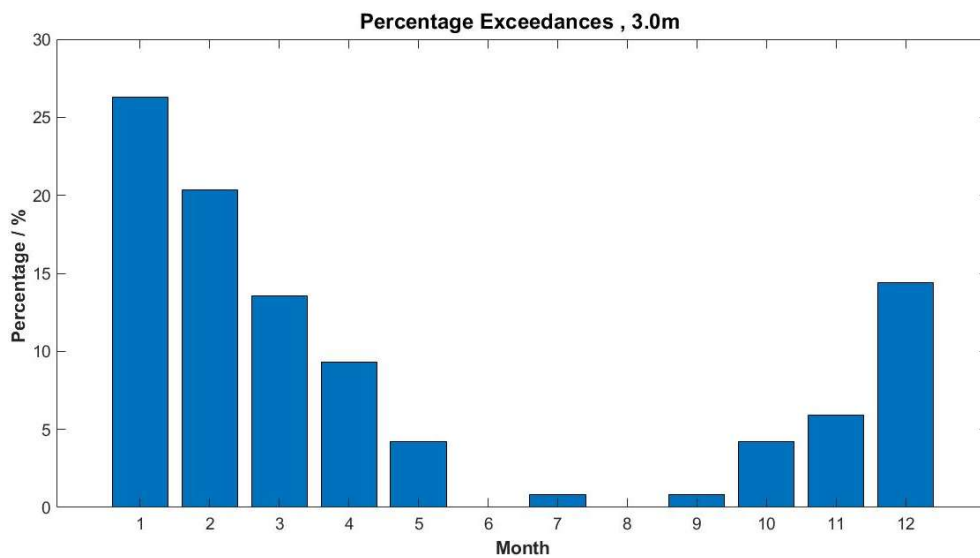
Figure 51: Percentage Exceedance of the 95-percentile value / month
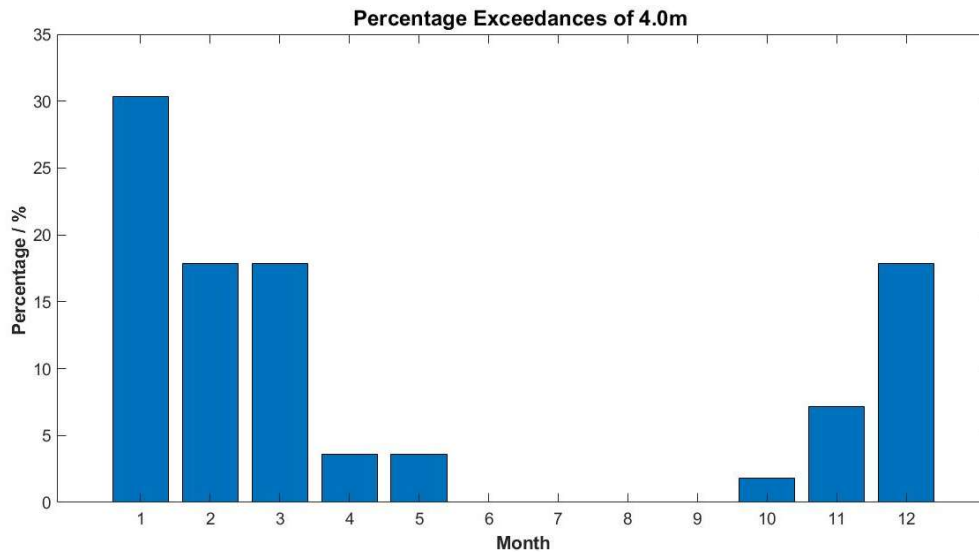
Figure 52: Percentage Exceedance of per month H_s = 4 m

On increasing the threshold value, the percentage occurrence for the month of April decreased from nearly 10% to under 5 % indicating that during this month most peaks had a value of between 3 and 4 m. It was decided to include this month with the peripheral months.

The period between April and October (both months included) were classified as peripheral months and only the data for the months between November and March (both months included) were used for the EVA.

6.2 Determining the Independent Extreme Events

As discussed in Section 4.2.6, independent extreme events were chosen based on the following

1. Significant Wave Height H_s exceeding an Upper Threshold Value;
2. Interarrival Time. A minimum time period of 1 day was set for peaks to be described as independent events;
3. Lower threshold value. The Significant Wave Height H_s value must decrease to a value under a lower threshold value in between storms for them to be considered independent.

The highest peak from each storm is subsequently chosen.

The H_s time series based on data generated by the *ROSARIO MARIA SWAN* model was plotted (Figure 53). The peripheral months were removed from the time series (Figure 54). Other wave parameters such as wave steepness and wave lengths were calculated using the T_z values generated by the *ROSARIO MARIA SWAN* model and adjusted using the correction factors derived during validation.
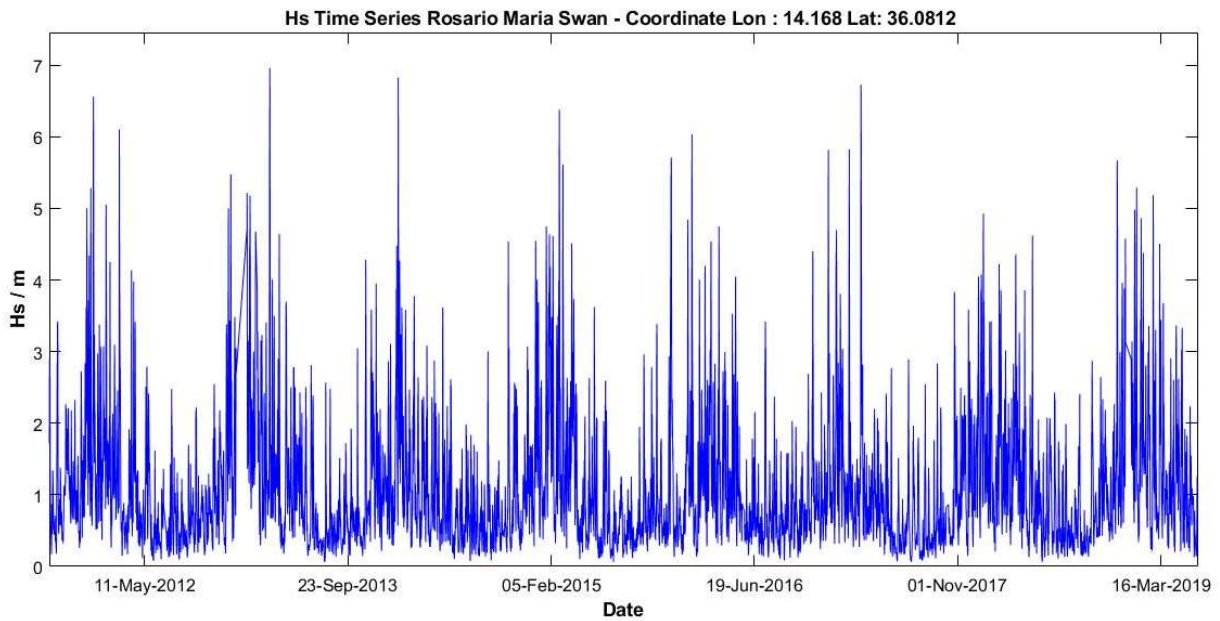
Figure 53: Model time series of H_s Period 2011-2019 for Lon 14.168 deg. and Lat. 36.0812 deg.
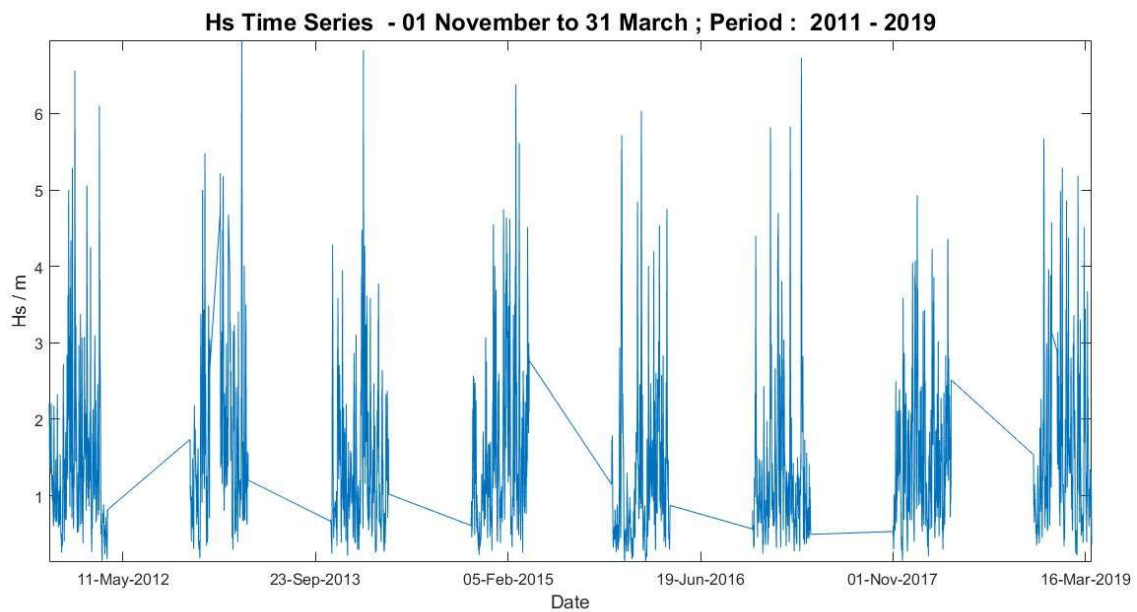
Figure 54: *Removal of the peripheral months from the corrected H_s timeseries*

The percentile values of the H_s time series excluding the peripheral months were found (Figure 55; Table 18). As previously discussed, the lower threshold value was assigned the 60-percentile value whilst the upper threshold value was given different values in range 2.5 m to 4.5 m.

It's worth noting that the percentile values for the reduced H_s time series increased by over 0.5m as compared to the values for the full H_s time series.
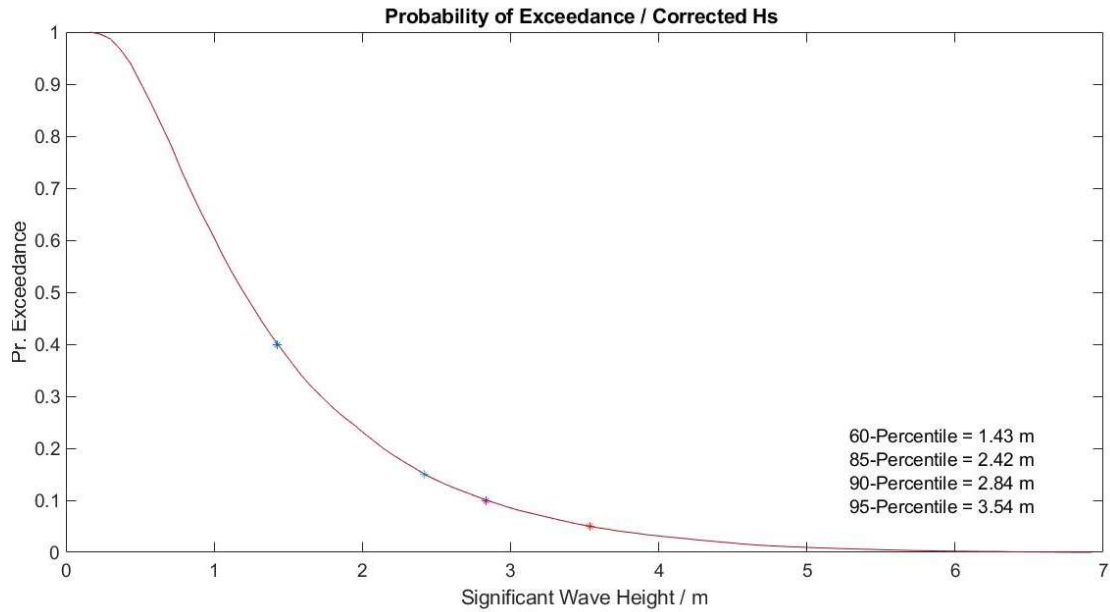
Figure 55: Pr. of Exceedance plot to determining the Percentile Values for the reduced time series

Percentile value	H_s /m
60	1.43
85	2.42
90	2.84
95	3.54

Table 18 : Percentile values on the reduced time series

The time series was inverted about the 60-percentile value to help identify the peak values under the lower threshold value (Figure 56 and Figure 57).
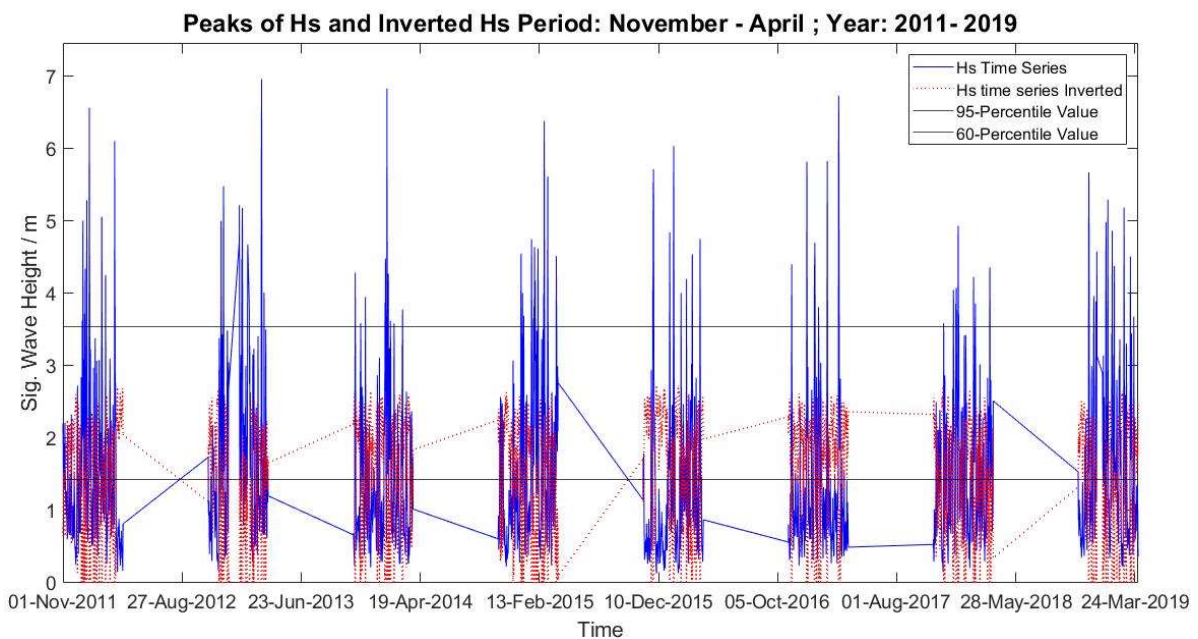
Figure 56: The H_s time series and the inverted time series for period 2011-2019
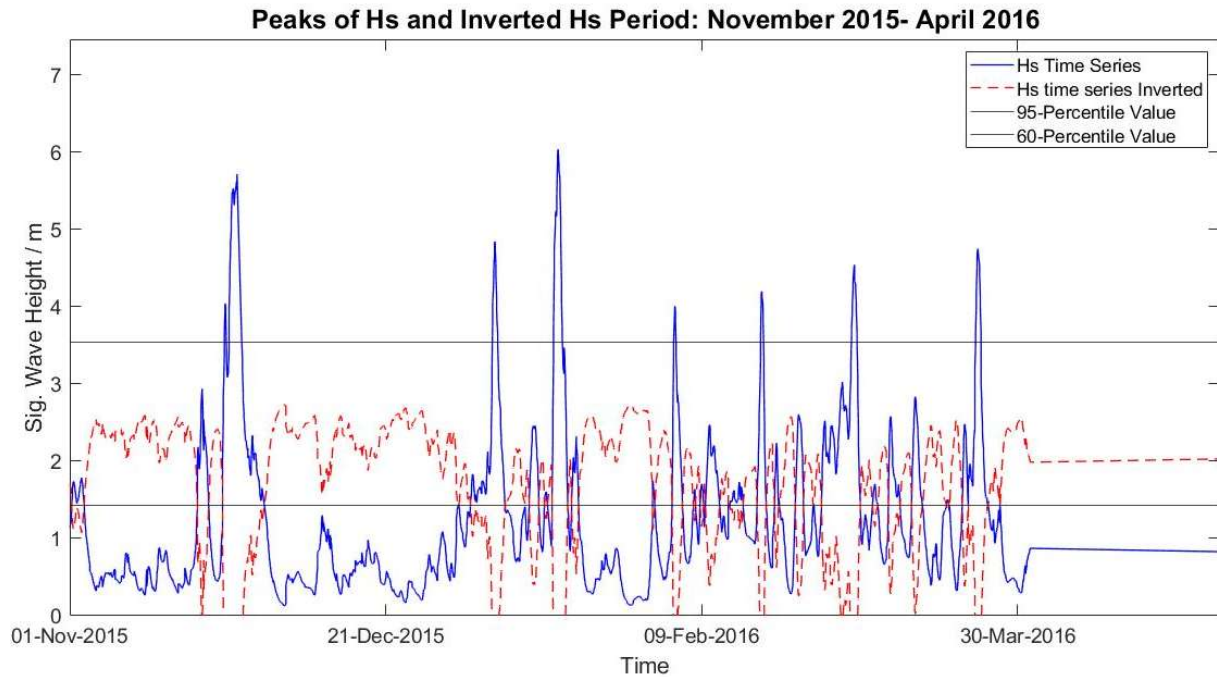
Figure 57: The time series and inverted time series for period Nov 2015 and April 2016

MATLAB was used to identify the peaks of the time series exceeding the upper threshold value and meeting the set interarrival time. It was also used to find the peaks of the inverted time series about the 60-percentile value.

Figures (58 and 59) indicate the H_s peaks selected based on the controlling parameters mentioned above. The peaks marked in red indicate the peaks that exceed the upper threshold value and satisfy the interarrival time. The peaks on the inverted time series are marked in green. These correspond to the peaks that fall under the lower threshold value. The highest peak of the independent extreme events are marked with the black circle.

So far the H_s correction factor based on the entire time series was used. But when carrying out EVA using the POT method, only H_s values exceeding a threshold value were used. From the validation section, a correction factor of 1.05 was found to give closer results to actual H_s values in excess of 2.3 m. Hence the model peak H_s data were adjusted by using the 1.05 correction factor.
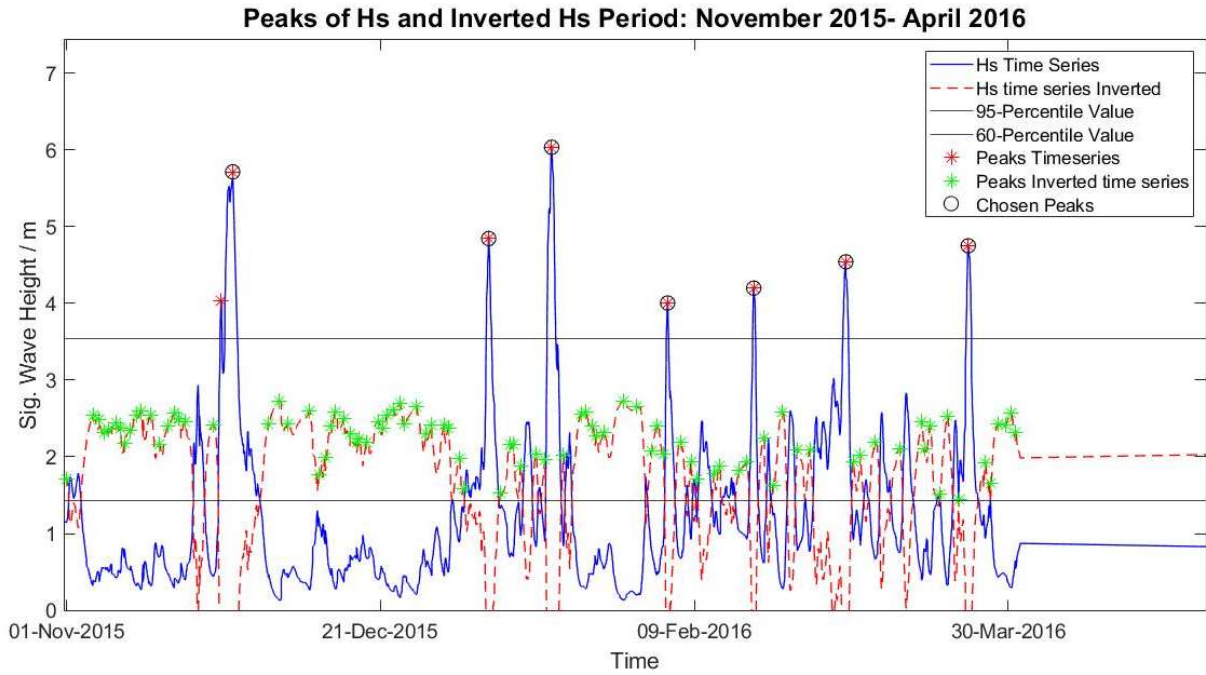
Figure 58: Selecting the peaks of the H_s time series, the Inverted time series and the peak of each storm for period 2015/2016
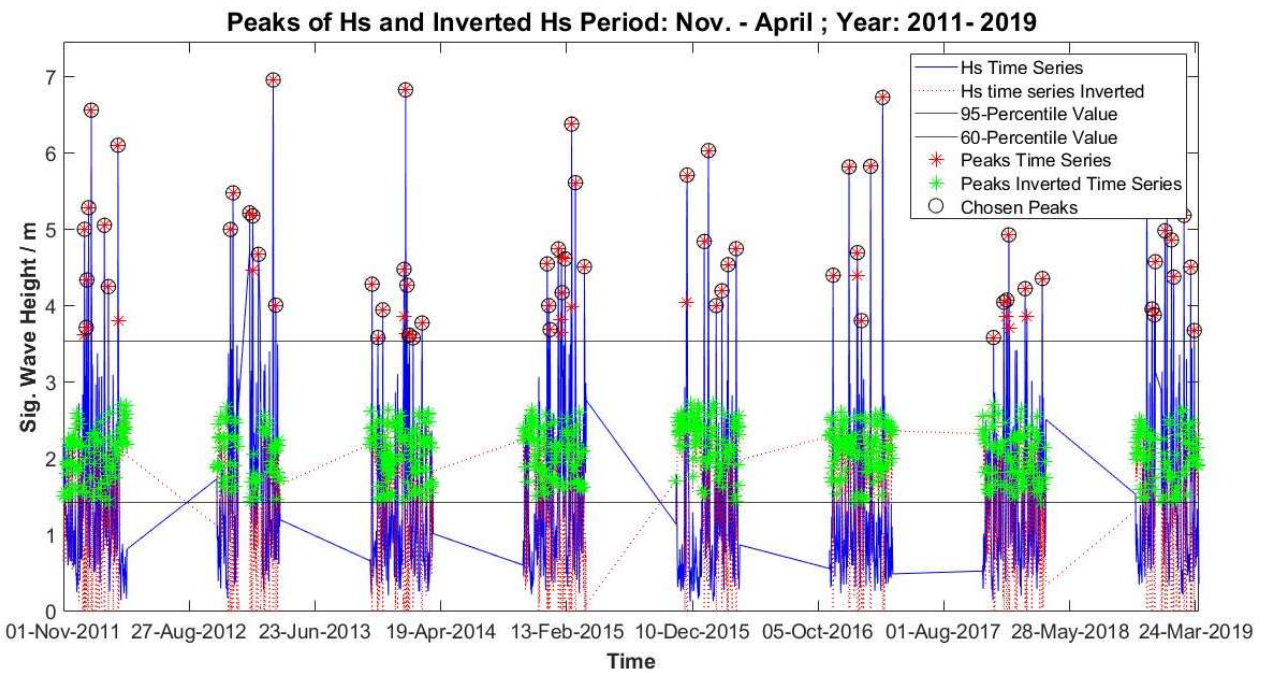
Figure 59: Selecting the peaks of the H_s and Inverted time series and the peak of each storm

6.3 Determining the Upper Threshold Value

There are no set rules for determining the upper threshold value, yet this value will affect the end results from the EVA. The Extreme Value analyses was carried out for various threshold values in close proximity to the 95-percentile value of 3.5 m.

The upper threshold value was selected based on the total number of extreme peaks found since a number of peaks are required for better statistical accuracy of the EVA. In this case having a relatively short time series, this might mean that peak data from not as extreme wave conditions were also considered in the EVA.

The process of finding the peak H_s values and the subsequent fitting of the three distributions was repeated for each threshold values. *MATLAB* was used to fit the Gumbell, the Weibull and the Generalised Pareto Distributions to these peak H_s values by using the Maximum Likelihood Method in order to determine the 100-year return value of H_s. The same procedure can be followed should the forecasted H_s value for a different time period be required.

To determine the distribution that fits most closely to the actual data, a visual inspection of the plots and the Correlation Coefficient, *r,* were used. But since the data is only available for a short period of time (8 years) as compared to the forecasted period (100 years), the interpretation of *r* is very subjective. Alternative methods of comparison as listed in Section 3.2.2.3 should be looked into. This together with a longer time series should improve the outcome of the analyses.

Distribution Parameters for different threshold values

Threshold Value /m	Gumbell Distribution		Weibull Distribution		GP Distribution	
	Parameter	Value	Parameter	Value	Parameter	Value
2.5	U	3.6112	α	1.8132	Shape	-0.4362
	α	0.9249	β	1.4183	Scale	2.3677
3.0	U	4.0917	α	1.7513	Shape	-0.4958
	α	0.8388	β	1.5097	Scale	2.3568
3.5	U	4.5863	α	1.6977	Shape	-0.5759
	α	0.7351	β	1.6874	Scale	2.3670
4.0	U	4.9646	α	1.5139	Shape	-0.6016
	α	0.6575	β	1.7017	Scale	2.1574
4.25	U	5.1457	α	1.4159	Shape	-0.6155
	α	0.6230	β	1.6964	Scale	2.0465
4.5	U	5.3264	α	1.3188	Shape	-0.6460
	α	0.5986	β	1.6449	Scale	1.9706

Table 19: Distributions Parameters based on different threshold values

6.4 Plots of the Return Periods

Figures (60 – 64) represent the plots of the Return Period based on different Upper Threshold Values of 2.5 m, 3.0 m, 3.5 m, 4.25 m and 4.5 m respectively.
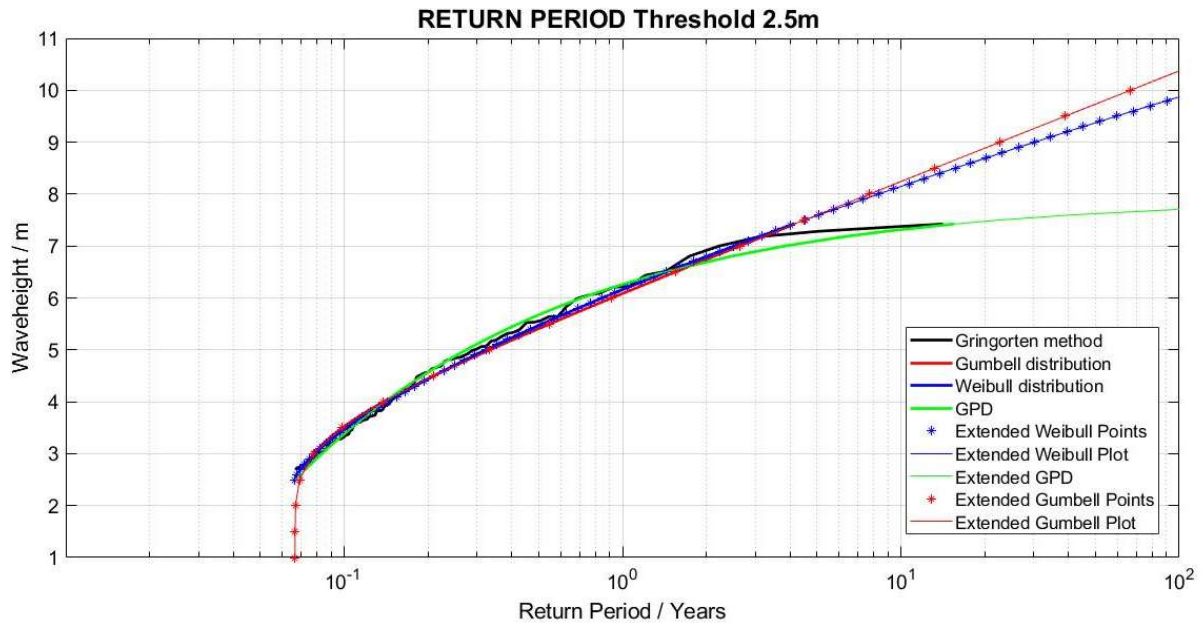
Figure 60: Return Period based on H_s threshold 2.5 m
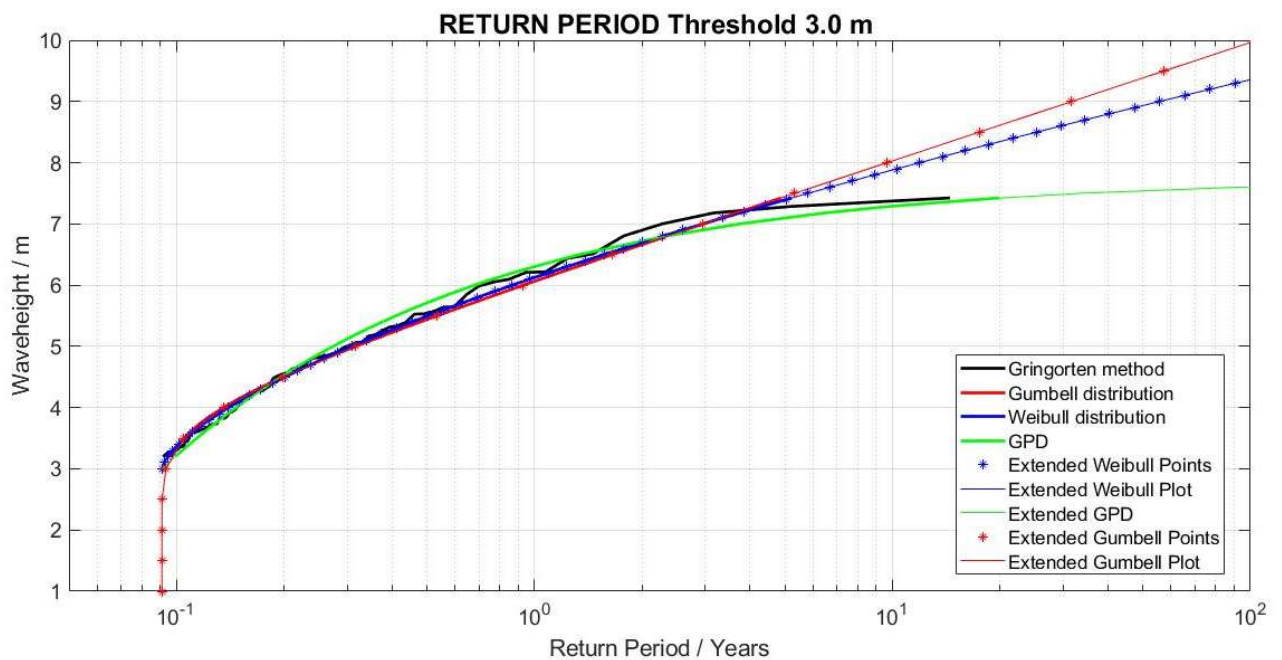
Figure 61: Return Period based on H_s threshold 3.0 m
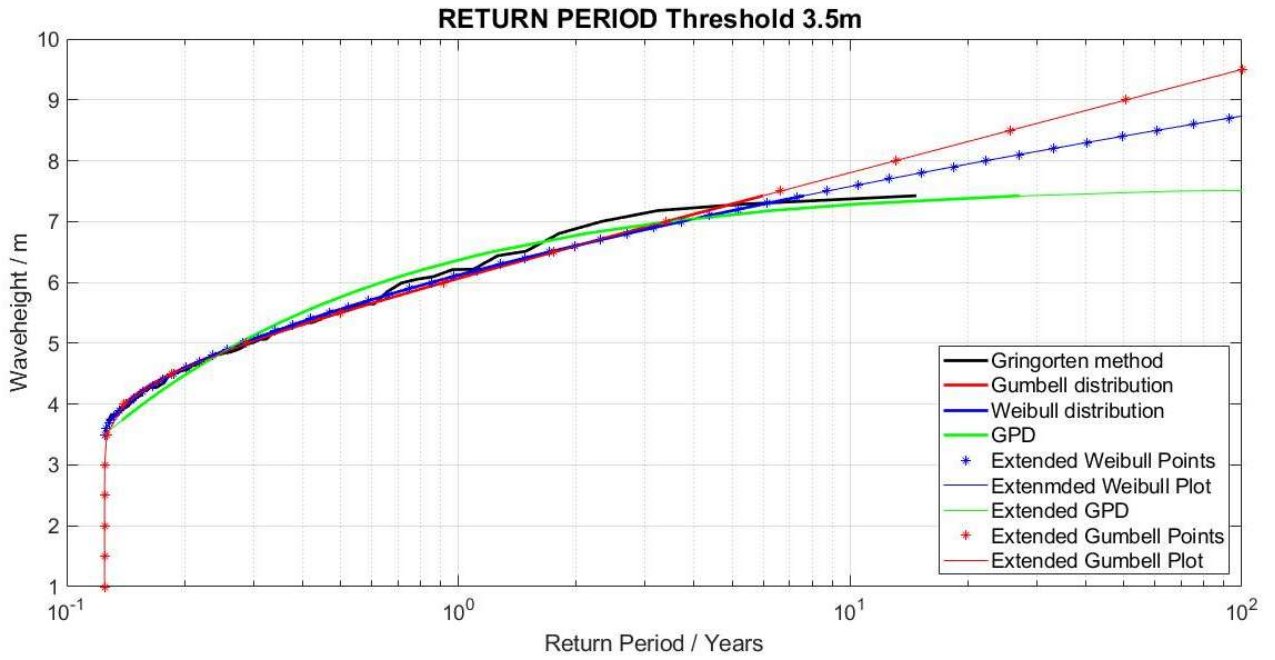
Figure 62: Return Period based on H$_s$ threshold 3.5 m
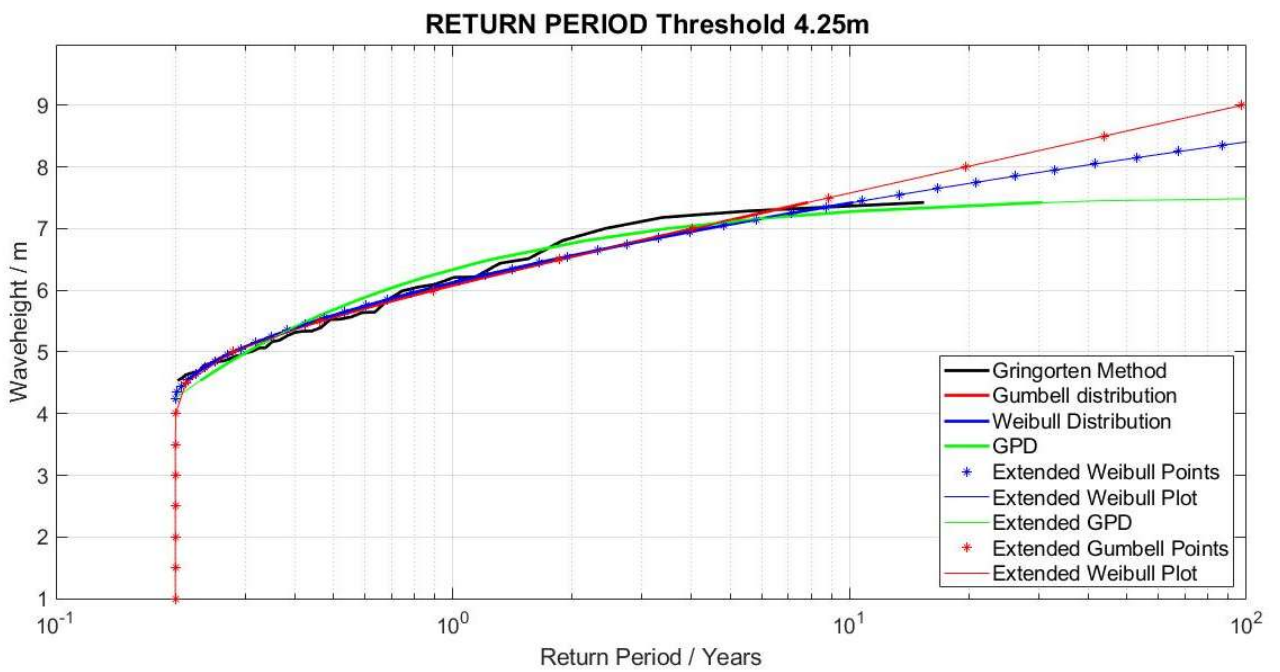
Figure 63: Return Period based on H$_s$ threshold 4.25 m
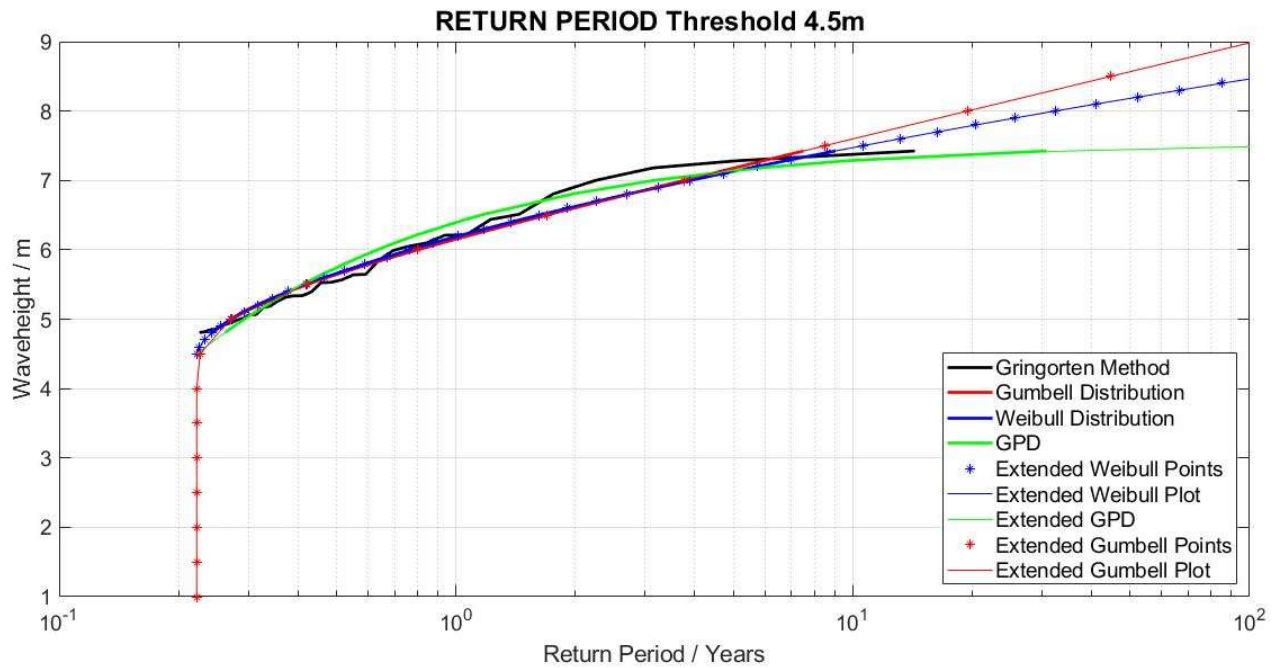
Figure 64: Return Period based on H_s threshold 4.5 m

6.5 The *100-year* Return Periods with different threshold values

The total number of peaks, number of peaks per year and the *100-year* Return Values for different threshold values and distributions are found in Table (20).

Threshold setting /m	Number of Peaks Total	Nr. of peaks per year (8 years)	Gumbell / m	Weibull / m	GPD / m
2.5	119	15	10.5	9.75	7.8
3.0	89	11	10.0	9.5	7.6
3.5	63	8	9.5	8.8	7.5
4.0	50	6.25	9.2	8.5	7.5
4.25	43	5	9.0	8.5	7.5
4.5	36	4.5	9.0	8.5	7.5

Table 20: The 100-year return period based on different H_s threshold values

Based on the total number of peaks over the time period and the yearly number of peaks, the most viable range of threshold values is between 3.5 m and 4.25 m. When changing the upper threshold value upwards of 4.0 m, the changes in the forecasted values is minimal for all 3 distributions.

Yet there is a 0.30 m difference between the 100-year return values for both the Gumbell and Weibull distribution when increasing the threshold value from 3.5 m to 4.0 m. For the GPD, for threshold values of 3.5 m and above, the 100-year return value remained at 7.5 m.

Table (21) is showing the values of the Cumulative Distribution and Probability of Exceedance for the highest hindcast values. For the highest H_s value registered (7.42 m), the CD value for the Gumbel and Weibull distribution is 0.98. This means that this H_s value will be exceeded for longer time periods. But for the GPD, the CD value is 1.00 indicating that this is the highest H_s value that can be forecasted by this distribution irrespective of the period of time.

Ranked H_s Value	Rank	Gumbell		Weibull		GPD	
		Exc.	CDF	Exc.	CDF	Exc.	CDF
7.42	1	0.02	0.98	0.02	0.98	0.00	1.00
7.29	2	0.03	0.97	0.02	0.98	0.01	0.99

Table 21: CDF and Exceedance probability for highest registered H_s value

As previously discussed in section 4.2.5 and in accordance with recommendations by *Matthiesen et. al*, the upper threshold value was set based on the number of peaks selected from the available time series. Based on this, it was decided to set the upper threshold value at 3.5 m giving 60 peaks, fitting within the 50 and 75 peaks originally targeted. This value coincides also with the 95-percentile value of the H_s timeseries after excluding the peripheral months. Setting the upper threshold value as the 95-percentile value is also widely adopted in other research studies (24, 33).

Figure (65) and figure (66) show the Probability Distribution and the Cumulative Distribution Function plots based on an H_s upper threshold value of 3.5 m respectively.
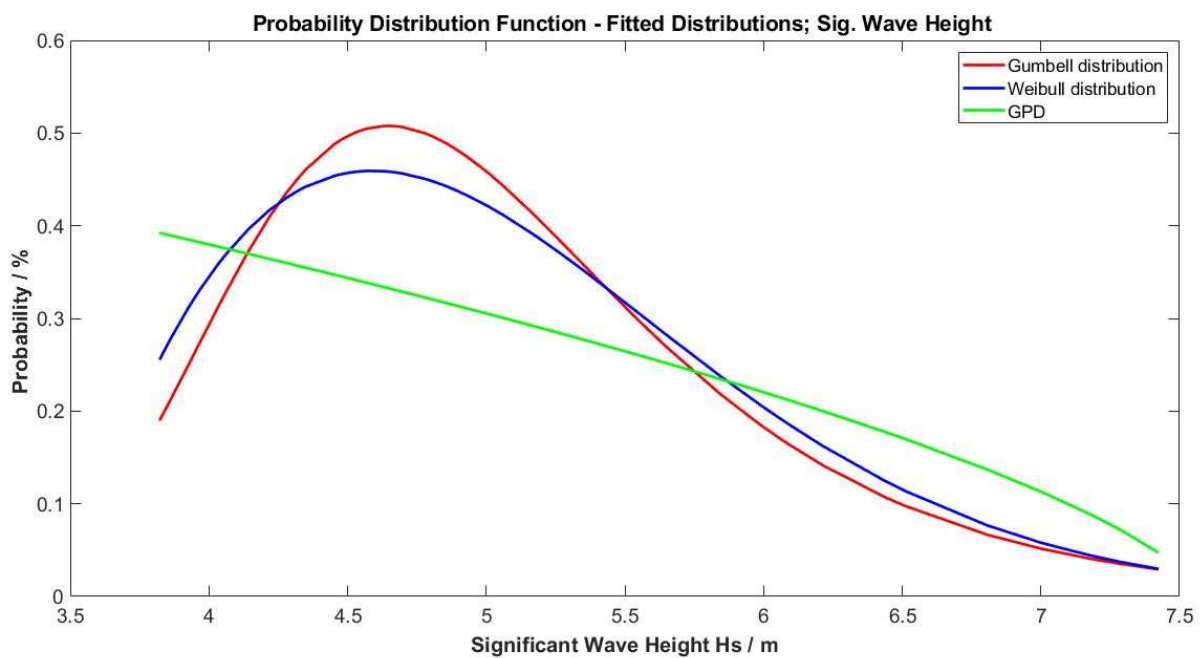
Figure 65: The H_s Probability Distribution Functions for threshold value of 3.5 m
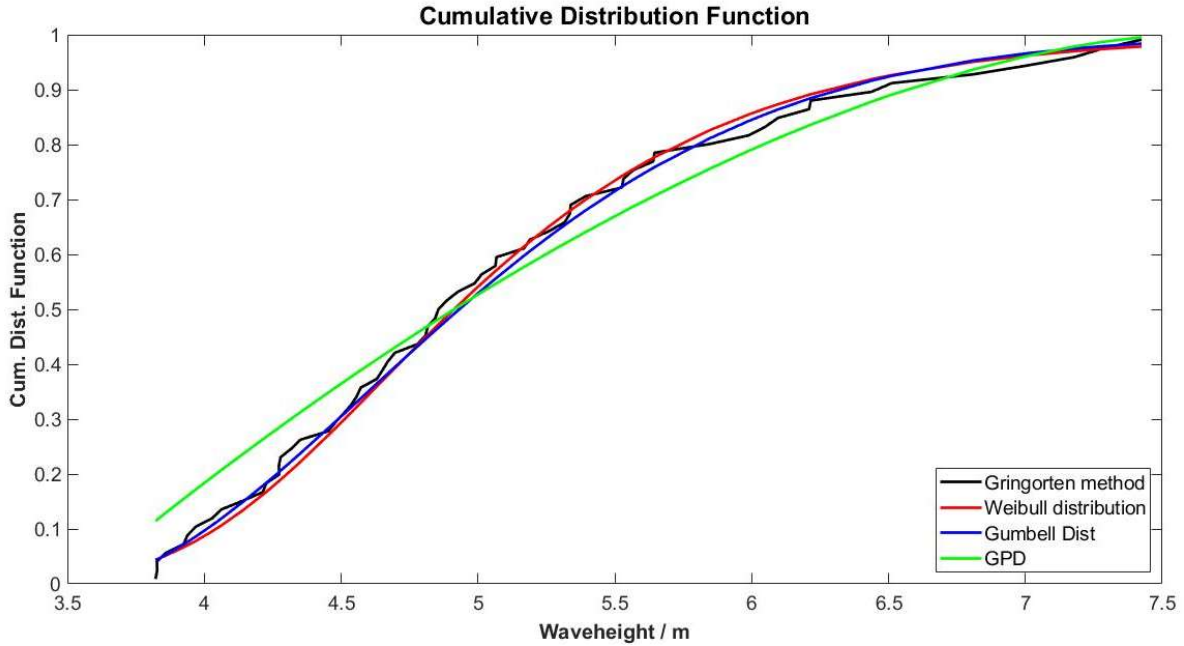
Figure 66: The Cumulative Distribution Function, Weibull, Gumbell, GPD with Actual Data

The Cumulative plots indicates a good closeness of fit between the actual data and the distributions with the GPD showing the most variation.

6.5 Selecting the Best Fitting Distribution

The distribution fitting closest to the actual data needed to be selected. As discussed, both a visual inspection of the plots and the correlation coefficient, *r*, were used. The correlation coefficient , *r*, was determined by comparing the wave heights (hindcast data and from the individual distributions) which have the same probability of exceedance. Figure (67) shows the plots of wave height against Probability of Exceedance for the actual data and for the three distributions. The Probability of Exceedance was given a range of values from 0.05 to 0.95 increasing in steps of 0.05. The respective H_s values are found. Scatter plots for the three distributions were produced and the linear regression lines were fitted (Figures 68- 70). The linear correlation coefficient , *r*, for each distribution was determined.

	r Gum	r Wei	r Gpd
r Value	0.9983	0.9959	0.98458

Table 22: Linear Correlation Coefficient r – Gumbell, Weibull and GPD
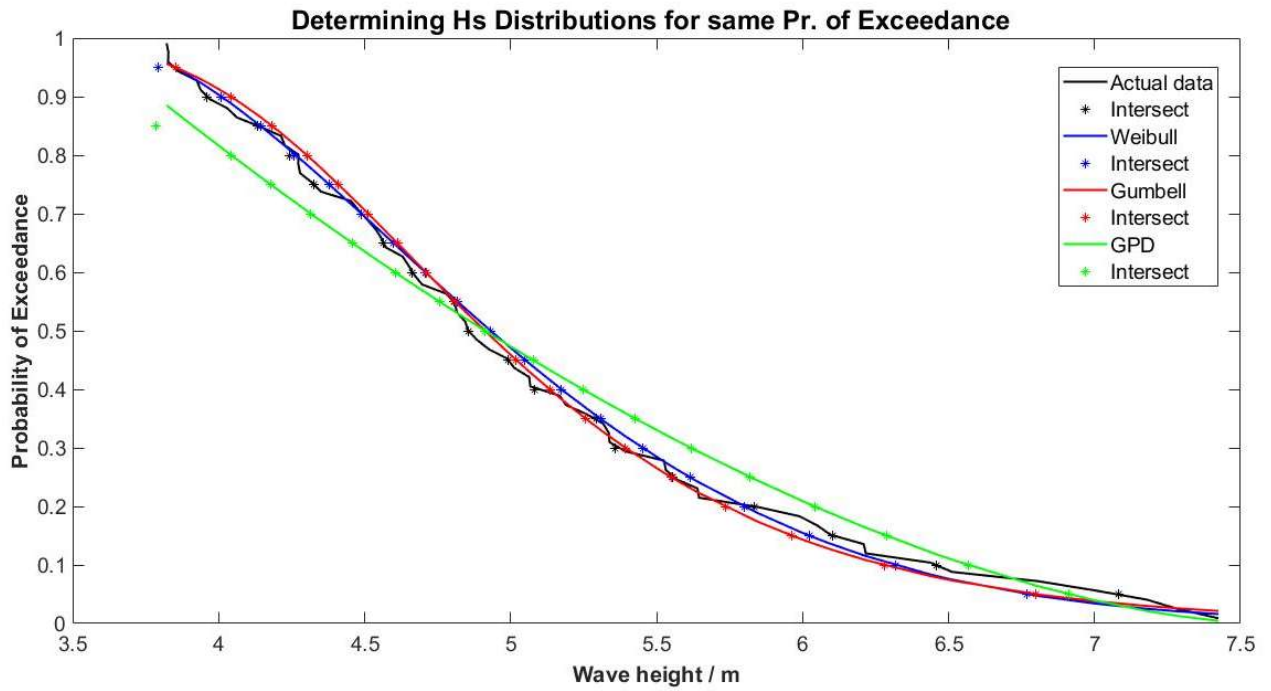
Figure 67: Probability of Exceedance - Weibull, Gumbell and GPD with Actual data
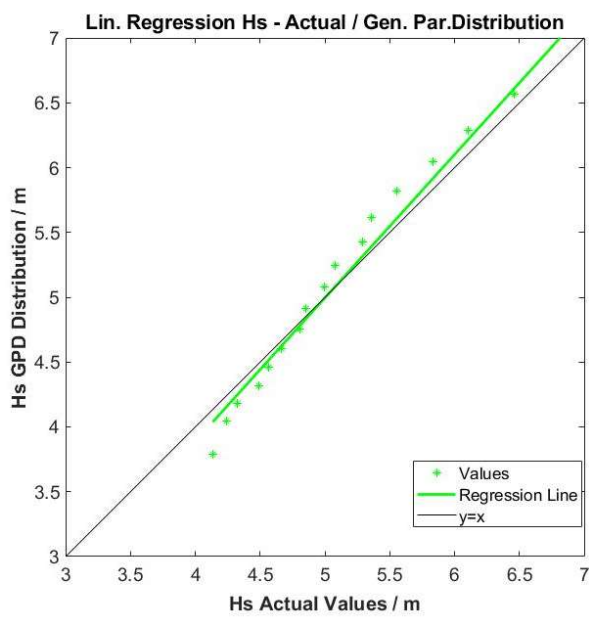
Figure 68: Linear Regression line – GPD
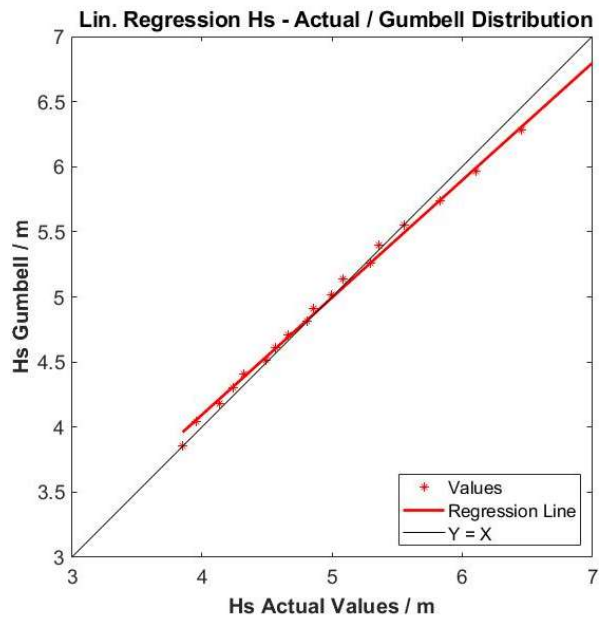
Figure 69: Linear Regression line – Gumbell
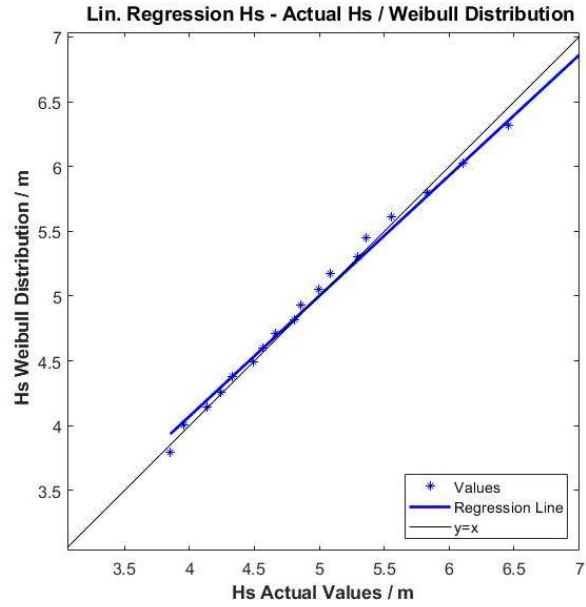
Figure 70: Linear Regression line – Weibull
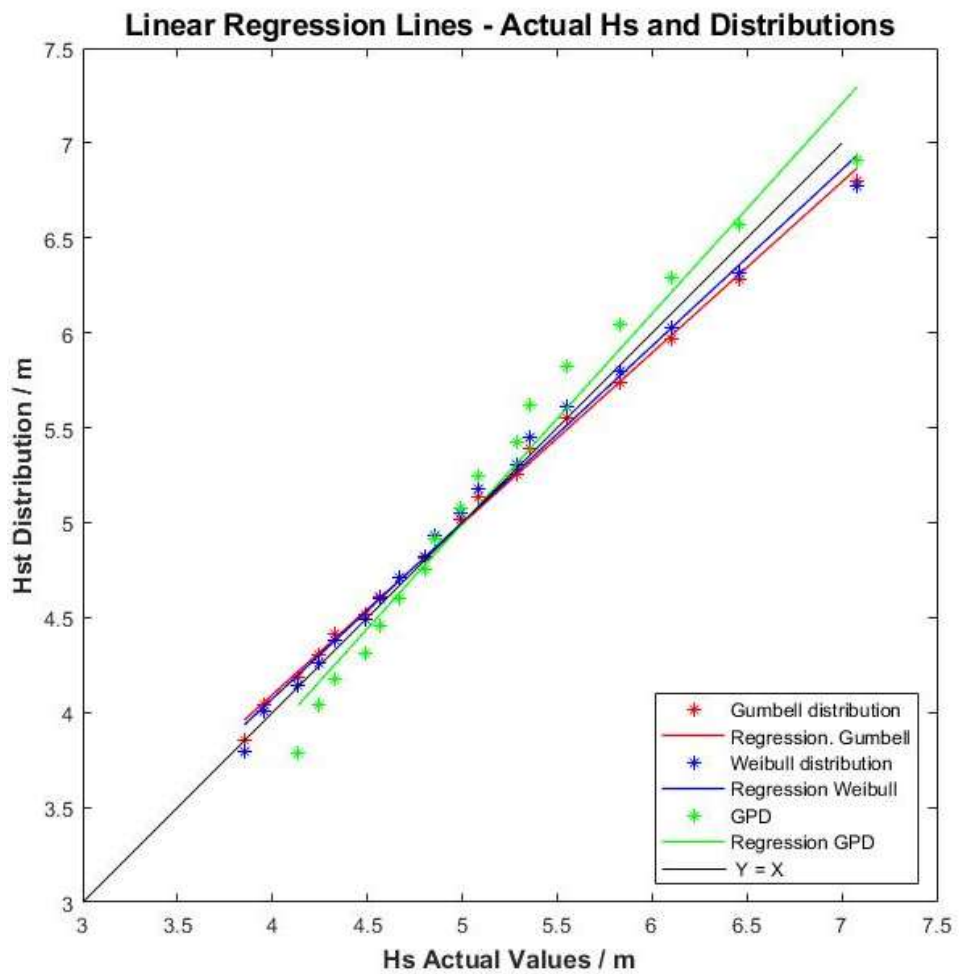
Figure 71: The three Regression lines superimposed with the ideal Regression Line

The values of the correlation coefficient for the three distributions are very high. From the visual inspection of the Probability of Exceedance against H_s plot, Figure (67), the Gumbell and the Weibull distribution have a closer fit to the actual data as compared to the GPD. This is also confirmed by the higher correlation coefficient r value of the two as compared to that of the GPD.

Threshold Value / m	Number of Peaks Total	Nr. of peaks per year (8 years)	Gumbell / m	Weibull / m	Difference / m
3.5	63	8	9.5	8.8	0.70
4.0	50	6.25	9.2	8.5	0.70
4.25	43	5	9.0	8.5	0.50
4.5	36	4.5	9.0	8.5	0.50

Table 23: the 100-year return values Gumbell and Weibull distributions at different Threshold values

The average difference between the forecasted 100-year return values using the Gumbell and Weibull distribution for different threshold values, is around 0.60 m or approximately 7.5 %.

On observing the graph with the 3 regression lines plotted together (Figure 68) the regression line of the Weibull distribution is closer to the ideal regression line (Y= X) as compared to the Gumbell. This indicates that the values obtained using the Weibull distribution are closer to the model values.

Based on the closer fit of the regression line to the ideal line, and the fact that the Weibull distribution is the more widely accepted distribution when using the POT method, the Weibull distribution was chosen over the Gumbell.

Hence, based on the use of the Weibull distribution and an H_s upper threshold value of 3.5 m, the forecasted H_s value with a 100-year return period for the location given by coordinates 14.16888º longitude and 36.08152º latitude is estimated to be 8.8 m.
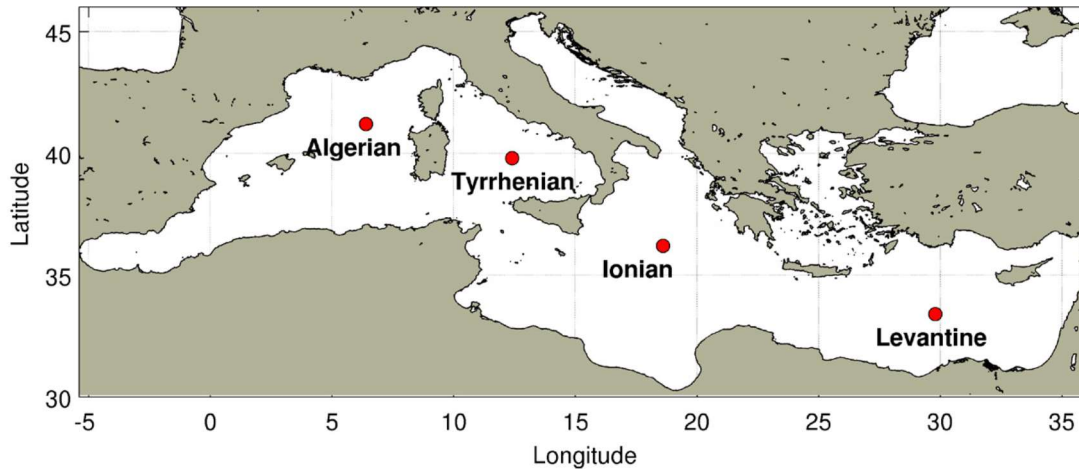
Figure 72: (This figure was used previously) Map showing the areas analysed by Kapelonis et al.; Reference (31)

The study carried out by *Kapelonis et. al*. produced the forecasted 100-year return H_s values for other areas in the Mediterranean based on data covering a 30-year period and using the POT method of analyses. These values and the value obtained by this study for Malta are as follows:

Location	Threshold Value / m	H_{s100} / m
Malta	3.5	8.8
Algerian*	4.5	11.4
Tyrrhenian*	4.5	8.4
Ionian*	4.2	8.5
Levantine*	3.7	8.6

*Table 24: The 100-year Hs Return Values obtained for Malta and other zones in the Mediterranean; * Reference (31)*

6.6 Wave Data at Extreme Events

Having set the upper threshold value at 3.5m, a total of 63 peaks were found during the period between September 2011 and June 2019, an average of nearly 8 storms per year. So far, the focus was primarily on the values of the Significant Wave Height. Using the data generated by the model and the linear wave theory, other wave parameters occurring during these extreme events can be calculated, collected and analysed. The values of the wave parameters related to these extreme events can be found in Appendix (E) while a summary of the results can be found hereunder.

1. Months exhibiting most extreme events

Month	Nr. of peaks	Per. %
January	18	28.5
December	14	22.2
February	13	20
March	12	19
November	6	9.5
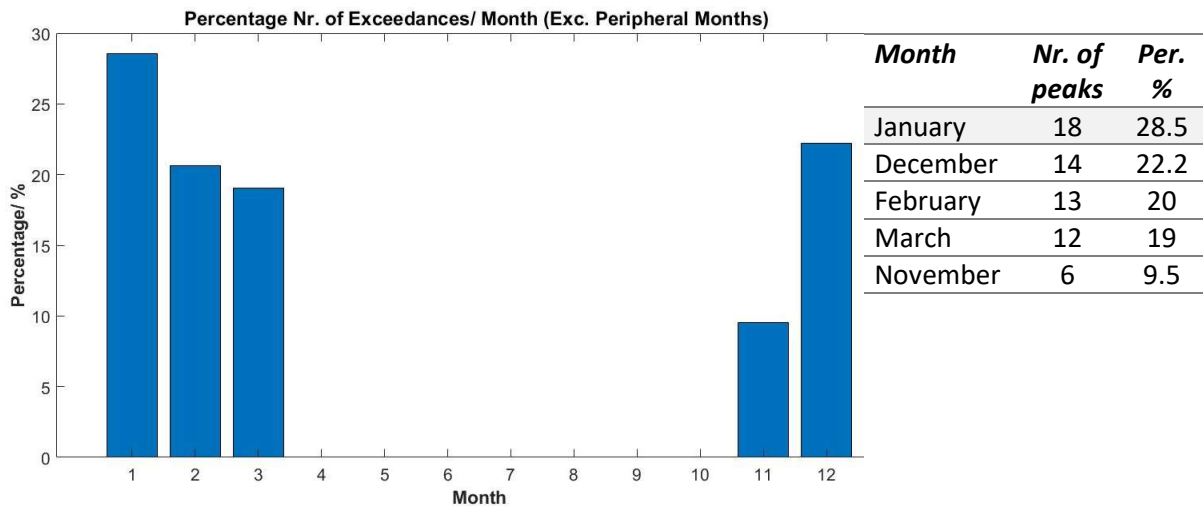
Figure 73: Month / Percentage Nr. of Occurances of Peak Events

2. Range of H_s Values

Nom. Hs	Range Hs	Nr. of peaks	Per. %
3.5	3.25 - 3.75	7	11.1
4.0	3.75 - 4.25	13	20.6
4.5	4.25 – 4.75	18	28.5
5.0	4.75 – 5.25	10	15.8
5.5	5.25 – 5.75	6	9.5
6.0	5.75 – 6.25	4	6.3
6.5	6.25 – 6.75	3	4.7
7.0	6.75 – 7.25	2	3.1
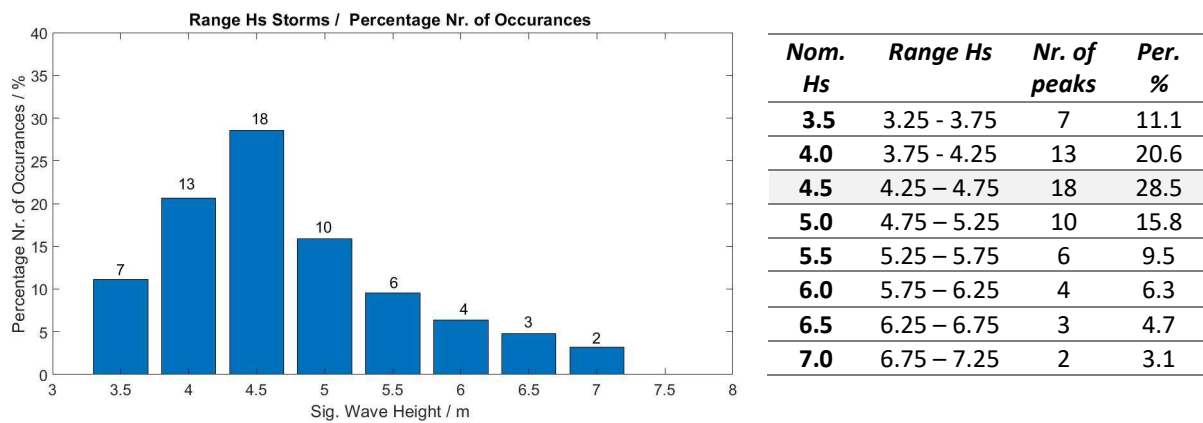
Figure 74: H_s Values / Percentage Nr. of Occurances of Peak Events

3. Wave Direction

Nominal	Range	Nr. of peaks	Per. %
45	22.5 – 67.5	7	11.1
225	202.5 – 247.5	1	1.6
270	247.5 – 292.5	41	65.1
315	292.5 – 337.5	14	22.2
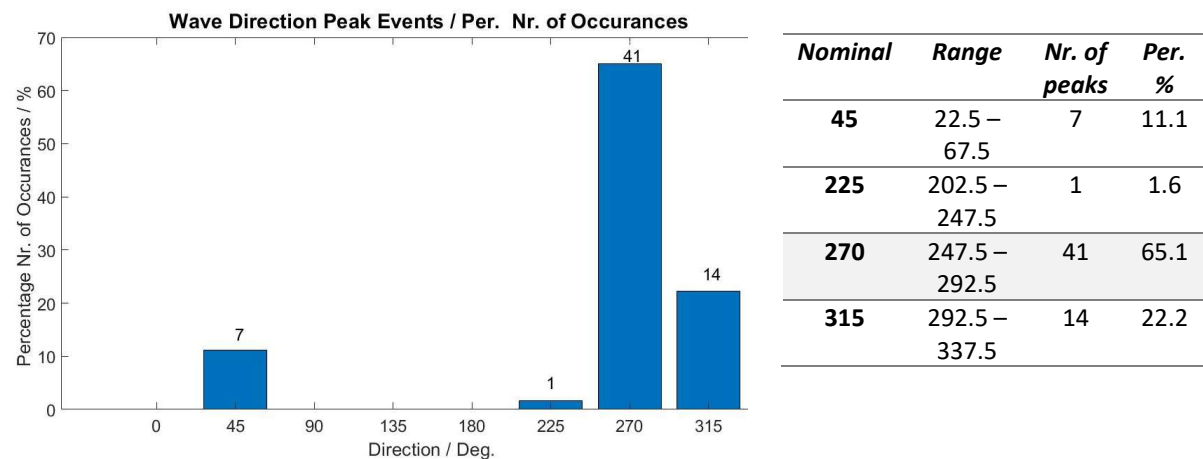
Figure 75: Wave direction / Percentage Nr. of Occurances of Peak Events

The wave direction with the highest number of extreme events was 247.5º– 292.5º that is the bin centred at 270º (65%), followed by 315º (22.2%) and the 45º (11.1%) respectively.

This study is based on the model analyses at the grid location of 14.168^0 longitude and $36.0812º$ latitude, off Dwejra to the east of Gozo. This means this area is mostly exposed to the Westerly winds whilst being protected by the island of Gozo from the Easterly winds.

4. H_s analyses for Waves heading 270 deg.

Nom. H_s	Range H_s	Nr. of peaks	Per. %
3.5	3.25 - 3.75	4	6.3
4.0	3.75 - 4.25	6	9.5
4.5	4.25 – 4.75	13	20.6
5.0	4.75 – 5.25	8	12.7
5.5	5.25 – 5.75	4	6.3
6.0	5.75 – 6.25	1	1.6
6.5	6.25 – 6.75	3	4.8
7.0	6.75 – 7.25	2	3.2
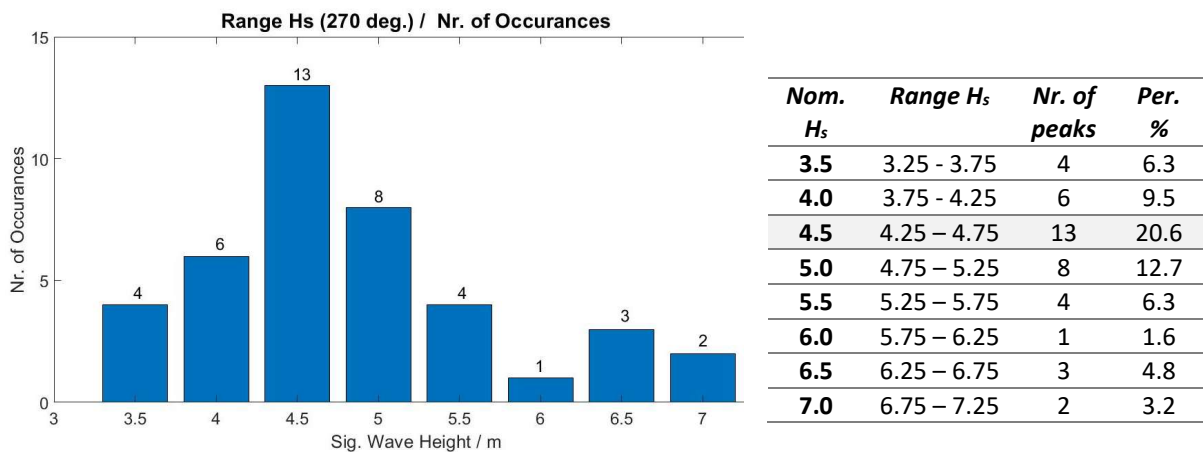
Figure 76: H_s range / Percentage Nr. of Occurances at 270 deg. Wave direction

5. Mean Up-cross period T_z:

Nom.	Range	Nr. of peaks	Per. %
5.5	5.25 - 5.75	4	6.3
6.0	5.75 - 6.25	23	36.5
6.5	6.25 – 6.75	24	38.1
7.0	6.75 – 7.25	8	12.7
7.5	7.25 – 7.75	4	6.3
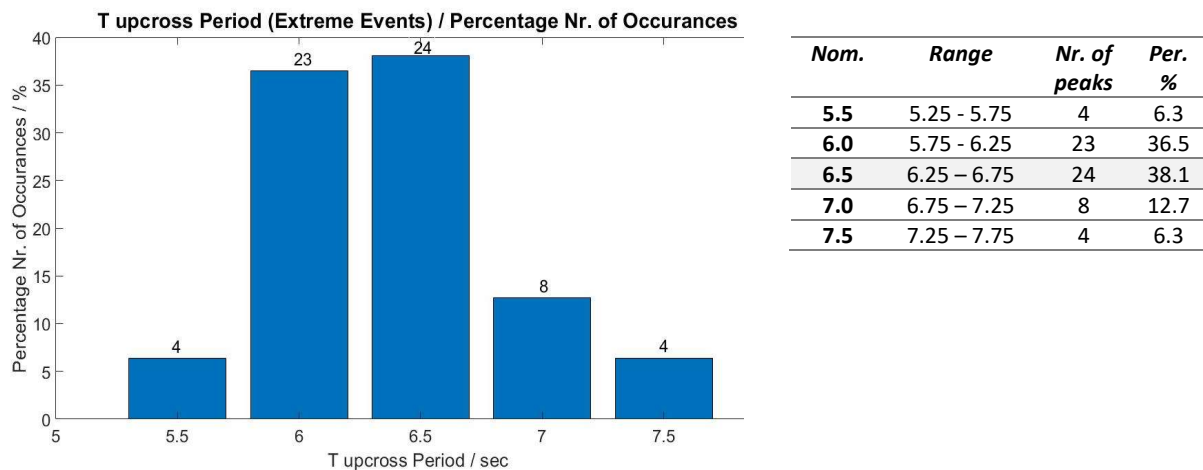
Figure 77: Mean up-cross Period / Percentage Nr. of Occurances of Peak Events

6. Wave Length, *L*

Nominal	Range	Nr. of peaks	Perc. %.
50	45 - 55	12	19.0
60	55 - 65	27	42.8
70	65 – 75	13	20.6
80	75 – 85	9	14.3
90	85 – 95	2	3.2
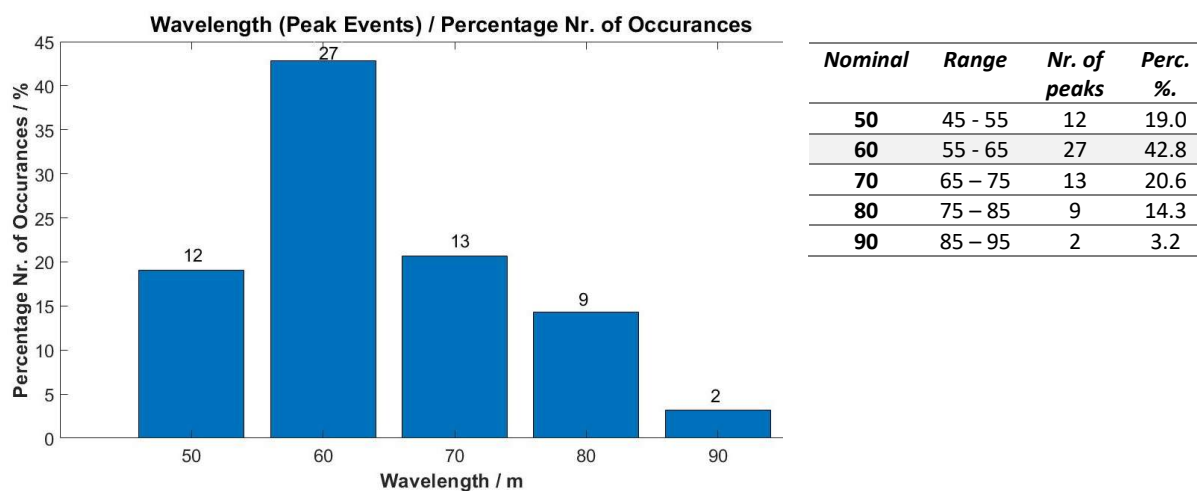
Figure 78: Wave Length / Percentage Nr. of Occurances of Peak Events

7. Wave Steepness

Nominal	Nr. of peaks	Per %
0.060	1	1.6 %
0.065	3	4.8 %
0.070	20	31.7 %
0.075	22	34.9 %
0.080	16	25.4 %
0.085	0	0.0 %
0.090	1	1.6 %
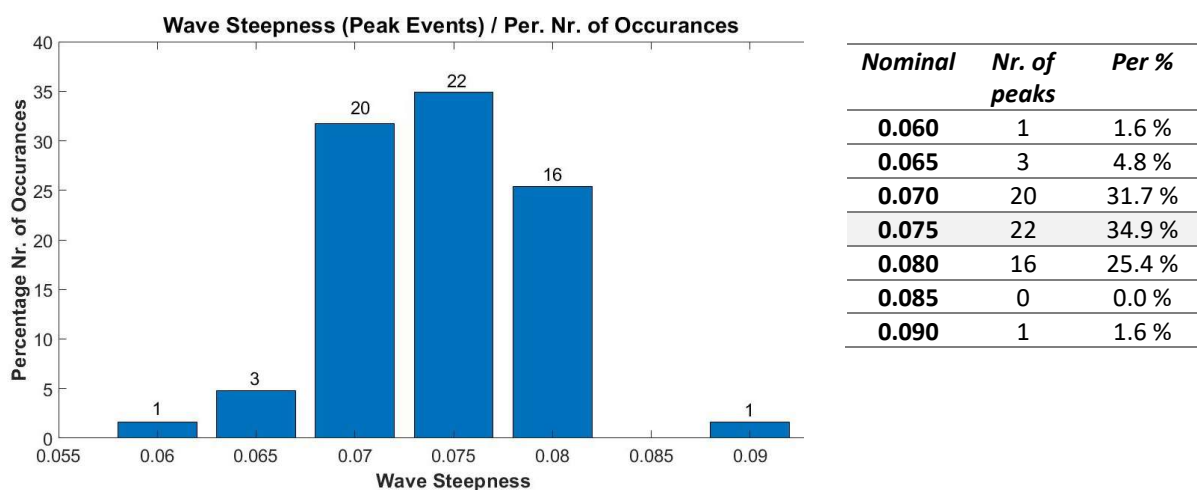
Figure 79: Wave Steepness / Percentage Nr. of Occurances of Peak Events

6.7 Determining T_z for the *100-year* return H_s

The equation for wave steepness for *deep water waves* is given by Equation (7):

$$Wave\ Steepness\ =\ \frac{2\,\pi\,H_s}{g\,T_z^2}$$

For open sea conditions, the wave steepness can be taken as 1/18 (that is 0.055), whilst for areas that are fetch limited a value of 1/14 (that is 0.07) can be taken (9). This compares favourably with the values extracted from the peak storm events above, where 35% of the peak storms had a wave steepness of 0.075 (Figure 79). As determined above, the 100-year return

value of H_s, $H_{s\ 100}$, was found to be 8.8 m. Substituting these values in the above equation for wave steepness, a mean wave period T_z of 8.7 s is obtained.

Hence the wave conditions at the once in a 100-year extreme event is forecasted to have a Significant Wave Height H_s of 8.8 m and a Mean Wave period T_z of 8.7 s.

6.8 Determining h_{max} for the peak H_s

Using the Rayleigh Distribution, and *Equation (15)* the highest wave h_{max} occurring during the 3-hour storm involving the 100-year return value of H_s is given by:

$$h_{max} = K\ H_{sn}\ [0.5 \times (\ln N_n)]^{1/2}$$

Where N_n = number of waves in the 3 hours of storm having the N-year return H_s
              = $(3\ x\ 60\ x\ 60)\ /\ T_{zn}$ = 1245
     H_{sn} = 100-year return value of H_s = 8.8 m
     K = constant equal to 0.90 *(reference 34)*
     T_{zn} = Mean Wave period = 8.67 s

Substituting in above equation, yields the value of the highest Individual Wave, h_{max100} ,that is forecasted to occur in the next 100 years based on the assumption it will occur in same extreme event containing H_{s100}. The value obtained is h_{max100} = 15 m.

6.9 Determining h_{max} using Battjes Method

Using *Equation (25)*,

$$M(h) = \sum\sum \frac{N\left(H_s,\ T_z\right) \times P(h\ //H_s)\ \times 10{,}800}{T_z}$$

Where $M(h)$ = Expected total number of waves exceeding value h in 1 year
        $N\left(H_s,\ T_z\right)$ = total number of counts in matrix $H_s\ /\ T_z$
        $P\ (h\ ///Hs)$ = Pr. of Exceedance of h for a value of H_s based on Rayleigh Dist.
                    = $\exp\left(-\ 2\ h^2\ /\ Hs^2\right)$
        T_z = Mean Wave period

The expected total number of waves exceeding value h in 1 year, *M(h),* was plotted against wave height h (Figure 80).
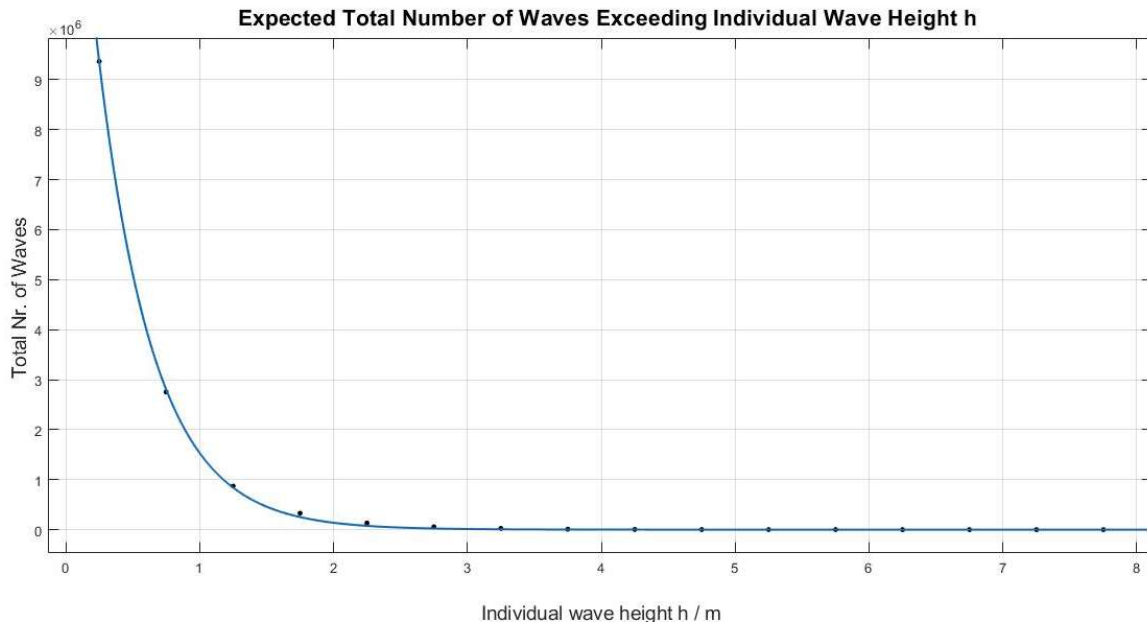
Figure 80: Nr. of Waves Exceeding Individual Wave Height 'h' against Individual Wave Height 'h'

Using *MATLAB*, an exponential function

$$M(h) = A \, e^{B \, h} \qquad\qquad (41)$$

was fitted to the curve. The coefficients *A* and *B* (with 95% confidence bounds) for the curve were found to be

$$A = \; 2.092\text{e}+07 \qquad \text{and} \qquad B = \quad -1.649$$

M(h) is the probable number of waves exceeding individual wave height *h* in one year. The objective is to find the individual wave height *h* that occurs once every *n-years*. Hence *M(h)* has a value of (1/*n*), that is 1/100 to find the highest *h* for the next 100 years. Substituting this value in *Equation (41)* above yields:

$$\text{Individual Wave Height } h_{\,max\,100} = 13.0 \text{ m}$$

The Highest Individual Wave h_{max100} that is forecasted to occur once in 100-years as computed using the Battjes method is $h_{max100} = 13$ m.

This is in actual fact less than that obtained by the Rayleigh distribution and the H_{s100} value.

In Section (5.2.8) the H_s and the $h_{\,max3hr}$ values were compared for the data collected on the 6th December 2020. A difference of 3.5 m was found based on an H_s value of 5 m that is $h_{\,max3hr}$ is 70 % higher as compared to H_s. In order to determine which of the two values for $h_{\,max\,100}$ is more accurate, the same percentage difference will be applied between H_s and $h_{\,max\,100}$.

Applying the same percentage (70 %), to the 100-year return value of H_s, $H_{s\ 100}$ that is 8.8 m, yields

$h_{max\ 100}$ = 15 m

Hence the highest Individual Wave h_{max100} *that will occur once in a 100-years,* h_{max100} *, is forecasted to be 15 m.*

6.10 Joint Frequency Table / Scatter table H_s / T_z Plot

Sea statistics are generally defined by two main parameters the H_s and T_z. For this purpose, wave climatic data can be presented by combining Significant Wave Height H_s and Mean Wave Period T_z in one table.

The H_s and T_z data were obtained from the hindcast data generated by the *ROSARIO MARIA SWAN* model for the period 2011 – 2019. The T_z values were corrected using the correction factors from the validation section. H_s and T_z are grouped in bins of 0.5 m and 0.5 s with the first bins having values of 0 – 0.25 m.

Highest occurring bin is for H_s values of 0.25 m – 0.75 m and T_z 3.25 – 3.75 s with a percentage occurrence 15.65 %.

Tz Range

Hs Range	Hs Nominal Value	Tz Range →	0 - 0.25	0.25 - 0.75	0.75 - 1.25	1.25 - 1.75	1.75 - 2.25	2.25 - 2.75	2.75 - 3.25	3.25 - 3.75	3.75 - 4.25	4.25 - 4.75	4.75 - 5.25	5.25 - 5.75	5.75 - 6.25	6.25 - 6.75	6.75 - 7.25	7.25 - 7.75	7.75 - 8.25	8.25 - 8.75	Sum
		Tz Nominal Value →	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5	8	8.5	
0.00-0.25	0		0.00	0.00	0.00	0.00	0.01	0.41	2.44	3.39	0.54	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.83
0.25-0.75	0.5		0.00	0.00	0.00	0.00	0.00	0.43	4.84	15.65	13.24	4.44	1.17	0.21	0.03	0.00	0.00	0.00	0.00	0.00	40.01
0.75-1.25	1		0.00	0.00	0.00	0.00	0.00	0.00	0.08	5.57	10.78	4.69	1.73	0.68	0.22	0.04	0.02	0.00	0.00	0.00	23.82
1.25-1.75	1.5		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	4.16	6.49	1.37	0.40	0.17	0.10	0.08	0.00	0.00	0.00	12.79
1.75-2.25	2		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	3.40	3.01	0.47	0.10	0.04	0.03	0.02	0.00	0.00	7.14
2.25-2.75	2.5		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	2.86	0.82	0.09	0.05	0.01	0.00	0.00	0.00	4.00
2.75-3.25	3		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.46	1.51	0.26	0.03	0.01	0.00	0.00	0.00	2.27
3.25-3.75	3.5		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.76	0.58	0.03	0.00	0.00	0.00	0.00	1.37
3.75-4.25	4		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.60	0.07	0.00	0.00	0.00	0.00	0.74
4.25-4.75	4.5		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.28	0.23	0.00	0.00	0.00	0.00	0.51
4.75-5.25	5		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.20	0.01	0.00	0.00	0.00	0.23
5.25-5.75	5.5		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.08	0.00	0.00	0.00	0.15
5.75-6.25	6		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.05	0.00	0.00	0.00	0.07
6.25-6.75	6.5		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.03	0.00	0.00	0.06
6.75-7.25	7		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7.25-7.75	7.5		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7.75-8.25	8		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total			0.00		0.00	0.00	0.01	0.84	7.36	24.64	28.79	19.25	10.62	4.93	2.33	0.84	0.33	0.06	0.00	0.00	100.00

Legend:
- > 10%
- 10 > Per > 5
- 5 > Per > 1
- > 1

Table 25: the Hs / Tz Scatter Table

Chapter 7 : Conclusion and Areas for Further Work

7.1 Conclusion

1. The Significant Wave Height data generated by the *ROSARIO MARIA SWAN* model was validated against the data provided by the in-situ Datawell buoy which was located off the coast of Gozo. The validation was carried out for the full time series of H_s and for values of H_s greater than 2.3 m which was eventually used in the Extreme Value Analyses.

2. The Mean Wave Period data generated by the *ROSARIO MARIA SWAN* and by the *ROSARIO SKIRON SWAN* Wave Models were also validated against the Buoy Data.

3. Wave and weather conditions are major cause of incidents at sea. Despite all the available information and regulations, incidents at sea do happen and with a large percentage (66%) occurring at less than extreme condition. Other wave parameters besides Significant Wave Height H_s were found to have an effect on vessel stability mainly:

    1. Wave Length, *L*;
    2. Wave Steepness;
    3. Individual Wave Height, *h*;

    Using the forecasted and validated data generated by the *ROSARIO MARIA SWAN* model, and using the Linear Wave Theory, 2-D operational plots of the forecasted data of the above three wave parameters were produced. The area, resolution and time frames of these plots are as per model data.

4. The Extreme Value Analyses was carried out using the hindcast data generated by the *ROSARIO MARIA SWAN* model for the period 2011 – 2019. The Peak Over Threshold Method was used. Various Upper Threshold Values were analysed withing the range of 2.5 m to 4.5 m to better understand the effect of this parameter on the predicted 100-year return value of H_s. Also, 3 distributions were fitted to the peak data, the Weibull, the Gumbell and the Generalised Pareto Distribution to determine which distribution produced the more realistic values.

    After analysing the Cumulative Distribution plots, the Linear Regression lines and the Linear Correlation Coefficient, *r*, the Weibull distribution was selected as the one which replicates more reliably the model data.

Setting the Upper Threshold Value at 3.5 m (which is also the 95-percentile value of the H_s time series when the peripheral months are removed from the time series) and using the Weibull distribution, the predicted 100-year return value of H_s, the H_{s100}, was calculated as 8.8 m.

5. Using the Rayleigh distribution and the H_{s100}, the forecasted value of the 100-year return value of individual wave height h, the h_{100}, was found to be 15 m. Using the Battjes method the forecasted *100-year* return value of *h*, the h_{100} was 13 m.

    The value attained when using the Rayleigh Distribution, that is an h_{100} of 15 m, was considered to be the more realistic.

7.2 Areas for Further Work

1. In 2003, Transport Malta commissioned a report to plot the wave zones around the Maltese islands based on Significant Wave Height H_s for a whole year period and for the summer period. This was carried out in accordance to the EU Directive 98/18/EC to designate area of operation of passenger vessels based on the vessel class. These plots can be updated using the more accurate data generated by the *ROSARIO MARIA SWAN* Model.

2. As previously discussed, the plots of individual wave height, *h*, are presented in 3 formats, the highest individual wave to occur in the next three hours, sea zones based on probability of exceedance of wave height, and zones based on *h* value. The end user should be consulted to understand which of the 3 plots presents the most useful and clear information.

3. The Extreme Value Analyses was carried out for one location, off the coast of Gozo in the vicinity of Dwejra. The same exercise should be carried out in more locations which are also exposed to the easterly winds or areas considered to be more strategically critical.

4. Throughout this project, the model data used was adjusted based on the validation exercise carried out against in-situ data. The H_s data used in the EVA used the correction factor derived from the validation exercise for H_s values greater than 2.3 m. The 2.3 m value was chosen since the upper threshold in EVA was given values in the range of 2.5 m and 4.5 m. Eventually, an upper threshold value of 3.5 m was selected. Hence, the validation exercise should be repeated for H_s values higher than 3.5 m which could give a different correction factor from the 1.05 obtained for H_s values greater than 2.3 m.

5. The hindcasted data of H_s generated by the *SWAN* model covered a period between 2011 and 2019. The upper threshold value was chosen based on the number of peaks selected. A longer time series would allow the use of a higher upper threshold value meaning that less peaks not belonging to the extreme wave group would be selected. This will result in a more accurate Extreme Value Analyses.

6. When fitting the Generalised Pareto distribution, the *N-year* return value of H_s reached its upper limit during the 8-year period of data available. A better understating of why this has occurred should be carried out.

7. A noticeable difference was attained when determining the 100-year return value of the individual wave, h_{100}, using the Battjes method and the Rayleigh distribution. The Battjes method is very much based on the size of the H_s / T_z bins since the mean bin value is used in the calculations. In this exercise bin sizes of 0.5 m and 0.5 s were selected for H_s and T_z respectively. Reduction in the bin sizes, could reduce the margin of error and possibly reduce the variance between the results attained using the two methods.

Bibliography

1.   Douglas Faulkner FREng. Shipping safety A matter of concern. Ingenia. 2003;

2.   Zhang Z, Li XM. Global ship accidents and ocean swell-related sea states. Nat Hazards Earth Syst Sci. 2017;17(11):2041–51.

3.   Toffoli A, Lefèvre JM, Bitner-Gregersen E, Monbaliu J. Towards the identification of warning criteria: Analysis of a ship accident database. Appl Ocean Res. 2005;27(6):281–91.

4.   Unit TM-MSI. Marine Safety Investigation Report Nr. 03/2018 - Hephaestus. 2018;(July):47.

5.   Malta T of. Victims' Bodies to be flown to France.

6.   IMO. IMO MSC.1/Circ 1228 Revised Guidance to the Master for Avoiding Dangerous Situations in adverse wether and sea conditions ; 11 January 2007. 2007;

7.   Caires S. Extreme Value Analysis : Wave Data. JCOMM Tech Rep No 57. 2011;(57):33.

8.   Malta Significant Wave Height Study Main Report. 2003;(December).

9.   Organization WM. Guide to Wave Analysis and Forecasting. Vol. 1998, WMO-No. 702. 1998. 1–109 p.

10.   Ocean E, Boor E. Wind Generated. Encyclopedia of Ocean Sciences. 2001.

11.   EuropeanCouncil. EU Directive 2013/53/EU. 2013;90–131.

12.   European Commission. Directive 2009/45/EC of the European Parliament and of the Council on safety rules and standards for passenger ships. Regulation. 2010;117(2).

13.   Shipping M. Non-Convention Vessel (NCV) Code.

14.   European Commission. EU Directive 2017 / 2108. 2018;2016(68):48–119.

15.   An P, Club I. P & I Loss Prevention Bulletin Marine Weather Ship Handling in Rough Sea Head and countering / Following Seas. 2019;45(April).

16.   M.R.Renilson. The Broaching of Vessels in following Seas. Vol. 3. University of Glasgow; 2012.

17.   Peters WS, Belenky V, Bassler CC, Spyrou KJ. On Vulnerability Criteria for Parametric Roll and Surf-riding. ISSW 12th Proc [Internet]. 2011;2(1):1–6. Available from: http://www.shipstab.org/index.php/conference-workshop-proceedings/issw2011-washington

18.   Maritime T, Committee S. IMO MSC/Circ.707. 1995.

19.   Hanzu-Pazara R, Duse, Varsami C, Andrei C, Dumitrache R. The influence of ship's stability on safety of navigation. IOP Conf Ser Mater Sci Eng. 2016;145(8).

20.   Spyrou KJ. Design Criteria for Parametric Rolling. Ocean Eng Int. 2005;9(1):11–27.

21.   HSE MO. Offshore Technology Report 2001/078. 2001;

22. Report OT. Green water susceptibility of FPSO/FSUs on the UKCS. Offshore Res Focus. 2001;(131):4.

23. Niclasen A, Simonsen K, Karin A. Wave forecasts and small-vessel safety : A review of operational warning parameters. 2010;23:1–21.

24. Jonathan P, Ewans K. Statistical modelling of extreme ocean environments for marine design : A review. 2013;62:91–109.

25. Mathiesen M, Goda Y, Hawkes PJ, Mansard E, Martín MJ, Peltier E, et al. Methodes conseillées pour l'analyse des houles extreme. J Hydraul Res. 1994;32(6):803–14.

26. Ferreira JA, Guedes-Soares C. An application of the peaks over threshold method to predict extremes of significant wave height. J Offshore Mech Arct Eng. 1998;120(3):165–76.

27. Mazas F, Hamm L. A multi-distribution approach to POT methods for determining extreme wave heights. Coast Eng. 2011;58(5):385–94.

28. Soares GG, Henriques AC. Statistical uncertainty in long-term distributions of significant wave height. J Offshore Mech Arct Eng. 1996;118(4):284–91.

29. Katalinić M, Parunov J. Uncertainties of Estimating Extreme Significant Wave Height for Engineering Applications Depending on the Approach and Fitting Technique—Adriatic Sea Case Study. J Mar Sci Eng. 2020;8(4):259.

30. van Vledder G, Goda Y, Hawkes P, Mansard E, Martin J, Mathiesen M, et al. Case studies of extreme wave analysis: a comparative analysis. 1994;(January):978–92.

31. Kapelonis ZG, Gavriliadis PN, Athanassoulis GA. Extreme value analysis of dynamical wave climate projections in the Mediterranean Sea. 2015;66:210–9.

32. Harley M. Coastal Storm Definition. Coast Storms Process Impacts. 2016;1–21.

33. Martzikos N, Prinos P, Memos C, Tsoukala V. Mediterranean Coastal Storms in a Changing Climate. 2019;(May). Available from: www.puertos.es

34. Tucker MJ. An Improved "Battjes" Method for Predicting the Probability of Extreme Waves. 1989;ll(4):212–8.

35. Mackay E, Johanning L. Long-term distributions of individual wave and crest heights. Ocean Eng [Internet]. 2018;165(May):164–83. Available from: https://doi.org/10.1016/j.oceaneng.2018.07.047

36. IOI - MOC Blue Ocean Energy Project ; 2012

37. CAP499.23. Subsidiary legislation 499.23 Commercial Vessels Regulations. 2014;1–20.

38. Recreational Craft Directive 94_25.pdf.

39. Malta T. Code of Practice for the Safety of Commercial Vessels. 2011;11(June):1–148.

40. Cargo P, Tonnage G, Convention S, Length LL, Convention TI, Vessels F, et al. Recommendations for the Safety of Cargo Vessels of less than Convention Size. 2007;(99):1–30.

Appendix A: Vessel Classification

Sea vessels are divided in various categories depending on their purpose, size and area of operation. These categories are governed by different codes and regulations, both on a local level through the local maritime authority, Transport Malta, and on an international level through the IMO. The local Maltese Authority, broadly sub divide vessels into either Commercial, Non-Convention Vessels, Recreational craft and Passenger vessels amongst others each are defined as follows:

1. A '*Passenger Vessel'* means a ship which carries more than 12 passengers; (12)

2. A *'Commercial vessel'* means a vessel, whether power-driven or not, engaged in or intended for use in any trade, business or calling or operating for hire or reward within the ports, internal waters and territorial waters of Malta; (37)

3. *'Non Convention Vessels'* are defined as commercial vessels operating exclusively in Maltese waters which due to their tonnage , usage or dimensions , *if operating in international waters* , they do not form part of the IMO conventions ; (37)

4. A '*Recreational Craft'* means any watercraft of any type, excluding personal (?) watercraft, intended for sports and leisure purposes of hull length from 2.5 m to 24 m, regardless of the means of propulsion; (38)

The following Directives, Codes and Maltese Legal Notices apply accordingly:

1. *Passenger Vessels:*
   a. for domestic navigation, passenger vessels made of steel or equivalent material (aluminium alloy etc.), with length overall (LoA**)** greater or equal to 24m , *the EU regulation, 2009/45/EC* applies (12) ;

   b. for domestic navigation, passenger vessels made of steel or equivalent , with LOA of less than 24m but greater than 15m, the *Non-Convention Vessel (NCV) Code* applies (13)

   c. for domestic navigation, passenger vessels *NOT* made of steel or equivalent material with a LoA greater or equal to 15m, the *Non-Convention Vessel (NCV) Code* applies (13)

   d. for domestic navigation passenger vessels less than 15m LoA, which operate in Maltese waters, the *Code of Practice for the Safety of Commercial Vessels (CVC)* (39) applies; whilst for domestic navigation vessels less than 15m LoA operating in any country, other than Malta, the *IACS Recommendation No.99* (40) applies

   e. for passenger vessels on international navigation the SOLAS Convention applies.

2. *Non-Convention Vessels*

   a. Non-Convention Vessels greater or equal to 15m in overall length are governed by the *Non-Convention Code (NCV)*

   b. Non-Convention Vessels less than 15m LoA, the *Non-Convention Vessel Survey Guidelines (IACS 99)* are applicable,

3. *Commercial Vessels*

   a. For commercial vessels of any size operating within Maltese territorial waters, the *Code of Practice for the Safety of Commercial Vessels (CVC)* apply

4. *Recreational Vessels*

   a. For recreational craft from 2.5 m to 24 m the *EU Council Directive 2013/53/EU* applies (11)

The two EU directives which base area of operation of vessels on wave conditions are:

1.." Safety Rules and Standard for Passenger ships Directive"- 2009/45/EC (12)   and

2. "Recreational Craft Directive" - EU Council Directive 2013/53/EU (11)

The Recreational Craft directive is also adopted in the local codes of practice through the Non-Convention Vessels Code (13).

Codes of Practice and EU Directives

2.2.1   Code of Practice for the safety of Commercial Vessels CVC Codes (39)

Are based on the Subsidiary Legislation 499.23 – Commercial Vessels Regulations, issued 1st October 2002 (37). This Code of Practice regulates the owners of commercial vessels and the safety and manning requirements of such vessels. It does not apply amongst others to Pleasure craft and Fishing boats.

A commercial vessel may not be operated outside the areas assigned to it or be used for a purpose other than that prescribed in the commercial vessel certificate without the approval of the Authority.

2.2.2 *Code of practice Non Convention Vessels (NCV) Code* (13)

The Non-Convention Vessel (NCV) Code is applicable to all Non-Convention Vessels greater or equal to 15m in length overall. An area of operation of NCV vessels is as assigned by the authority as indicated on the survey certificate.

It does not apply for fishing vessels, pleasure crafts and all water sports vessels.

All vessels less than 24m in LoA should be fitted with a builder's plate that shall contain the design category (maximum wind and wave height limitations). The design categories are as per European Recreational Craft Directive (Directive 94/25/EC, updated to 2013/53/EU) quoted hereunder.

Appendix B: Validation of T$_z$ – *ROSARIO SKIRON SWAN*

B.1   Time Series Plot T$_z$
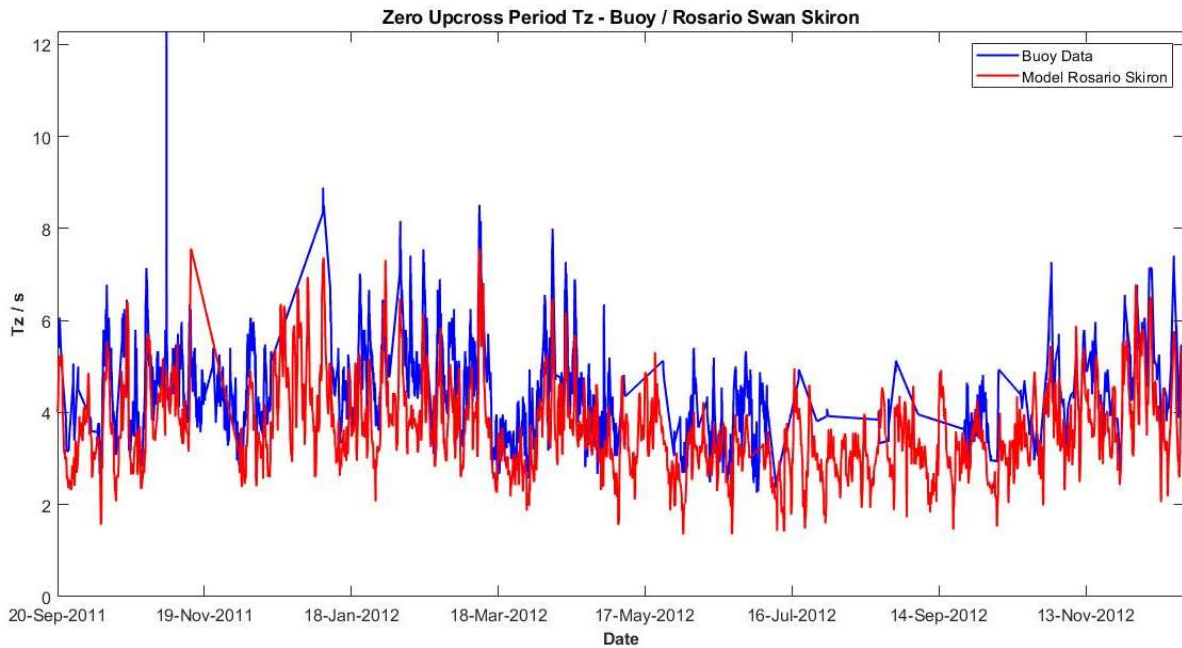
Figure 81 : T$_z$ Mean Up cross Period Time Series; ROSARIO SKIRON SWAN

B.2   T$_z$ Data Sorted by Direction
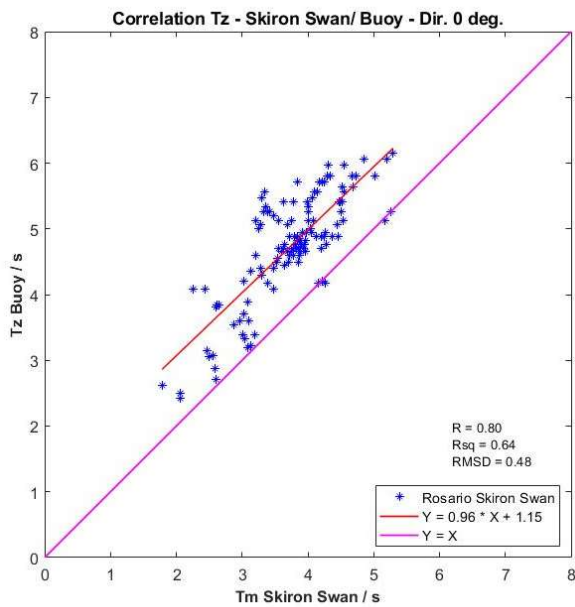
Figure 82 :T$_z$ Skiron Swan / Actual Values; 0 deg.   *Figure 83 :T$_z$ Skiron Swan / Actual Data; 45 deg.*
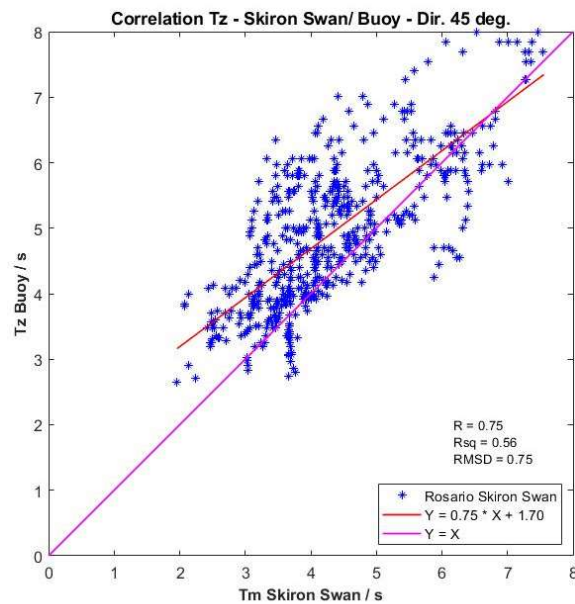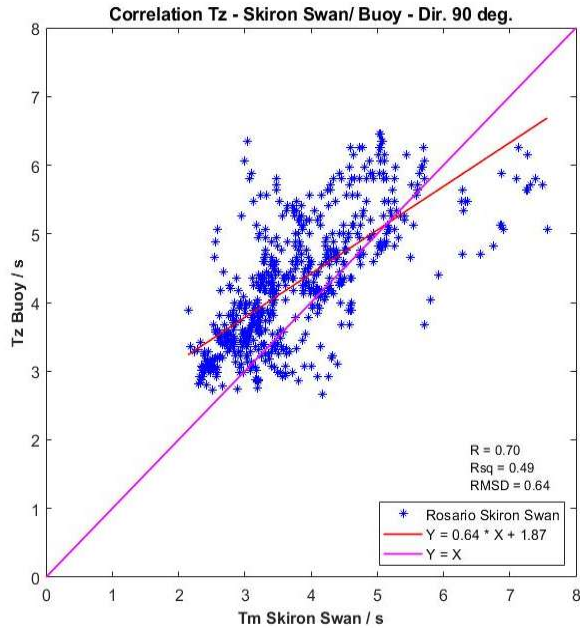
Figure 84: T$_{z\ Skiron}$ Swan / Actual Data; 90deg.   *Figure 85: T$_z$ Skiron Swan / Actual Data; 135 deg.*
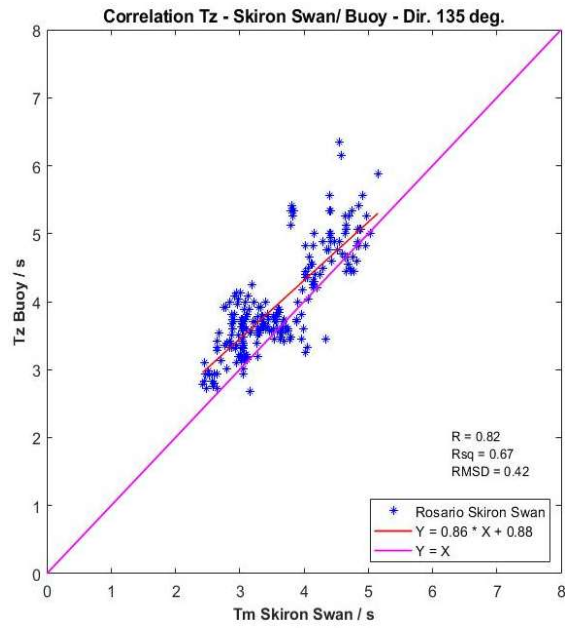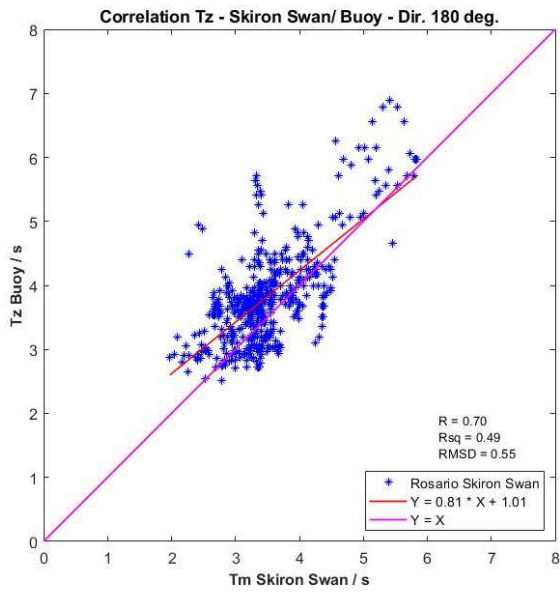
Figure 86: T$_z$ Skiron Swan / Actual Data; 180 deg.   *Figure 87: T$_z$ Skiron Swan / Actual Data; 225 deg.*
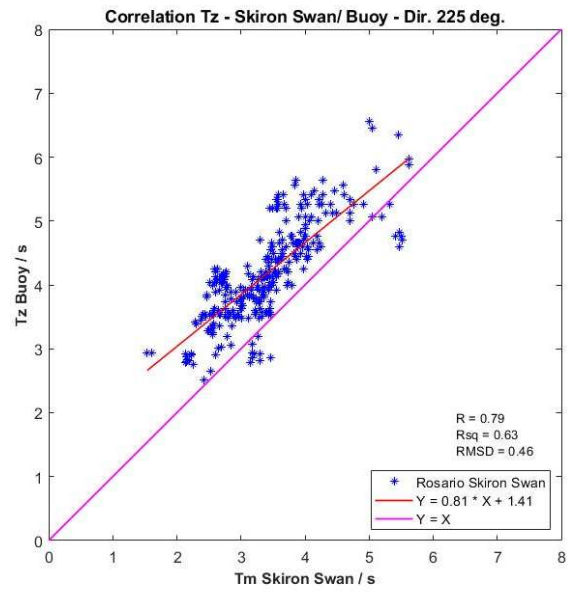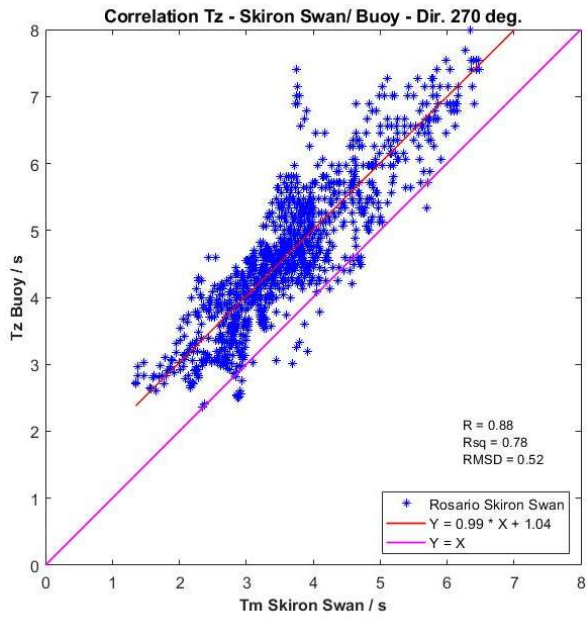
Figure 88: T$_z$ Skiron Swan / Actual Data; 270 deg.   *Figure 89: T$_z$ Skiron Swan / Actual Data; 315 deg.*
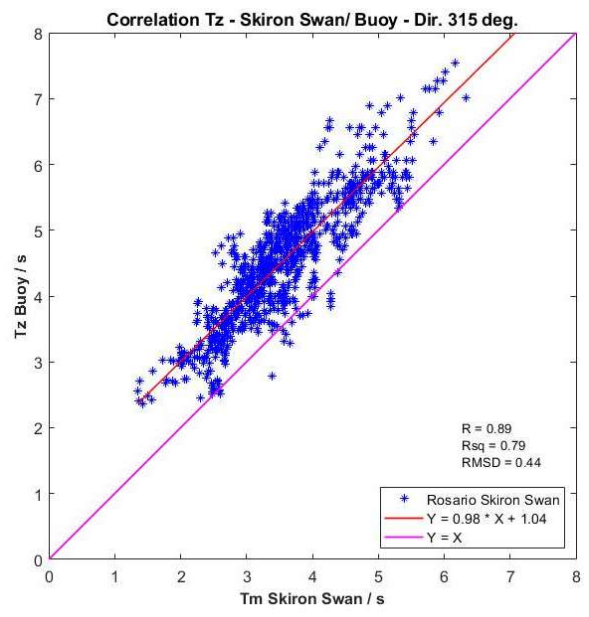
Appendix C: Validation of H$_s$ / T$_z$ – *ROSARIO MARIA SWAN*
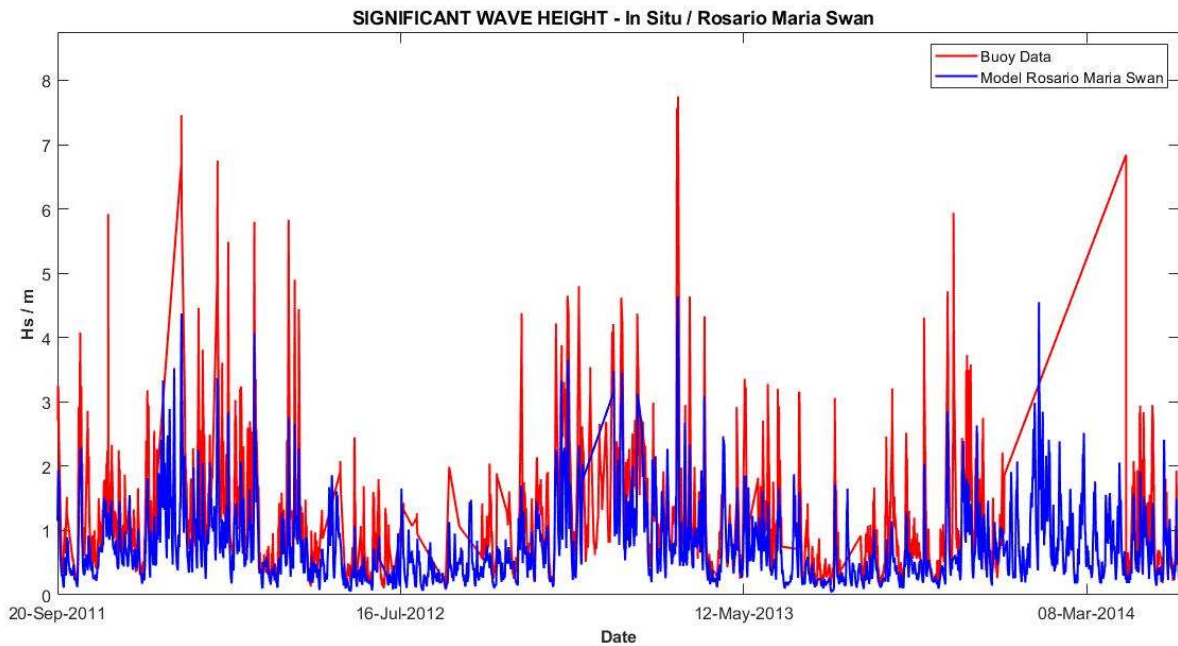
Figure 90: H$_s$ Time Series ROSARIO MARIA SWAN and Buoy Data; 2011- 2014

C.1   Time Series Plot T$_z$
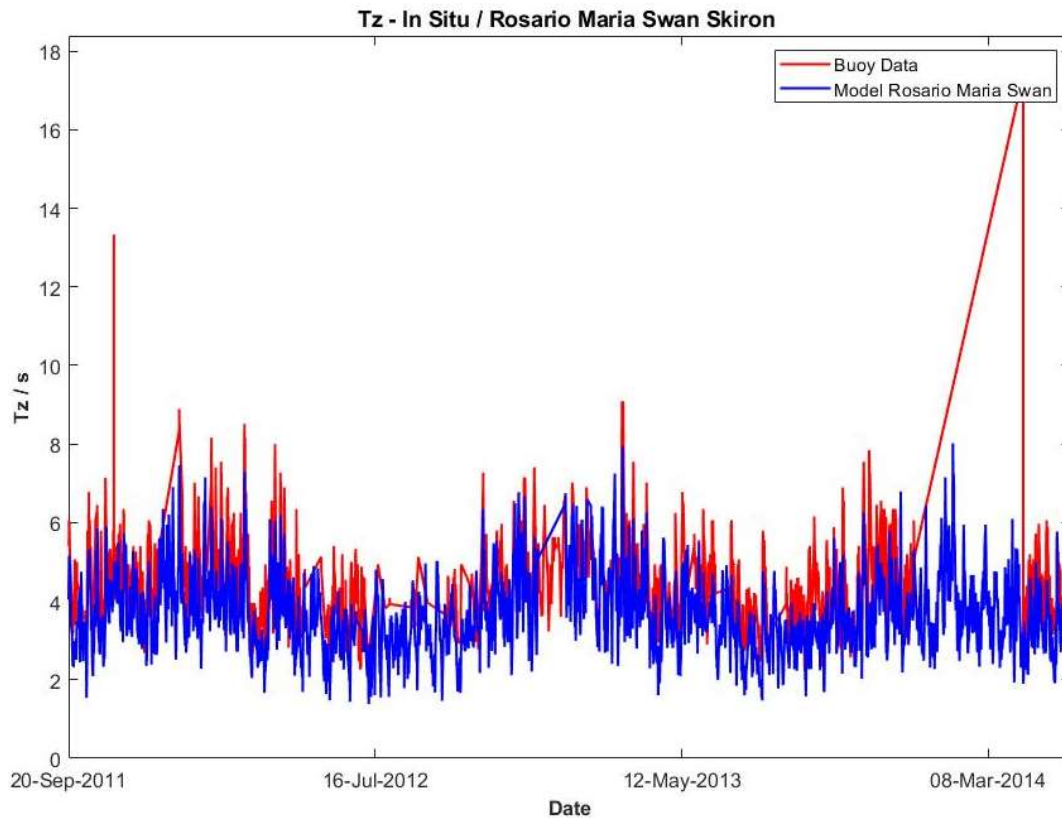
Figure 91: T$_z$ Time Series Actual Data / ROSARIO MARIA SWAN Model

C.2 T_z Data Sorted by Direction
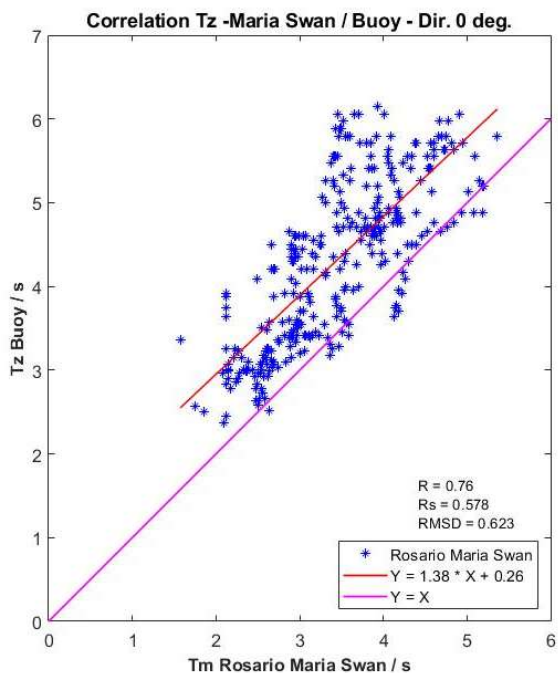
Figure 92: T_z Maria Swan / Actual Data; 360 deg.
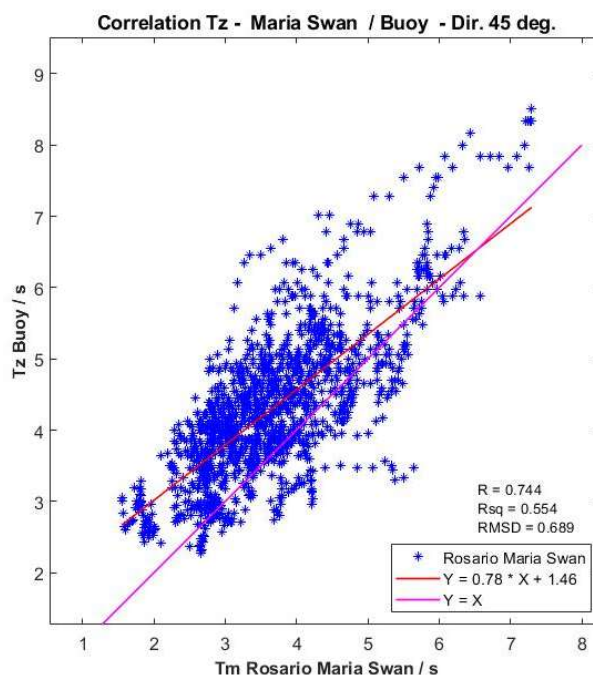
Figure 93: T_z Maria Swan / Actual Data; 45 deg.
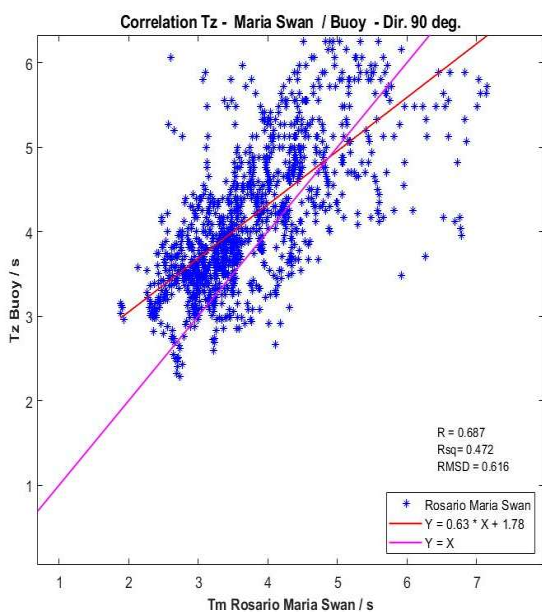
Figure 94: T_z Maria Swan / Actual Data; 90 deg.
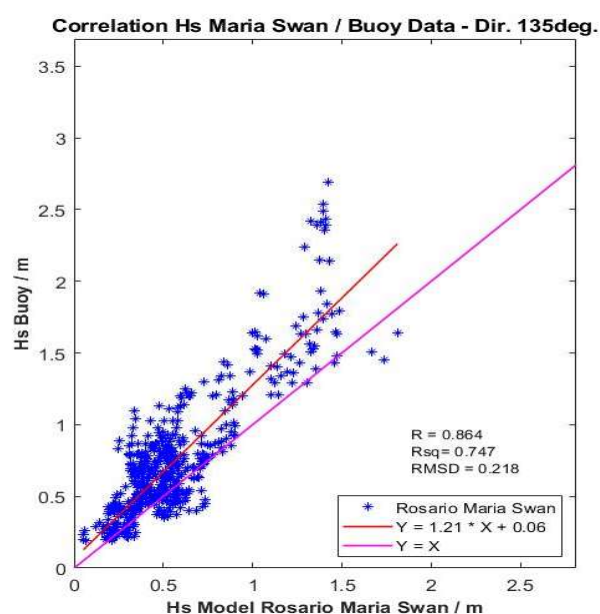
Figure 95: T_z Maria Swan / Actual Data; 135 deg.
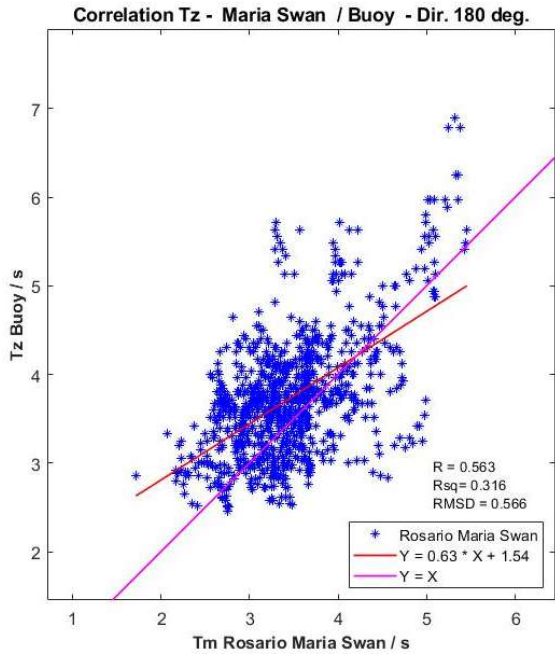
Figure 96: T$_z$ Maria Swan/Actual Data; 180 deg.
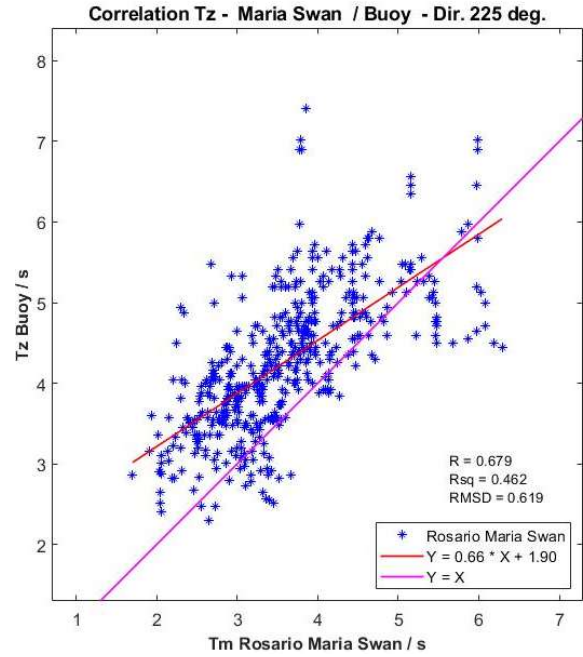
Figure 97: T$_z$ Maria Swan/Actual Data; 225 deg.
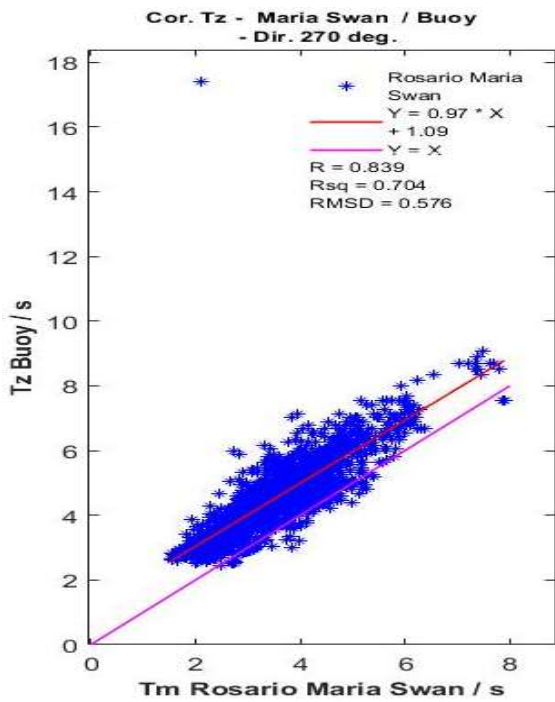
Figure 98: T$_z$ Maria Swan/Actual Data; 270 deg.
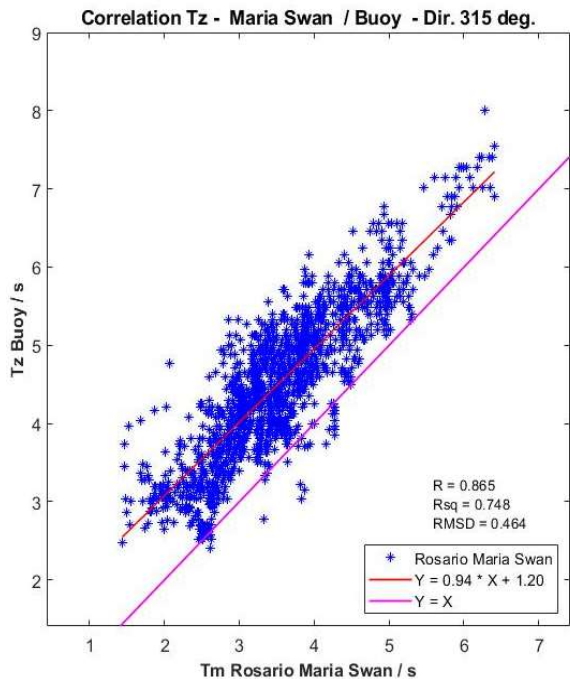
Figure 99: T$_z$ Maria Swan/Actual Data; 315 deg.

C.3　Correction factors for H_s exceeding threshold values of 2.3m
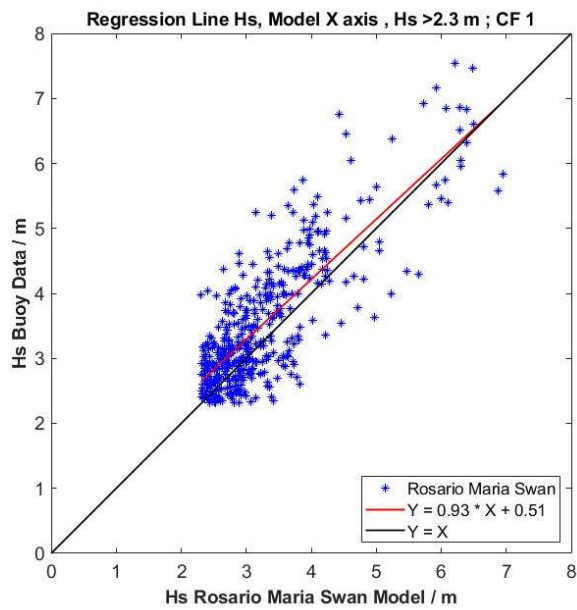
Figure 100: Correlation H_s Maria Swan; Hs>2.3m
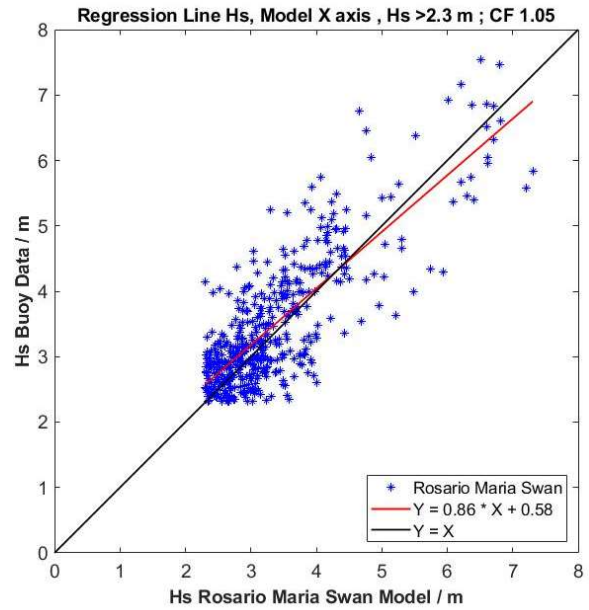
Figure 101: Correlation H_s >2.3m; CF 1.05
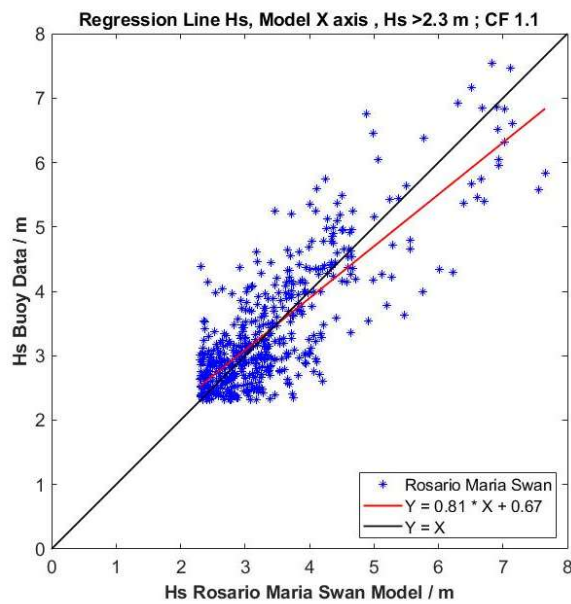
Figure 102: Correlation H_s >2.3m; CF 1.1
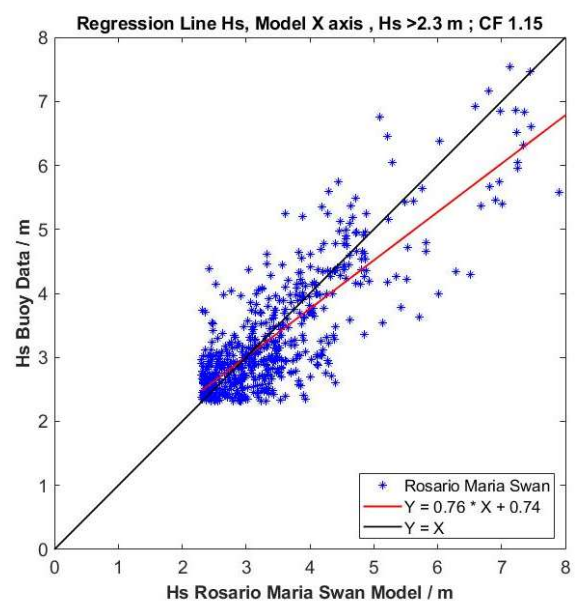
Figure 103: Correlation H_s >2.3m; CF 1.15

Appendix D: Plots Forecasted Wave Parameters

D.1 Wave Length
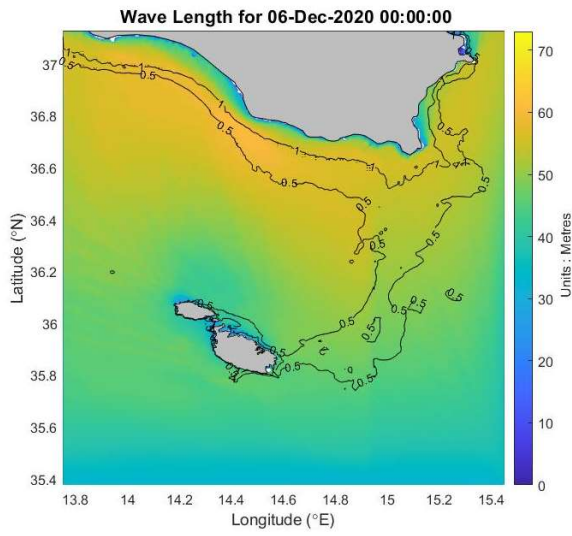
Figure 101 a: Wave Length 06/12/20 at 00.00
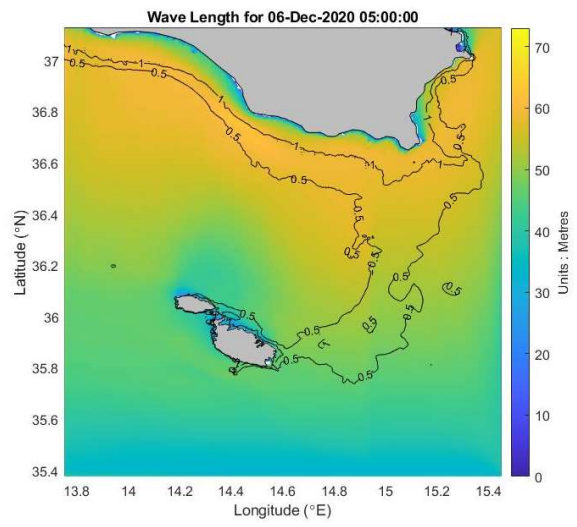
Figure 101 b: Wave Length 06/12/20 at 05.00
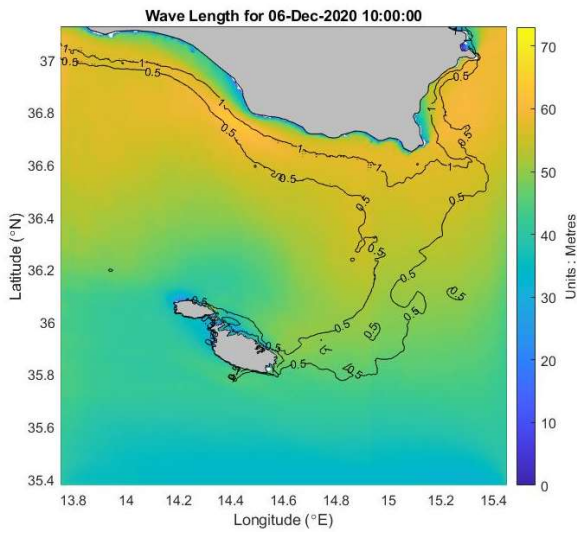
Figure 101 c: Wave Length 06/12/20 at 10.00
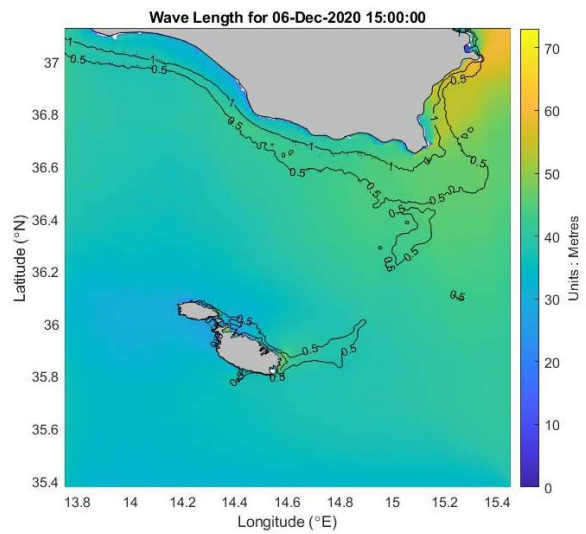
Figure 101 d: Wave Length 06/12/20 at 15.00
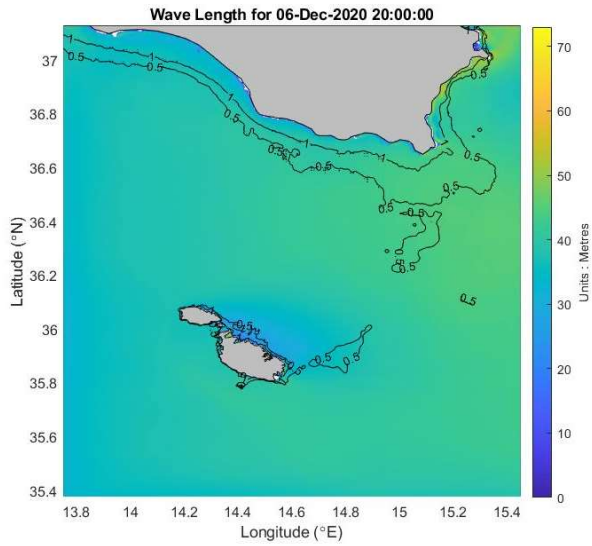
Figure 101 e: Wave Length 06/12/20 at 20.00
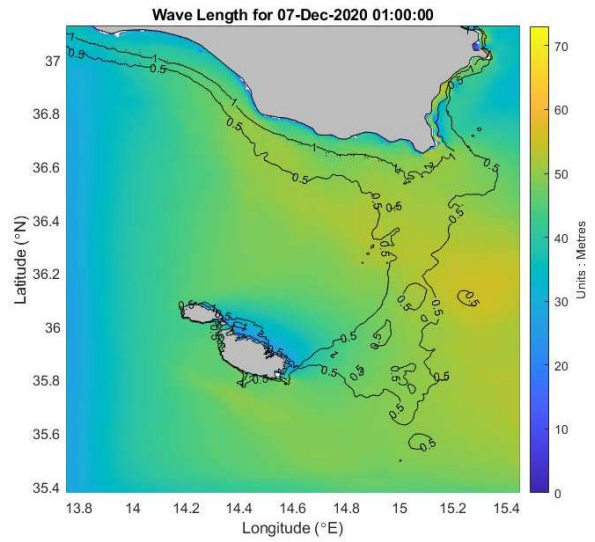
Figure 101 f: Wave Length 07/12/20 at 01.00
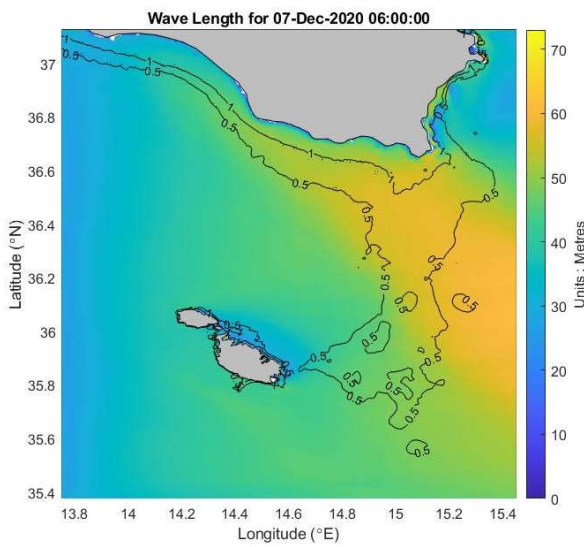
Figure 101 g: Wave Length 07/12/20 at 06.00
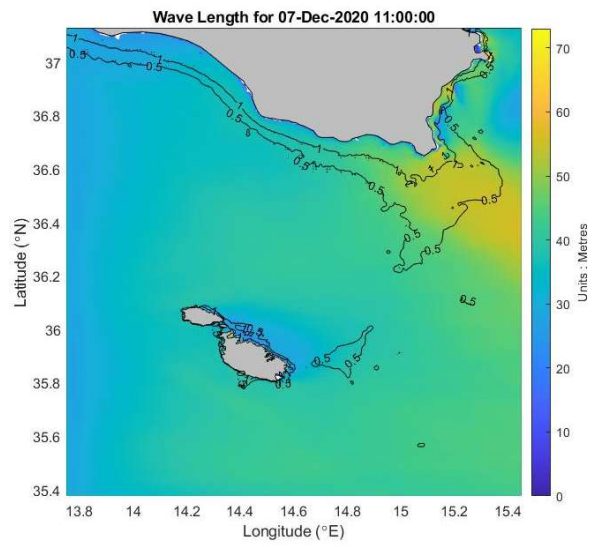
Figure 101 h: Wave Length 07/12/20 at 11.00

D.2 Wave Length with Wave Direction
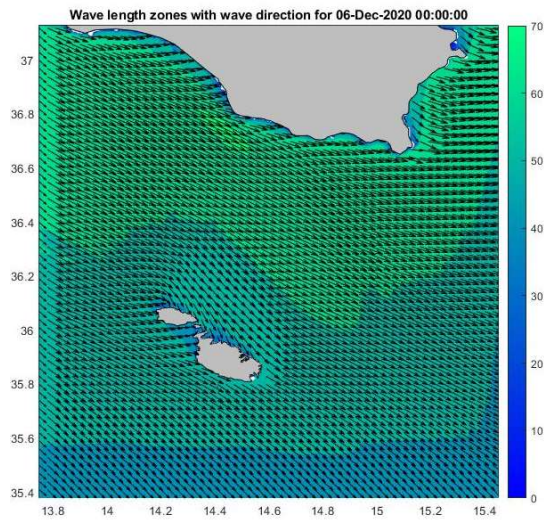
Figure 102 a: Wave Length 06/12/20; 00.00
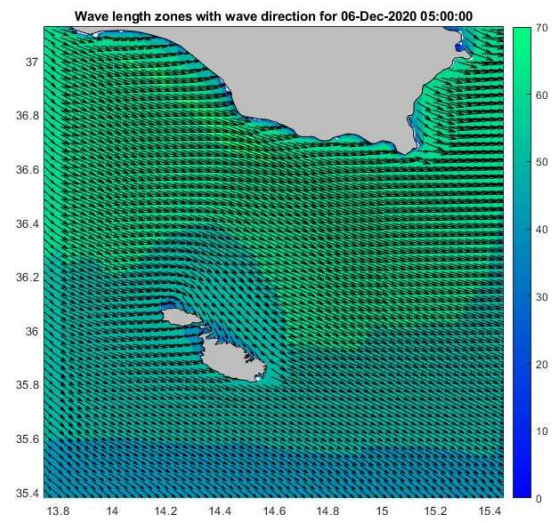
Figure 102 b: Wave Length 06/12/20; 05.00
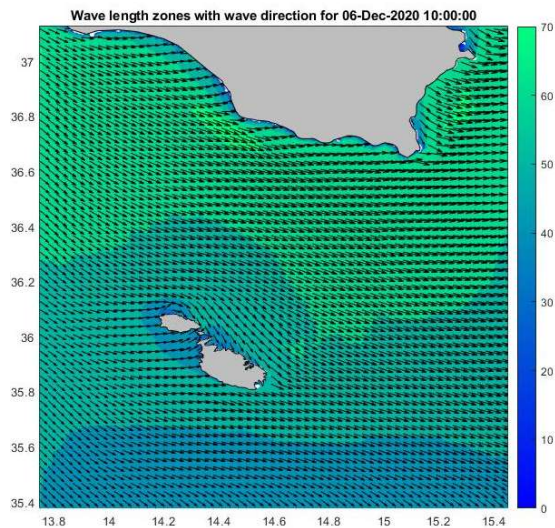
Figure 102 c: Wave Length 06/12/20; 10.00
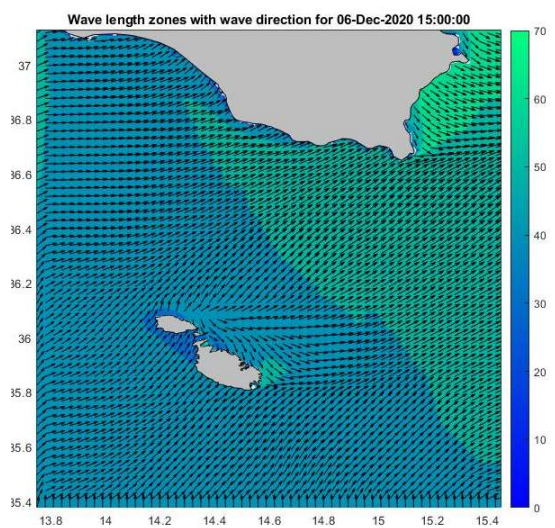
Figure 102 d: Wave Length 06/12/20; 15.00
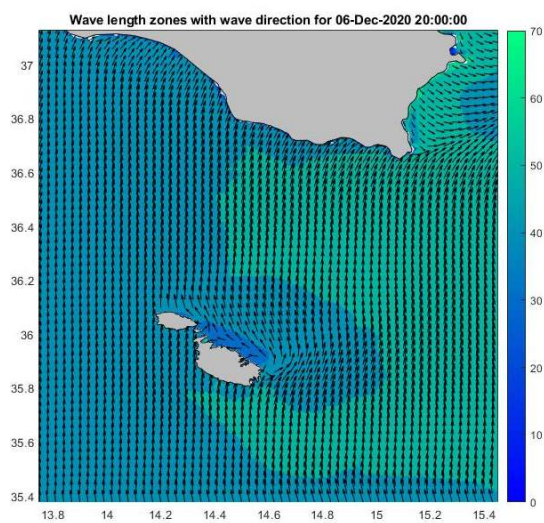
Figure 102 e: Wave Length 06/12/20; 20.00
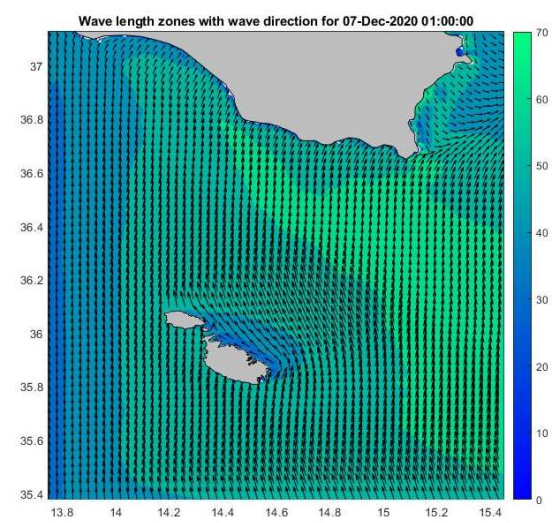
Figure 102 f: Wave Length 07/12/20; 01.00
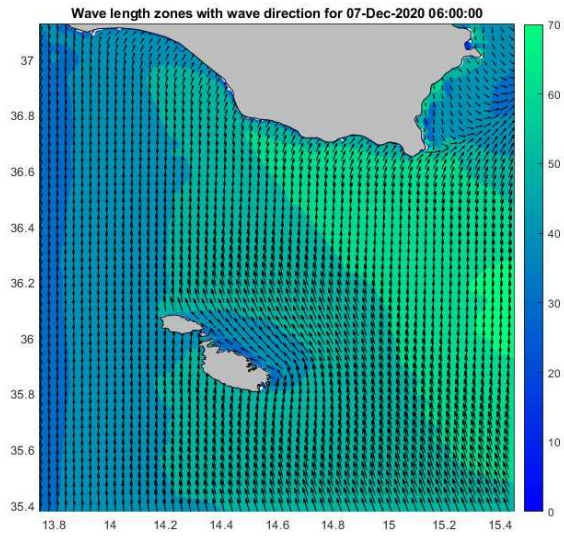
Figure 102 g: Wave Length 07/12/20; 06.00
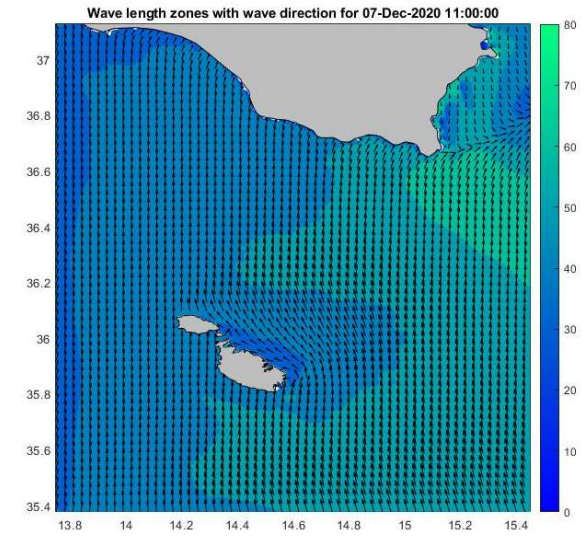
Figure 102 h: Wave Length 07/12/20; 11.00

D.3   Wave Steepness
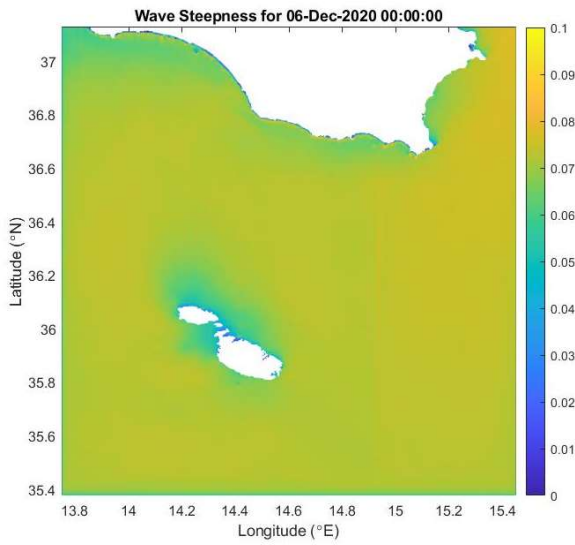
Figure 103 a: Wave Steepness 06/12/20; 00.00
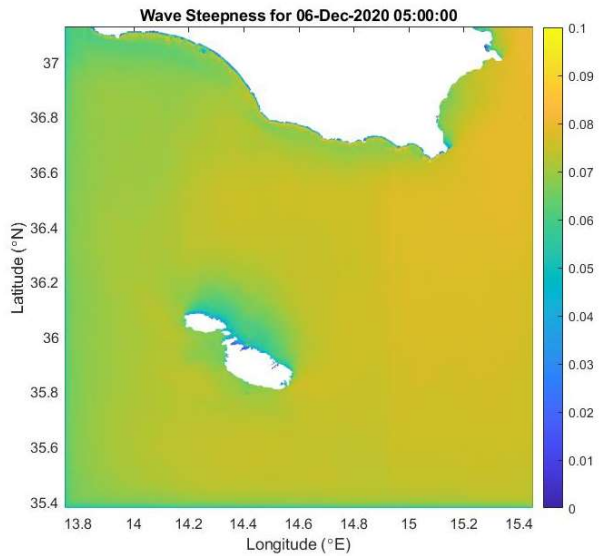
Figure 103 b: Wave Steepness 06/12/20; 05.00
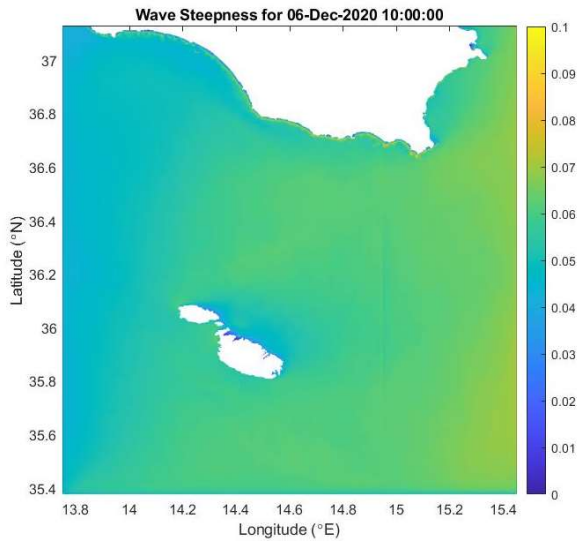
Figure 103 c: Wave Steepness 06/12/20; 10.00
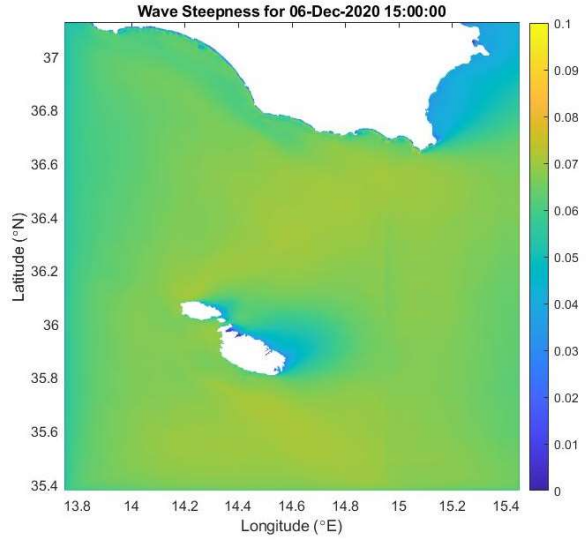
Figure 103 d:Wave Steepness 06/12/20; 15.00
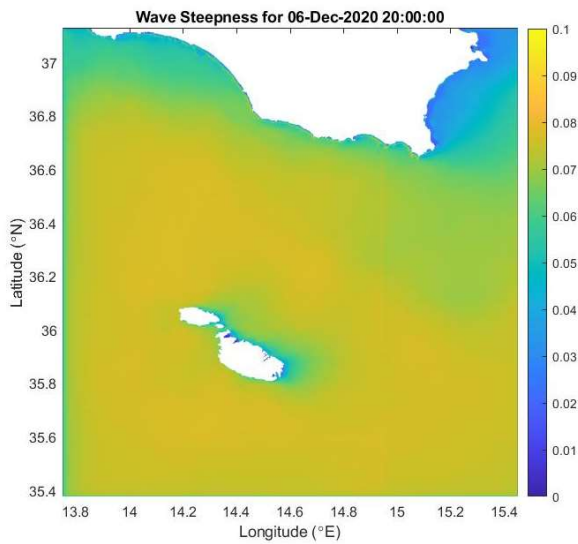
Figure 103 e: Wave Steepness 06/12/20; 20.00
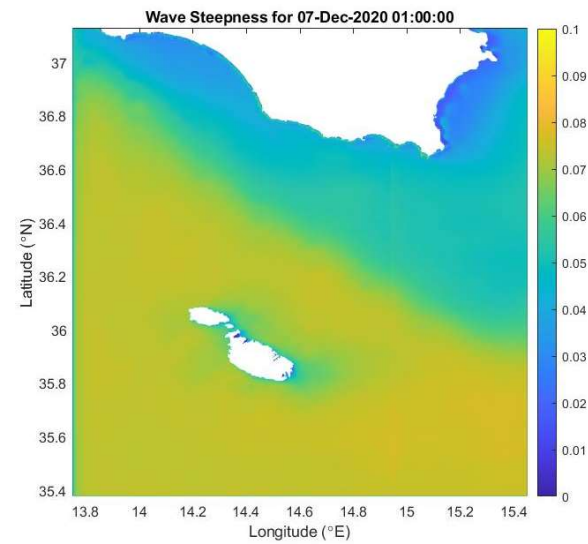
Figure 103 f: Wave Steepness 07/12/20; 01.00
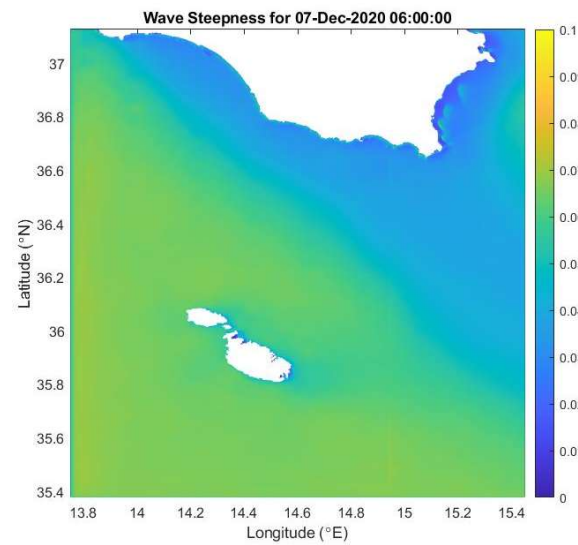
Figure 103 g: Wave Steepness 07/12/20; 06.00
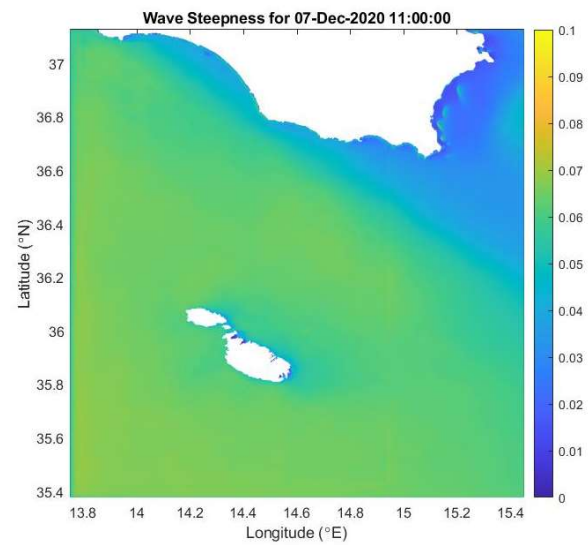
Figure 103 h: Wave Steepness 06/12/20; 11.00

D.4 Wave Steepness with Wave Direction
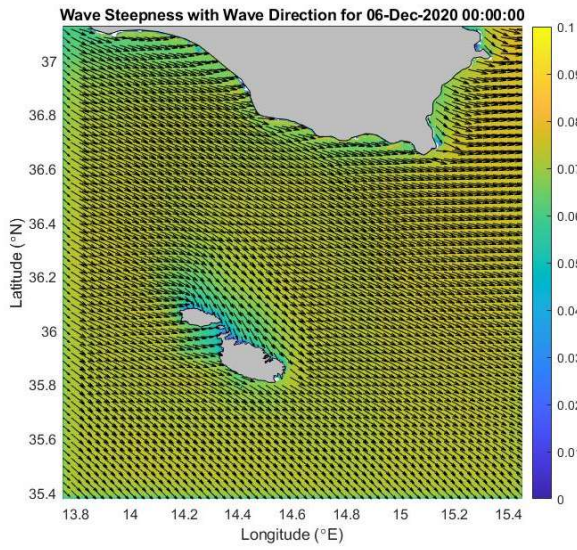
Figure 104 a: Wave Steepness 06/12/20; 00.00
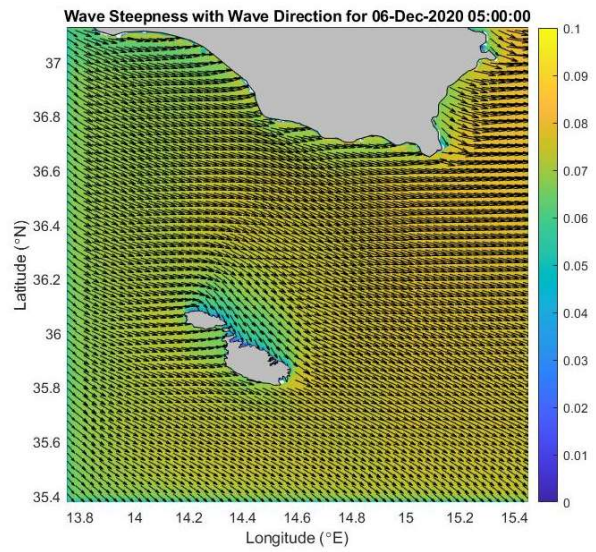
Figure 104 b: Wave Steepness 06/12/20; 05.00
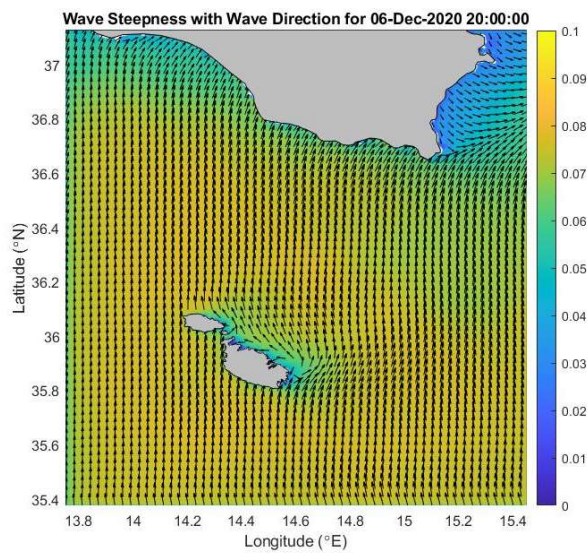
Figure 104 c: Wave Steepness 06/12/20; 10.00
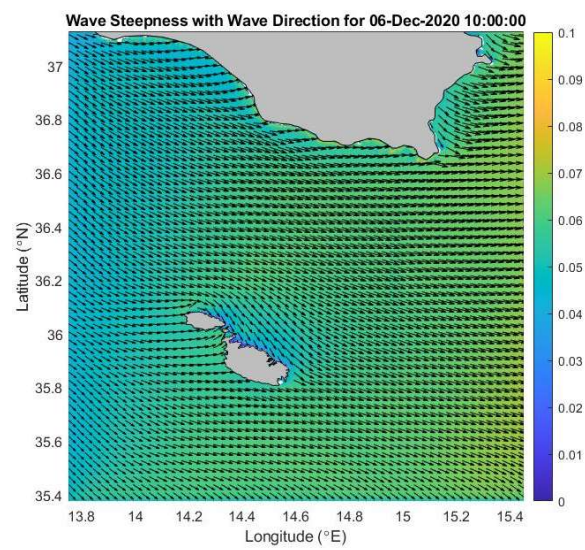
Figure 104 d: Wave Steepness 06/12/20; 15.00
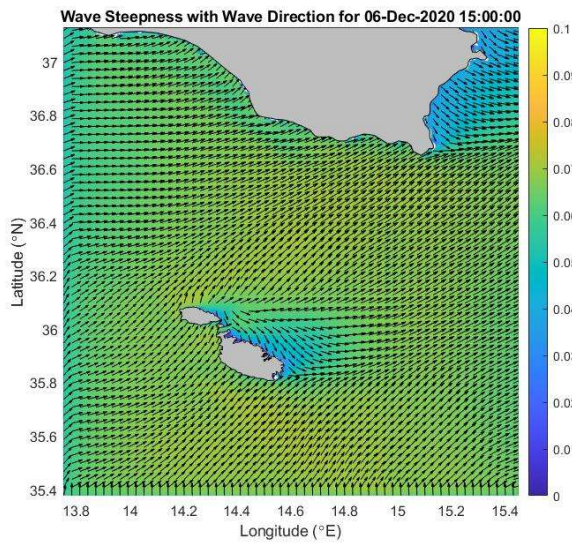
Figure 104 e: Wave Steepness 06/12/20; 20.00
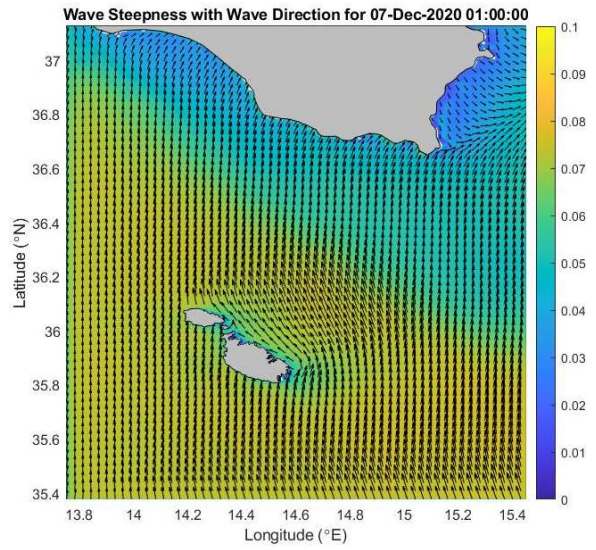
Figure 104 f: Wave Steepness 07/12/20; 01.00
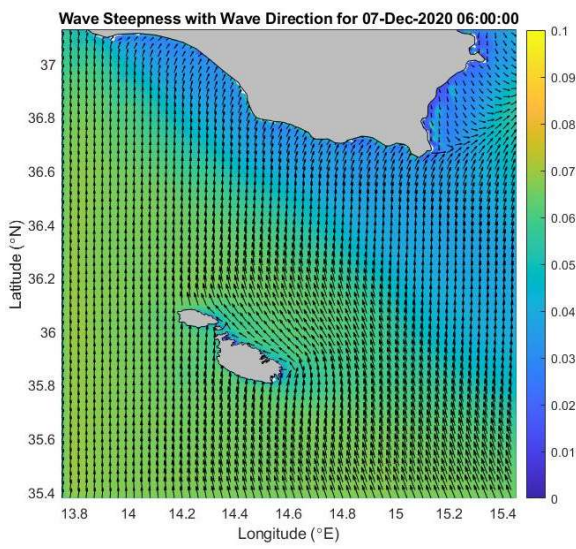
Figure 104 h: Wave Steepness 07/12/20; 06.00
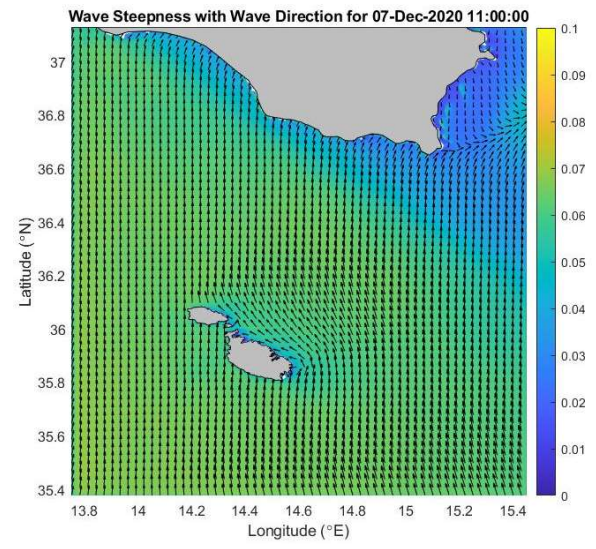
Figure 104 i: Wave Steepness 07/12/20; 11.00

D.5 Probability Wave Height 'h' is greater than 2.5m
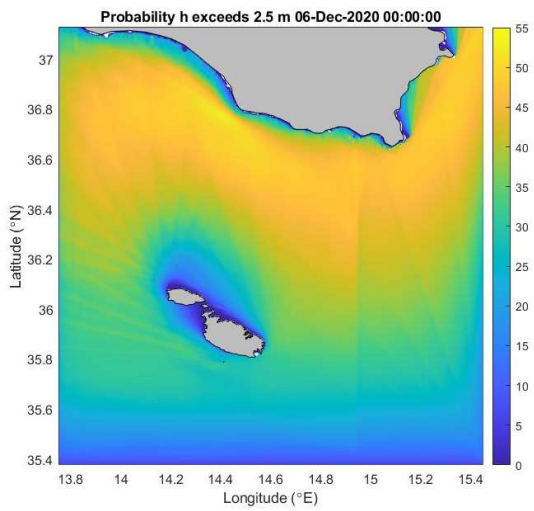
Figure 105 a: Pr. h>2.5m; 06/12/20 at 00.00
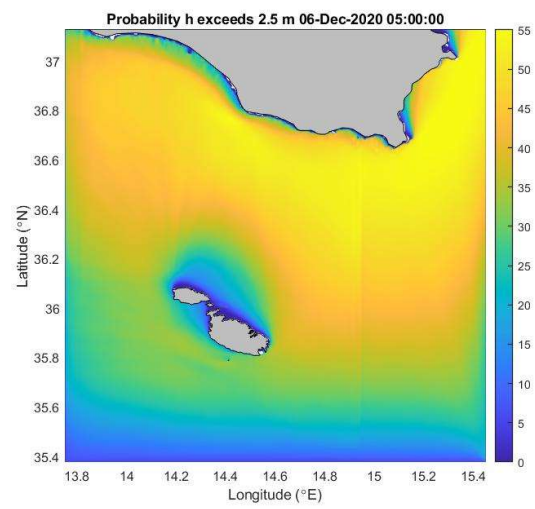
Figure 105 b: Pr. h>2.5m; 06/12/20 at 05.00
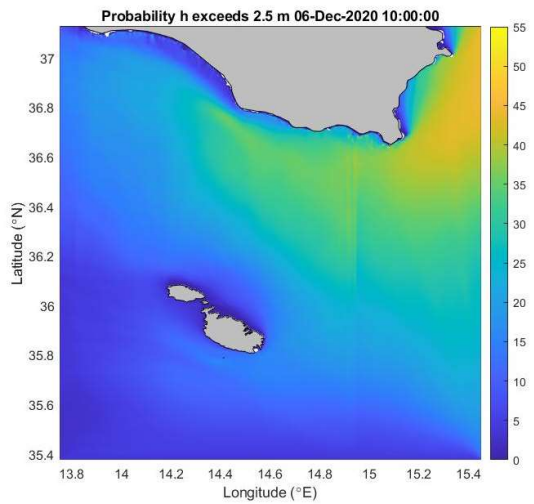
Figure 105 c: Pr. h>2.5m; 06/12/20 at 10.00
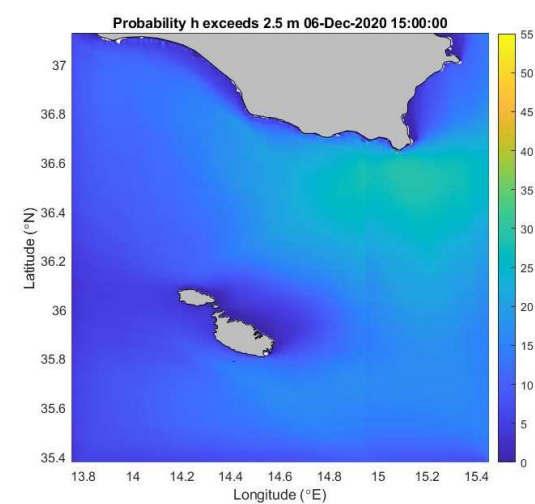
Figure 105 d: Pr. h>2.5m; 06/12/20 at 15.00
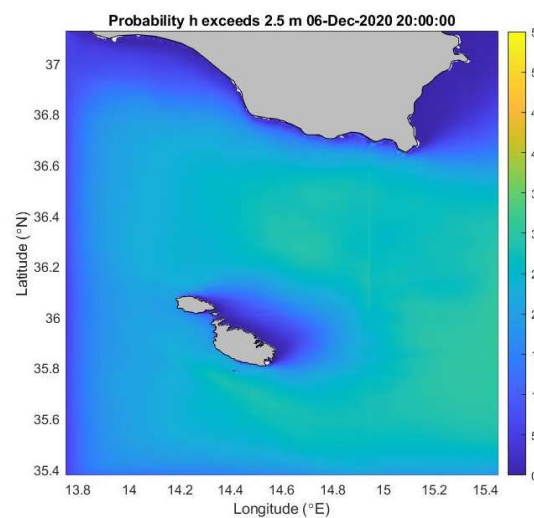
Figure 105 e: Pr. h>2.5m; 07/12/20 at 20.00
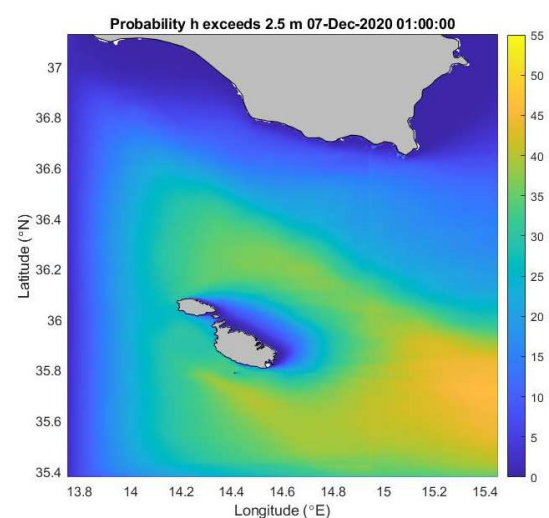
Figure 105 f: Pr. h>2.5m; 07/12/20 at 01.00
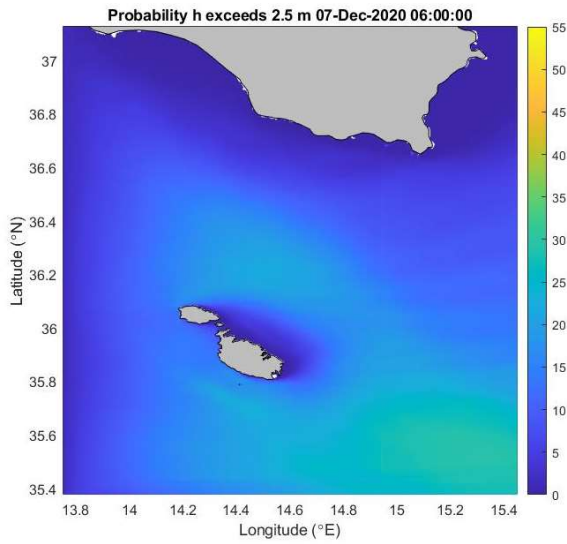
Figure 105 g: Pr. h>2.5m; 07/12/20 at 06.00   *Figure 105 g: Pr. h>2.5m; 07/12/20 at 11.00*
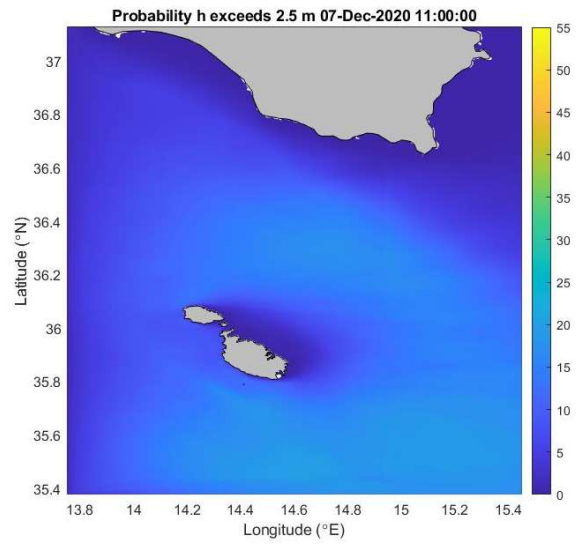
D.6   Individual Wave Height 'h' – h with Pr. exceedance > 20%
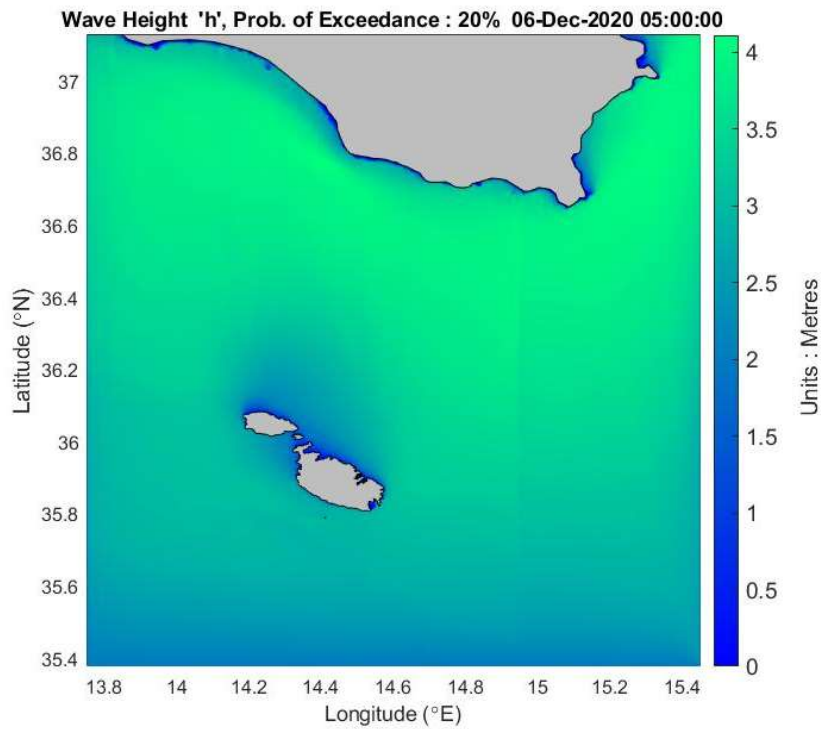
Figure 106: Wave Heights h having percentage exceedance of 20%

D.7   Highest Individual Wave Height 'h' -Max in 3 hr 'h $_{\text{max 3hr}}$'
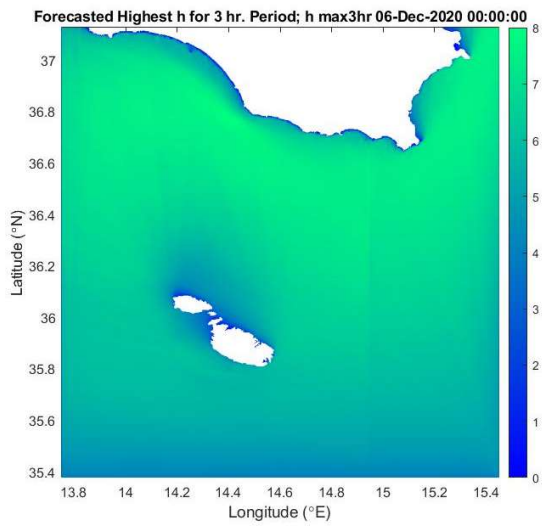
Figure 107 a: h $_{max\ 3hr}$ 06/12/2020 at 00.00
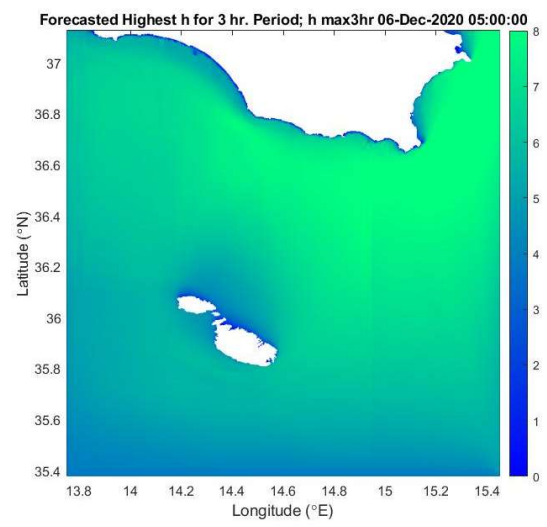
Figure 107 b: h $_{max\ 3hr}$ 06/12/2020 at 05.00
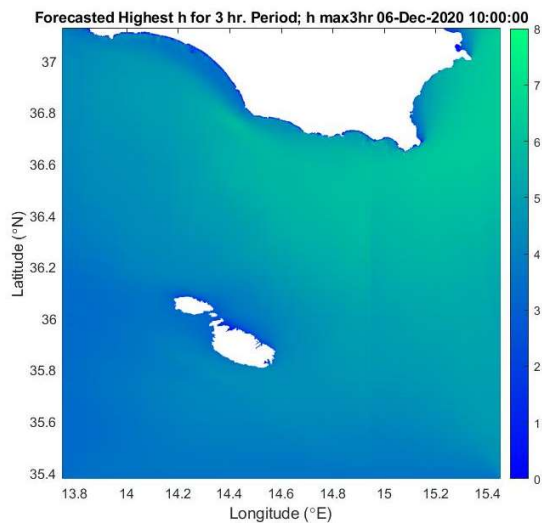
Figure 107 c: h $_{max\ 3hr}$ 06/12/20 at 10.00
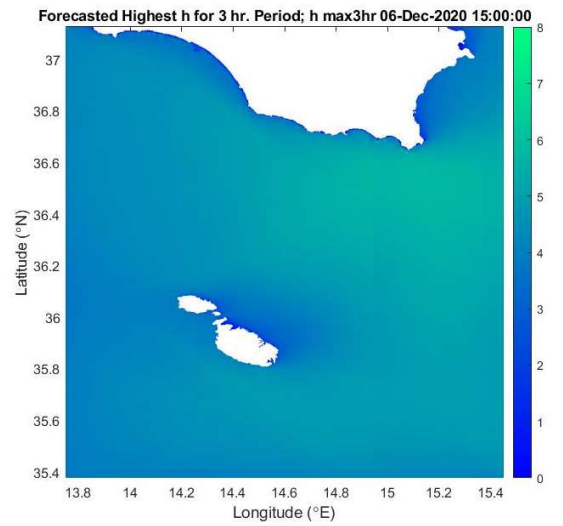
Figure 107 d: h $_{max\ 3hr}$ 06/12/20 at 15.00
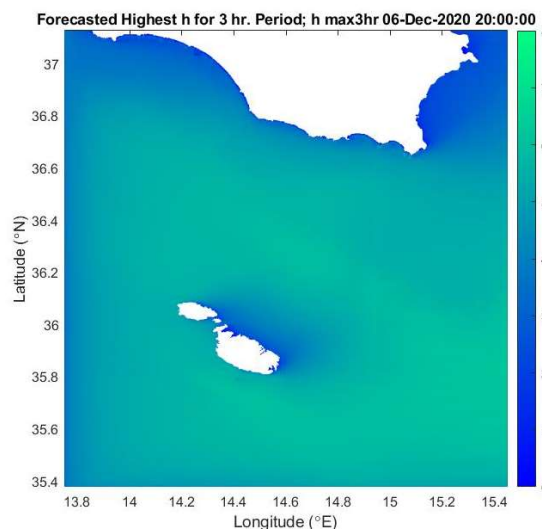
Figure 107 e: h $_{max\ 3hr}$ 06/12/20 at 20.00
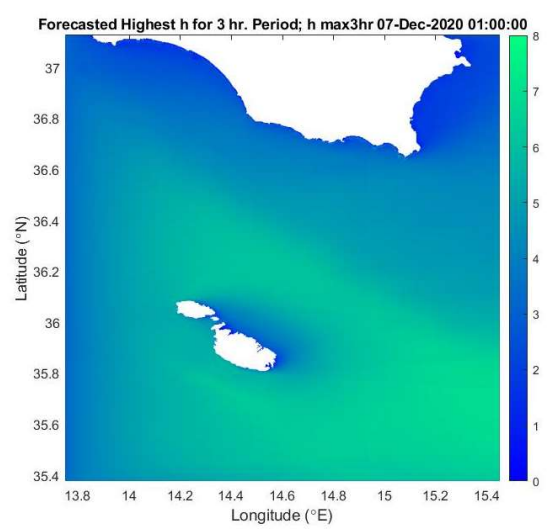
Figure 107 f: h $_{max\ 3hr}$ 07/12/20 at 01.00
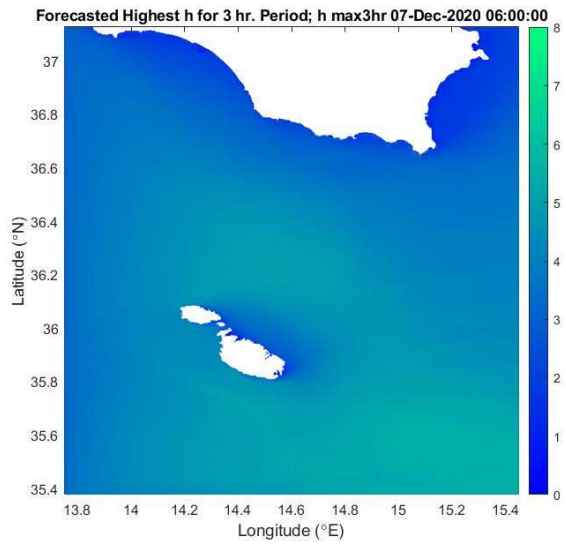
Figure 107 g : h $_{max\ 3hr}$ 07/12/2020 at 06.00
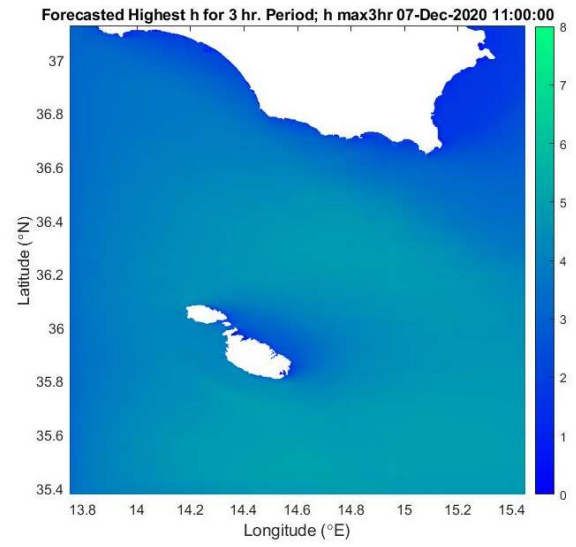
Figure 107 h: h $_{max\ 3hr}$ 07/12/2020 at 11.00

Appendix E: Wave Data at Extreme Events

	Date			Hs / m	Tm/ sec	Steepness	Wave length / m	Dir. / dig
Year	Month	Day	Hour					
2011	12	21	11	5.00	6.29	0.08	61.76	260
2011	12	25	13	3.72	5.93	0.07	54.96	287
2011	12	27	4	4.34	6.05	0.08	57.16	60
2011	12	31	4	5.29	6.58	0.08	67.52	293
2012	1	6	13	6.56	7.45	0.08	86.76	286
2012	2	6	18	5.06	6.35	0.08	63.01	294
2012	2	16	9	4.25	6.11	0.07	58.19	279
2012	3	10	5	6.10	7.21	0.08	81.15	60
2012	12	3	12	5.00	6.75	0.07	71.21	285
2012	12	9	12	5.48	6.67	0.08	69.46	301
2013	1	18	10	5.22	6.75	0.07	71.16	275
2013	1	25	17	5.18	6.46	0.08	65.19	276
2013	2	8	11	4.68	6.60	0.07	67.93	279
2013	3	15	11	6.96	7.85	0.07	96.24	273
2013	3	21	14	4.01	5.91	0.07	54.50	294
2013	11	6	5	4.28	6.26	0.07	61.09	292
2013	11	19	23	3.59	5.42	0.08	45.86	221
2013	12	1	20	3.95	5.77	0.08	52.03	275
2014	1	21	12	4.48	6.12	0.08	58.51	272
2014	1	25	5	6.83	7.15	0.09	79.87	276
2014	1	28	12	4.27	5.76	0.08	51.70	296
2014	2	2	20	3.62	5.56	0.08	48.19	307
2014	2	12	5	3.58	5.95	0.06	55.24	280
2014	3	5	11	3.78	5.86	0.07	53.66	270
2014	12	29	0	4.55	6.51	0.07	66.06	268
2015	1	1	12	4.00	5.82	0.08	52.81	38
2015	1	4	16	3.69	5.68	0.07	50.34	288
2015	1	24	5	4.75	6.58	0.07	67.66	277
2015	2	2	14	4.17	6.28	0.07	61.52	275
2015	2	9	13	4.62	5.93	0.08	54.89	324
2015	2	25	4	6.38	7.72	0.07	92.99	275
2015	3	6	1	5.61	7.26	0.07	82.27	277
2015	3	27	15	4.51	6.37	0.07	63.44	287
2015	11	27	11	5.71	7.21	0.07	81.10	272
2016	1	7	6	4.84	6.50	0.07	65.87	279
2016	1	17	7	6.03	7.00	0.08	76.53	302
2016	2	4	18	4.00	5.80	0.08	52.56	300
2016	2	18	12	4.20	5.85	0.08	53.38	294
2016	3	4	4	4.54	6.53	0.07	66.52	282
2016	3	23	16	4.75	6.34	0.08	62.83	250
2016	11	9	13	4.40	6.11	0.08	58.32	282

2016	12	17	16	5.82	7.46	0.07	86.92	63
2017	1	6	17	4.70	6.08	0.08	57.74	333
2017	1	16	2	3.81	5.58	0.08	48.68	265
2017	2	7	3	5.83	7.56	0.07	89.19	272
2017	3	7	21	6.73	7.56	0.08	89.19	288
2017	11	27	4	3.59	5.67	0.07	50.27	290
2017	12	21	21	4.05	6.18	0.07	59.61	58
2017	12	29	0	4.08	6.23	0.07	60.69	272
2018	1	2	19	4.93	6.55	0.07	66.90	288
2018	2	10	13	4.23	5.95	0.08	55.29	299
2018	3	23	13	4.36	6.33	0.07	62.65	280
2018	11	27	23	5.67	7.14	0.07	79.55	278
2018	12	10	9	3.96	5.97	0.07	55.66	287
2018	12	15	20	3.88	5.67	0.08	50.23	296
2018	12	18	1	4.58	6.30	0.07	61.97	288
2019	1	10	2	4.98	6.43	0.08	64.64	273
2019	1	14	18	5.29	6.61	0.08	68.28	283
2019	1	25	17	4.86	6.43	0.08	64.52	291
2019	1	31	1	4.38	6.87	0.06	73.59	277
2019	2	24	10	5.18	6.62	0.08	68.39	52
2019	3	12	9	4.51	6.19	0.08	59.84	301
2019	3	21	4	3.68	5.48	0.08	46.82	61

Table 26 : Wave data for the extreme events, Model data, period 2011 / 2019

Appendix F: Data Extracted from Distributions

Sorted hs	Rank		Gringorten		Gumbell		Weibull		GPD	
			Exc.	CDF	Exc.	CDF	Exc.	CDF	Exc.	CDF
7.42	1		0.01	0.99	0.02	0.98	0.02	0.98	0.00	1.00
7.29	2		0.02	0.98	0.03	0.97	0.02	0.98	0.01	0.99
7.18	3		0.04	0.96	0.03	0.97	0.02	0.98	0.02	0.98
7.00	4		0.06	0.94	0.04	0.96	0.03	0.97	0.04	0.96
6.81	5		0.07	0.93	0.05	0.95	0.05	0.95	0.06	0.94
6.51	6		0.09	0.91	0.07	0.93	0.07	0.93	0.11	0.89
6.44	7		0.10	0.90	0.08	0.92	0.08	0.92	0.12	0.88
6.22	8		0.12	0.88	0.11	0.89	0.12	0.88	0.16	0.84
6.21	9		0.14	0.86	0.11	0.89	0.12	0.88	0.17	0.83
6.10	10		0.15	0.85	0.13	0.87	0.14	0.86	0.19	0.81
6.05	11		0.17	0.83	0.13	0.87	0.14	0.86	0.20	0.80
5.99	12		0.18	0.82	0.15	0.85	0.16	0.84	0.21	0.79
5.85	13		0.20	0.80	0.17	0.83	0.19	0.81	0.24	0.76
5.65	14		0.21	0.79	0.22	0.78	0.24	0.76	0.29	0.71
5.64	15		0.23	0.77	0.22	0.78	0.24	0.76	0.29	0.71
5.57	16		0.25	0.75	0.24	0.76	0.26	0.74	0.31	0.69
5.53	17		0.26	0.74	0.26	0.74	0.27	0.73	0.32	0.68
5.53	18		0.28	0.72	0.26	0.74	0.28	0.72	0.32	0.68
5.39	19		0.29	0.71	0.30	0.70	0.32	0.68	0.36	0.64
5.34	20		0.31	0.69	0.32	0.68	0.34	0.66	0.37	0.63
5.34	21		0.33	0.67	0.32	0.68	0.34	0.66	0.38	0.62
5.32	22		0.34	0.66	0.33	0.67	0.35	0.65	0.38	0.62
5.26	23		0.36	0.64	0.35	0.65	0.37	0.63	0.40	0.60
5.19	24		0.37	0.63	0.38	0.62	0.39	0.61	0.42	0.58
5.17	25		0.39	0.61	0.39	0.61	0.40	0.60	0.42	0.58
5.07	26		0.40	0.60	0.43	0.57	0.44	0.56	0.45	0.55
5.06	27		0.42	0.58	0.43	0.57	0.44	0.56	0.45	0.55
5.01	28		0.44	0.56	0.45	0.55	0.47	0.53	0.47	0.53
4.99	29		0.45	0.55	0.46	0.54	0.48	0.52	0.48	0.52
4.93	30		0.47	0.53	0.49	0.51	0.50	0.50	0.50	0.50
4.89	31		0.48	0.52	0.51	0.49	0.52	0.48	0.51	0.49
4.86	32		0.50	0.50	0.53	0.47	0.53	0.47	0.52	0.48
4.84	33		0.52	0.48	0.53	0.47	0.54	0.46	0.52	0.48
4.82	34		0.53	0.47	0.55	0.45	0.55	0.45	0.53	0.47
4.81	35		0.55	0.45	0.55	0.45	0.55	0.45	0.53	0.47
4.78	36		0.56	0.44	0.57	0.43	0.57	0.43	0.54	0.46
4.70	37		0.58	0.42	0.61	0.39	0.61	0.39	0.57	0.43
4.67	38		0.60	0.40	0.62	0.38	0.62	0.38	0.58	0.42
4.65	39		0.61	0.39	0.63	0.37	0.63	0.37	0.58	0.42
4.63	40		0.63	0.37	0.64	0.36	0.64	0.36	0.59	0.41
4.57	41		0.64	0.36	0.67	0.33	0.66	0.34	0.61	0.39

4.56	42		0.66	0.34	0.68	0.32	0.67	0.33	0.62	0.38
4.54	43		0.67	0.33	0.69	0.31	0.68	0.32	0.62	0.38
4.51	44		0.69	0.31	0.70	0.30	0.69	0.31	0.63	0.37
4.48	45		0.71	0.29	0.72	0.28	0.70	0.30	0.64	0.36
4.45	46		0.72	0.28	0.73	0.27	0.72	0.28	0.65	0.35
4.35	47		0.74	0.26	0.78	0.22	0.76	0.24	0.69	0.31
4.32	48		0.75	0.25	0.79	0.21	0.78	0.22	0.70	0.30
4.28	49		0.77	0.23	0.81	0.19	0.79	0.21	0.71	0.29
4.27	50		0.79	0.21	0.81	0.19	0.80	0.20	0.72	0.28
4.27	51		0.80	0.20	0.81	0.19	0.80	0.20	0.72	0.28
4.23	52		0.82	0.18	0.83	0.17	0.82	0.18	0.73	0.27
4.21	53		0.83	0.17	0.84	0.16	0.82	0.18	0.74	0.26
4.14	54		0.85	0.15	0.87	0.13	0.85	0.15	0.76	0.24
4.06	55		0.86	0.14	0.89	0.11	0.88	0.12	0.79	0.21
4.03	56		0.88	0.12	0.90	0.10	0.89	0.11	0.81	0.19
3.97	57		0.90	0.10	0.92	0.08	0.91	0.09	0.83	0.17
3.94	58		0.91	0.09	0.93	0.07	0.92	0.08	0.84	0.16
3.92	59		0.93	0.07	0.93	0.07	0.93	0.07	0.85	0.15
3.86	60		0.94	0.06	0.95	0.05	0.95	0.05	0.87	0.13
3.83	61		0.96	0.04	0.96	0.04	0.96	0.04	0.88	0.12
3.83	62		0.98	0.02	0.95	0.05	0.96	0.04	0.88	0.12
3.82	63		0.99	0.01	0.96	0.04	0.96	0.04	0.89	0.11

Sorted Hs	Probability Distribution Function			Return Period			
	Gumbell	Weibull	GPD	Gringorten Equation	Gumbell	Weibull	GPD
7.42	0.03	0.03	0.05	14.09	5.84	7.78	28.15
7.29	0.04	0.04	0.07	5.06	4.83	6.03	9.75
7.18	0.04	0.04	0.09	3.08	4.19	5.00	5.88
7.00	0.05	0.06	0.11	2.22	3.29	3.67	3.18
6.81	0.07	0.08	0.14	1.73	2.53	2.65	1.96
6.51	0.10	0.11	0.17	1.42	1.70	1.67	1.15
6.44	0.11	0.12	0.18	1.20	1.54	1.49	1.03
6.22	0.14	0.16	0.20	1.04	1.15	1.08	0.76
6.21	0.14	0.16	0.20	0.92	1.14	1.07	0.76
6.10	0.16	0.18	0.21	0.83	0.99	0.92	0.66
6.05	0.17	0.19	0.22	0.75	0.93	0.86	0.63
5.99	0.18	0.21	0.22	0.68	0.86	0.80	0.59
5.85	0.22	0.24	0.23	0.63	0.72	0.66	0.51
5.65	0.27	0.28	0.25	0.58	0.56	0.52	0.43
5.64	0.27	0.28	0.25	0.54	0.56	0.52	0.42
5.57	0.29	0.30	0.26	0.51	0.51	0.47	0.40
5.53	0.30	0.31	0.26	0.48	0.49	0.46	0.39
5.53	0.30	0.31	0.26	0.45	0.49	0.45	0.39
5.39	0.34	0.34	0.27	0.43	0.42	0.39	0.35
5.34	0.36	0.35	0.28	0.40	0.39	0.37	0.33
5.34	0.36	0.35	0.28	0.38	0.39	0.37	0.33
5.32	0.37	0.36	0.28	0.37	0.38	0.36	0.33
5.26	0.39	0.37	0.28	0.35	0.36	0.34	0.32
5.19	0.41	0.39	0.29	0.33	0.33	0.32	0.30
5.17	0.41	0.39	0.29	0.32	0.32	0.31	0.30
5.07	0.44	0.41	0.30	0.31	0.29	0.28	0.28
5.06	0.44	0.41	0.30	0.30	0.29	0.28	0.28
5.01	0.46	0.42	0.30	0.29	0.28	0.27	0.27
4.99	0.46	0.42	0.31	0.28	0.27	0.26	0.26
4.93	0.48	0.43	0.31	0.27	0.25	0.25	0.25
4.89	0.48	0.44	0.31	0.26	0.24	0.24	0.25
4.86	0.49	0.44	0.32	0.25	0.24	0.23	0.24
4.84	0.49	0.44	0.32	0.24	0.23	0.23	0.24
4.82	0.50	0.45	0.32	0.24	0.23	0.23	0.24
4.81	0.50	0.45	0.32	0.23	0.23	0.23	0.23
4.78	0.50	0.45	0.32	0.22	0.22	0.22	0.23
4.70	0.51	0.46	0.33	0.22	0.21	0.21	0.22
4.67	0.51	0.46	0.33	0.21	0.20	0.20	0.22
4.65	0.51	0.46	0.33	0.20	0.20	0.20	0.21
4.63	0.51	0.46	0.33	0.20	0.20	0.20	0.21
4.57	0.50	0.46	0.34	0.19	0.19	0.19	0.20

4.56	0.50	0.46	0.34	0.19	0.18	0.19	0.20
4.54	0.50	0.46	0.34	0.19	0.18	0.18	0.20
4.51	0.50	0.46	0.34	0.18	0.18	0.18	0.20
4.48	0.49	0.46	0.35	0.18	0.17	0.18	0.19
4.45	0.49	0.45	0.35	0.17	0.17	0.17	0.19
4.35	0.46	0.44	0.35	0.17	0.16	0.16	0.18
4.32	0.45	0.44	0.36	0.17	0.16	0.16	0.18
4.28	0.44	0.43	0.36	0.16	0.15	0.16	0.18
4.27	0.43	0.43	0.36	0.16	0.15	0.16	0.17
4.27	0.43	0.43	0.36	0.16	0.15	0.16	0.17
4.23	0.41	0.42	0.36	0.15	0.15	0.15	0.17
4.21	0.41	0.42	0.36	0.15	0.15	0.15	0.17
4.14	0.37	0.40	0.37	0.15	0.14	0.15	0.16
4.06	0.33	0.37	0.38	0.14	0.14	0.14	0.16
4.03	0.31	0.36	0.38	0.14	0.14	0.14	0.16
3.97	0.27	0.33	0.38	0.14	0.14	0.14	0.15
3.94	0.26	0.32	0.38	0.14	0.13	0.14	0.15
3.92	0.25	0.31	0.39	0.13	0.13	0.13	0.15
3.86	0.21	0.28	0.39	0.13	0.13	0.13	0.14
3.83	0.19	0.26	0.39	0.13	0.13	0.13	0.14
3.83	0.19	0.26	0.39	0.13	0.13	0.13	0.14
3.82	0.19	0.26	0.39	0.13	0.13	0.13	0.14

Battjes Method – Pr. Nr. of waves exceeding 'h;

h	Probable nr of waves exceeding h		h	Probable nr of waves exceeding h
0.25	13872355		5.25	17404
0.75	6032721		5.75	9985
1.25	2620102		6.25	5703
1.75	1215337		6.75	3235
2.25	595489		7.25	1820
2.75	307992		7.75	1013
3.25	166428		8.25	557
3.75	92841		8.75	301
4.25	52782		9.25	161
4.75	30276		9.75	84

Appendix G: Programs - Validation

Program 1 – To find model grid point closest to buoy
Program Name: *Grid point closest to buoy coordinate*

```matlab
% Finding Swan data at point closest to the buoy and EXTRACT THE DATA
% NECCESSARY
% coordinates of mooring of buoy :
%       Longitude : 14.16888 ;
%       Latitude : 36.08152


% DATA COLLECTED EVERY HOUR ON THE HOUR

%we set the date. we must nake it variable with the data
YEAR = 2020;
MONTH =1;
DATE = 1;


StartDate = datenum(YEAR, MONTH, DATE, 0, 0, 0);


inf = ncinfo('output0101.nc'); % info on the file
Hs = ncread('output0101.nc','hs'); % significant wave height
PPer = ncread('output0101.nc','tp'); % peak wave period
Dir = ncread('output0101.nc','theta0'); %wave direction (direction of
propagation)
Time = ncread('output0101.nc', 'time'); % time vector
lat = ncread ('output0101.nc','latitude'); %Latitude of data
lon = ncread('output0101.nc','longitude'); % Longitude of data


% converts time vector at the start date
defDate = datenum([1970 1 1 0 0 0]);


% numel - number of elements in array time; 61
% Time is a matrix in seconds from initial date defDate-01 01 1970
% adding Time(t) in 'seconds' to defDate cause Time is in sec

for t = 1:1:numel(Time)
    TimeSeries(t) = addtodate(defDate, Time(t), 'second');
end

% coordinates of buoy : longitude : 14.16888 ; Latitude : 36.08152
lonB= 14.16888;
latB= 36.08152;

% find closest point minimum value and absolute
% XMinVal is the least value of min(abs(lon - lonB)
% xMinIdx is the index of the point
% abs - absolute ; since it can be positice and negative

[xMinVal xMinIdx] = min(abs(lon - lonB));
[yMinVal yMinIdx] = min(abs(lat - latB));


t1 = TimeSeries(1,1);
name= ['Model_Data_',num2str(t1)];

% COLLECTING DATA IN ONE MATRIX
% timeseries all values of Hs at that point for all t
```

```matlab
data=zeros(61,4);

data(:,1) = TimeSeries(1,:); % Time
data(:,2) = Hs(xMinIdx, yMinIdx, :); % Sig Wave Height Hs
data(:,3) = PPer(xMinIdx, yMinIdx, :); % peak wave period
data(:,4) = Dir(xMinIdx, yMinIdx, :); % wave direction (direction of
propagation)

data = squeeze(data);

eval([name,'=data;']);

% closest coordinates
lonC = num2str(lon(xMinIdx));
latC =  num2str(lat(yMinIdx));

% plot(TimeSeries, timeSeriesHs, '-');
% dateaxis('x', 0);
% title(['Hs for lon: ', num2str(lon(xMinIdx)), '\circE, lat: ',
num2str(lat(yMinIdx)), '\circN']);
```

Program 2 – To determine Linear Regression Fit

Program Name*:* Fit2.m

```matlab
% TO DETERMINE CLOSENESS OF FIT Hs Buoy
% VERY IMPORTANT - USED IN PROGRAM "VALIDATION SKIRON_BUOY"

% Function - [outputs,output 2] = function name (input1, input2)

function [RValue,M,RSquared,Inter, ErEst] = Fit2(HsModel,HsBuoy)

if isempty(HsModel)
    RValue = nan;
    M=nan;
    RSquared=nan;
    Inter=nan;
    ErEst= nan
else

%%% To fit line with least squares y = mx + a
% m = summation (x- x mean) * (y-ymean)/ summation (x- x mean)^2
Xmean = mean(HsModel);
Ymean = mean(HsBuoy);

for L=1:1:numel(HsModel(:,1));
   ValX(L,1) = HsModel(L,1) - Xmean; % = x- xmean
   ValY(L,1) = HsBuoy(L,1) - Ymean; % = y - ymean
   ValXY(L,1) =  ValX(L,1)* ValY(L,1); % = (x- xmean) * (y - ymean)
   ValXSq(L,1) = ValX(L,1)*ValX(L,1); % = (x- x mean)^2
   ValYSq(L,1) = ValY(L,1)*ValY(L,1); % = (y- y mean)^2
end

TotValXY = sum(ValXY); % Sum (x- xmean) * (y - ymean)
TotValXSq = sum(ValXSq); % Sum(x- x mean)^2
TotValYSq = sum(ValYSq); % Sum (y - y mean)^2

SqrtTotValXSq = sqrt(TotValXSq); % = SQRT [Sum(x- x mean)^2]
```

```matlab
SqrtTotValYSq = sqrt(TotValYSq); % = SQRT [Sum(y - y mean)^2]

M = TotValXY / TotValXSq ; % = Sum [(x- xmean) * (y - ymean)] / [Sum(x- x
mean)^2]

% to evaluate intercept A = Ymean - b*Xmean
Inter = Ymean - (Xmean * M);

%%%% To determine R Squared using the least square method
% rsquared = 1 - Sum (y - y estimated)^2 / Sum (y - y mean)^2

for L=1:1:numel(HsModel(:,1));
 YEst(L,1) = (M * HsModel(L,1))+Inter; %  y estimated
 YLE(L,1) = (HsBuoy(L,1) - YEst(L,1)) ; % YlessEst. = (y - y estimated)
 YLES(L,1)= YLE(L,1)*YLE(L,1); % (y - y estimated)^2
end

TotYLES = sum(YLES); % sum( (y - y estimated)^2
RSquared = 1 - (TotYLES / TotValYSq) ; % Rsq = 1 - Sum (y - y estimated)^2
/ Sum (y - y mean)^2

n = numel(HsModel(:,1));
ErEst = sqrt(TotYLES/n);

%%%%% to determine R = Sum[(x- xmean) * (y - ymean)] / SQRT [Sum(x- x
mean)^2]* SQRT [Sum(y - y mean)^2]
RValue= TotValXY / (SqrtTotValXSq *SqrtTotValYSq);

end
```

Program 3 – Determine the Linear regression Lines
Program Name: Validation Skiron_Buoy .m

```matlab
% to extract Data from Time series Rosario Swan Skiron and Buoy Data
% To find vakues for same date
% to find LINEAR CORRELATION USING THE LEAST SQUARE METHOD


%TO READ SKIRON SWAN DATA COORDINATES : LON : 14.168 LAT: 36.0812
dataSwanSk = dlmread('timeseries_Skiron2011_2012.txt');
SkDate = datenum([dataSwanSk(:,1) dataSwanSk(:,2) dataSwanSk(:,3)
dataSwanSk(:,4) zeros(size(dataSwanSk, 1), 1) zeros(size(dataSwanSk, 1),
1)]);
SkHs = dataSwanSk(:,5); % Sea surface Wave significant height
SkDir = dataSwanSk(:,6); % Sea surface wave FROM direction
SkTp = dataSwanSk(:,7); % Wave Peak Period
SkTm = dataSwanSk(:,8); % sea surface wave MEAN period from variance
spectral density first frequency moment

% TO READ BUOYDATA
buoydata = dlmread('buoydata.txt'); % to extract the files

% TO REMOVE DATA THAT ARE NOT CORRECT - which are not correct data. i did
% <0.005
idx = find(buoydata(:,8)< 0.005);
buoydata(idx,8) = nan;
```

```matlab
idx2 = find(buoydata(:,9)< 0.005);
buoydata(idx,9) = nan;

BD = isnan(buoydata);
index1 = find (BD ==1);
buoydata(index1) = 0;

% CREATE MATRIX BuoyDATA having Datenum , HS ,Tz
a=numel(buoydata(:,1));
BuoyDATA = zeros(a,3);

for i = 1:1:size(buoydata, 1)
BuoyDATA(i,1) = datenum([buoydata(i,4) buoydata(i,3) buoydata(i,2)
buoydata(i,5) buoydata(i,6) buoydata(i,7)]); % Date
BuoyDATA(i,2) = (buoydata(i,8)/100); % Hs in M
BuoyDATA(i,3) = buoydata(i,9); % T zero
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PLOTTING TIME SERIES
% plotting the Buoy time series
figure
hold on;
plot(BuoyDATA(:,1), BuoyDATA(:,2), 'b', 'LineWidth', 1.5);
xticks([min(BuoyDATA(:,1)):90:max(BuoyDATA(:,1))]);
dateaxis('x',1);

axis([ min(BuoyDATA(:,1)) max(BuoyDATA(:,1)) 0 max(BuoyDATA(:,2)+1)]);
title(['SIGNIFICANT WAVE HEIGHT Hs Time Series- In
Situ'],'FontSize',14,'FontWeight','bold');

ax = gca;
ax.FontSize = 12;
xlabel('Date','FontSize',12,'FontWeight','bold');
ylabel('Hs / m','FontSize',12,'FontWeight','bold');

legend('Buoy Data','Location','northeast');
box on;
hold off

% 2. Hs model and Hs Buoy
figure
hold on;
plot(BuoyDATA(:,1), BuoyDATA(:,2), 'r', 'LineWidth', 1.5);
plot(SkDate, SkHs, 'b', 'LineWidth', 1.5);
% xticks([min(BuoyDATA(:,1)):60:max(BuoyDATA(:,1))]);

axis([ min(SkDate(:,1)) max(SkDate(:,1)) 0 max(BuoyDATA(:,2)+0.5)]);
dateaxis('x',1);
title(['SIGNIFICANT WAVE HEIGHT - In Situ / Rosario Skiron Swan
'],'FontSize',14,'FontWeight','bold');
ax = gca;
ax.FontSize = 12;
xlabel('Date','FontSize',12,'FontWeight','bold');
ylabel('Hs / m','FontSize',12,'FontWeight','bold');

legend('Buoy Data', 'Model Rosario Skiron Swan','Location','northeast');
box on;
hold off
```

153

```matlab
% TO FIND GAPS IN DATA MORE THAN 5 DAYS - SAVE BUOYDATAT
TT1 = datenum([0 0 5 0 0 0]);
BuoyDATAT = BuoyDATA ;

for n=2:1:numel(BuoyDATAT(:,1));
    m = n-1 ;
    if (BuoyDATAT(n,1)-BuoyDATAT(m,1) >= TT1);
       BuoyDATAT(m,1) = BuoyDATAT(m,1);
    else
        BuoyDATAT(m,1) = 0;
    end
end

indexT = find(BuoyDATAT(:,1) > 0);
gap = zeros(numel(indexT),1);
gap(:,1) = BuoyDATAT(indexT,1);

date1 = zeros(numel(indexT),6);
for i=1:1:numel(indexT);
date1(i,:) = datevec(gap(i,1));
end

% 2a. Time series Tz - Buoy
figure;
plot(BuoyDATA(:,1), BuoyDATA(:,3), 'b', 'LineWidth', 1.5);
hold on;
axis([ min(BuoyDATA(:,1)) max(BuoyDATA(:,1)) 0 max(BuoyDATA(:,3)+1)]);

xticks([min(BuoyDATA(:,1)):120:max(BuoyDATA(:,1))]);
dateaxis('x',1);
title(['Zero Upcross Period Tz - In Situ Buoy
Data'],'FontSize',14,'FontWeight','bold');

ax = gca;
ax.FontSize = 12;
xlabel('Date','FontSize',12,'FontWeight','bold');
ylabel('Tz / s','FontSize',12,'FontWeight','bold');
legend('Buoy
Data','Location','northeast','FontSize',10,'FontWeight','bold');
box on;
hold off

% 2b. Plot Time series Tz - Buoy / Skiron Tm
figure;
plot(BuoyDATA(:,1), BuoyDATA(:,3), 'b', 'LineWidth', 1.5);
hold on;

plot(SkDate, SkTm, 'r', 'LineWidth', 1.5);
xticks([min(SkDate):60:max(SkDate)]);
ax = gca;
ax.FontSize = 12;

dateaxis('x',1);
axis([ min(SkDate) max(SkDate) 0 max(SkTp)]);
title(['Zero Upcross Period Tz - Buoy / Rosario Swan
Skiron'],'FontSize',14,'FontWeight','bold');

xlabel('Date','FontSize',12,'FontWeight','bold');
ylabel('Tz / s','FontSize',12,'FontWeight','bold');
```

```matlab
legend('Buoy Data', 'Model Rosario Skiron','Location','northeast');
box on;
hold off

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% TO FIND BUOY AND MODEL DATA FOR SAME PERIOD AND PUT IN MATRIX
% "BuoyDataSK"
% Model data every hour Buoy Data every 30 mins
BuoyDataSK = nan(size(SkDate,1), 6); % creating matrix for both Buoy and
Model Data

for d = 1:1:size(SkDate, 1)
    currDate = SkDate(d);
    % minIdx will give you the index of BuoyData which is closest to model
    % in time
    [minVal minIdx] = min(abs(currDate - BuoyDATA(:,1)));

    % if gap is less than 20 minutes take point . ELSE IT WILL GIVE NAN
    % BuoyDataSK - 1. date    2. Buoy Hs    3. Buoy Tz    4. Swan Hs
    % 5. Swan dir    6. Swan Tm
    if (minVal < 0.0139)
        BuoyDataSK(d,:) = [currDate BuoyDATA(minIdx,2) BuoyDATA(minIdx,3)
SkHs(d,1) SkDir(d,1) SkTm(d,1)];
    end

end

% TO REMOVE NAN's FROM ALL THE DATA
    index = find(BuoyDataSK(:,1)> 0) ;
    BuoyDataSkT= zeros(numel(index), 6);
    BuoyDataSkT(:,:)= BuoyDataSK(index,:);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% VALIDATION USING LEAST SQUARES METHOD
% 1.  HS BASED ON ALL DATA
HsSkFD = BuoyDataSkT(:,4); % X AXIS
HsBuoyFD = BuoyDataSkT(:,2);  % Y AXIS

% TO USE THE FUNCTION FIT 2 TO EXTRACT RTOTAL ETC
[RTotSk(1,1),MTotSk(1,1),RSTotSk(1,1), IntTotSk(1,1),ErEstTotSk(1,1)] =
Fit2(HsSkFD,HsBuoyFD); %FULL DATA SERIES

% TO PLOT DATA Model against Buoy
figure;
plot (HsSkFD, HsBuoyFD,'b*');  % Plotting Data hs
title(['Correlation Hs- Skiron Swan Model / In situ
Buoy'],'FontSize',14,'FontWeight','bold');

ax = gca;
ax.FontSize = 12;
xlabel('Hs Rosario Skiron Swan Model /
m','FontSize',12,'FontWeight','bold');
ylabel('Hs Buoy Data / m','FontSize',12,'FontWeight','bold');

hold on;

% TO PLOT THE ESTIMATE USING Y = MX + INTERCEPT
for L=1:1:numel(HsSkFD(:,1));
 % TO GET VALUES OF Y ESTIMATED -
```

```matlab
 HsEstSk(L,1) = (MTotSk(1,1) * HsSkFD(L,1))+IntTotSk(1,1);
end


%%%% TO PLOT Y ESTIMATED - the line with least squares
plot (HsSkFD, HsEstSk,'r-','LineWidth',1.5);
axis equal
box on
axis([0 5 0 7]);    % xlim([0 5]);% ylim([0 8]);

%%%%% TO PLOT THE 45 deg line
XTheor = [0 1 2 3 4  5 6 7 8];
YTheor = [0 1 2 3 4  5 6 7 8];
plot(XTheor,YTheor,'m-','LineWidth',1.25);


m = MTotSk;
n = IntTotSk;
legend('Rosario Skiron Swan',sprintf('Y = %.2f * X + %.2f',m,n),'Y = X',
'Location','southeast' );


hold off;

% 2.  TZ BASED ON ALL DATA - VALIDATION ON ALL DATA
TmSkFD = BuoyDataSkT(:,6); % X AXIS
TzBuoyFD = BuoyDataSkT(:,3);  % Y AXIS

% TO PLOT DATA
figure;
plot (TmSkFD,TzBuoyFD ,'b*');  % Plotting Data hs
title(['Correlation Tz -Skiron Swan Model / In situ
Buoy'],'FontSize',14,'FontWeight','bold');


ax = gca;
ax.FontSize = 12;
xlabel('Tm Rosario Skiron Swan Model /
s','FontSize',12,'FontWeight','bold');
ylabel('Tz Buoy Data / s','FontSize',12,'FontWeight','bold');
hold on;

% TO USE THE FUNCTION FIT 2 TO EXTRACT RTOTAL ETC
[RToSkTz(1,1),MToSkTz(1,1),RSToSkTz(1,1), IntToSkTz(1,1),ErEstTotSkTz(1,1)
] = Fit2(TmSkFD,TzBuoyFD); %FULL DATA SERIES

% TO PLOT THE ESTIMATE USING Y = MX + INTERCEPT
for L=1:1:numel(TmSkFD(:,1));
 TzEstSk(L,1) = (MToSkTz(1,1) * TmSkFD(L,1))+IntToSkTz(1,1); %  y estimated
end

%%%% To plot the line witha least squares
plot (TmSkFD, TzEstSk,'r-','LineWidth',1.5);
axis equal
box on
axis([0 8 0 9]);    % xlim([0 5]);% ylim([0 8]);

%%%%% to plot the theoretical 45 deg line

plot(XTheor,YTheor,'m-','LineWidth',1.25);

mT = MToSkTz;
nT = IntToSkTz;
```

```matlab
legend('Rosario Skiron Swan',sprintf('Y = %.2f * X + %.2f',mT,nT),'Y = X', ...
'Location','southeast' );
hold off;


% TO SORT DATA BY ANGLE
% 1. to put in bins
Direction =  BuoyDataSK(:,5); % Column 5 = direction . Check what it means
direction **

dDir = [];

namel = ['DataDir_',num2str(0)] ; % Naming of directory
namel = strrep(namel, '.', '');

idx3 = find(Direction >= 337.5);
idx4 = find(Direction < 22.5);
idx5 = union(idx3, idx4);
lenIdx0 = length(idx5);

% Number of counts
NrC(8,1) = lenIdx0;
dDir =  BuoyDataSK(idx5,:);
eval([namel,'= dDir;']);

Degrees = [22.5 67.5 112.5 157.5 202.5 247.5 292.5 337.5]; % 22.5
DegreeC = [45 90 135 180 225 270 315];

for N=1:1:length(Degrees)-1; %-1
    % The central value of the range
    angleC = DegreeC(N);
    namel = ['DataDir_',num2str(angleC)] ; % Naming of directory
    namel = strrep(namel, '.', ''); % to remove decimal place else error

    idx1 = find(Direction >= Degrees(N));
    idx2 = find(Direction < Degrees(N+1));
    idx = intersect(idx1, idx2);
    lenIdx = length(idx);

    % to put in counts
    NrC(N,1) = lenIdx; % Number of Counts

    dDir =  BuoyDataSK(idx,:); % : means all columns
    % BuoyDATASWAN - 1. date    2. Buoy Hs    3. Buoy Tz    4. Swan Hs
    % 5. Swan dir    6. Swan Tm
    eval([namel,'= dDir;']);
end

NrCTot = sum(NrC,'all');

DegreeC2 = [45 90 135 180 225 270 315 0];

figure;
bar(DegreeC2(1,:),NrC(:,1)*100/NrCTot);
hold on
title(['Histogram - Percentage Occurance / Wave
Direction.'],'FontSize',14,'FontWeight','bold');
```

```matlab
ax = gca;
ax.FontSize = 12;
xlabel('Wave Direction - From
direction','FontSize',12,'FontWeight','bold');
ylabel('Count / % ','FontSize',12,'FontWeight','bold');



% FOR THE HS DATA - TO USE THE FUNCTION FIT 2 TO EXTRACT VALUES OF
% R , M RSQAURED AND INTERSECTION

% FOR THE HS DATA SORTED IN ANGLES
for i = 1:1:8
    RSk(i,1) = DegreeC2(1,i);
    MSk(i,1) = DegreeC2(1,i);
    RSSk(i,1) = DegreeC2(1,i);
    IntSk(i,1)= DegreeC2(1,i);
    ErEstSk(i,1)= DegreeC2(1,i);
end

[RSk(1,2),MSk(1,2),RSSk(1,2),IntSk(1,2),ErEstSk(1,2)] =
Fit2(DataDir_45(:,4),DataDir_45(:,2));
[RSk(2,2),MSk(2,2),RSSk(2,2),IntSk(2,2),ErEstSk(2,2)] =
Fit2(DataDir_90(:,4),DataDir_90(:,2));
[RSk(3,2),MSk(3,2),RSSk(3,2),IntSk(3,2),ErEstSk(3,2)] =
Fit2(DataDir_135(:,4),DataDir_135(:,2));
[RSk(4,2),MSk(4,2),RSSk(4,2),IntSk(4,2),ErEstSk(4,2)] =
Fit2(DataDir_180(:,4),DataDir_180(:,2));

[RSk(5,2),MSk(5,2),RSSk(5,2),IntSk(5,2),ErEstSk(5,2)] =
Fit2(DataDir_225(:,4),DataDir_225(:,2));
[RSk(6,2),MSk(6,2),RSSk(6,2),IntSk(6,2),ErEstSk(6,2)] =
Fit2(DataDir_270(:,4),DataDir_270(:,2));
[RSk(7,2),MSk(7,2),RSSk(7,2),IntSk(7,2),ErEstSk(7,2)] =
Fit2(DataDir_315(:,4),DataDir_315(:,2));
[RSk(8,2),MSk(8,2),RSSk(8,2),IntSk(8,2),ErEstSk(8,2)] =
Fit2(DataDir_0(:,4),DataDir_0(:,2));

%%% FOR THE TZ DATA - TO USE THE FUNCTION FIT 2 TO EXTRACT VALUES OF
% R , M RSQAURED AND INTERSECTION
% FOR THE TZ DATA SORTED IN ANGLES
for i = 1:1:8
    RTSk(i,1) = DegreeC2(1,i);
    MTSk(i,1) = DegreeC2(1,i);
    RsTSk(i,1) = DegreeC2(1,i);
    IntTSk(i,1) = DegreeC2(1,i);
    ErEstTSk(i,1)= DegreeC2(1,i)
end

[RTSk(1,2),MTSk(1,2),RsTSk(1,2),IntTSk(1,2),ErEstTSk(1,2)] =
Fit2(DataDir_45(:,6),DataDir_45(:,3)); % 45 deg
[RTSk(2,2),MTSk(2,2),RsTSk(2,2),IntTSk(2,2),ErEstTSk(2,2) ] =
Fit2(DataDir_90(:,6),DataDir_90(:,3)); % 90 deg
[RTSk(3,2),MTSk(3,2),RsTSk(3,2),IntTSk(3,2),ErEstTSk(3,2)] =
Fit2(DataDir_135(:,6),DataDir_135(:,3)); % 135 deg
[RTSk(4,2),MTSk(4,2),RsTSk(4,2),IntTSk(4,2),ErEstTSk(4,2)] =
Fit2(DataDir_180(:,6),DataDir_180(:,3)); % 180 deg

[RTSk(5,2),MTSk(5,2),RsTSk(5,2),IntTSk(5,2),ErEstTSk(5,2)] =
Fit2(DataDir_225(:,6),DataDir_225(:,3));
```

158

```
[RTSk(6,2),MTSk(6,2),RsTSk(6,2),IntTSk(6,2),ErEstTSk(6,2)] =
Fit2(DataDir_270(:,6),DataDir_270(:,3));
[RTSk(7,2),MTSk(7,2),RsTSk(7,2),IntTSk(7,2),ErEstTSk(7,2)] =
Fit2(DataDir_315(:,6),DataDir_315(:,3));
[RTSk(8,2),MTSk(8,2),RsTSk(8,2),IntTSk(8,2),ErEstTSk(8,2)] =
Fit2(DataDir_0(:,6),DataDir_0(:,3));
```

Program 4 – To Plot Regression Lines Hs/ Tz for all angles
Program Name: Validation Skiron_BuoyPlots.m

```matlab
Degrees = [22.5 67.5 112.5 157.5 202.5 247.5 292.5 337.5 22.5];
DegreeC = [45 90 135 180 225 270 315 0];
dDir = [];

for N=1:1:length(Degrees)-1;
    % angle = Degrees(N);
    % The central value of the range
    angleC = DegreeC(N);

    %%%Sorting the data by ANGLE RANGE
    idx1 = find(Direction >= Degrees(N));
    idx2 = find(Direction < Degrees(N+1));
    idx3 = intersect(idx1, idx2);

    dDirA = BuoyDataSK(idx3,:); % : means all columns
    % BuoyDATASWAN - 1. date    2. Buoy Hs    3. Buoy Tz    4. Swan Hs
    % 5. Swan dir    6. Swan Tm

    HsModel = dDirA(:,4); % Hs MODEL sorted by angle -  X AXIS
    HsBuoy = dDirA(:,2);  % Hs BUOY sorted by angle - Y AXIS

        figure(N);
        plot (HsModel, HsBuoy,'b*');  % Plotting Data Hs Model / Buoy
        title(['Correlation Hs – Skiron Swan/ Buoy - Dir.
',num2str(angleC),' deg.'],'FontSize',14,'FontWeight','bold');
        hold on

        ax = gca;
        ax.FontSize = 12;
        xlabel('Hs Model Skiron Swan /
m','FontSize',12,'FontWeight','bold');
        ylabel('Hs Buoy / m','FontSize',12,'FontWeight','bold');

        axis equal
         box on
         % axis([ min(swanDate) max(swanDate) 0 max(swanTp)]);
         % axis([0 2 0 3]);

% Gradient and intercept
AHs = IntSk(N,2);
MHs = MSk(N,2);

YEstimate= [];
% TO PLOT THE ESTIMATE USING Y = MX + INTERCEPT
    for L=1:1:numel(HsModel(:,1));
        YEstimate(L,N) = (MHs* HsModel(L,1))+AHs; %  y estimated least
Square method
    end
```

```matlab
    plot(HsModel,YEstimate(:,N)','R-','LineWidth',1.25);

% to plot the 45 deg line
plot(XTheor,YTheor,'m-','LineWidth',1.25);

axis([0 (max(HsModel) +1) 0 (max(HsBuoy)+1) ]);
legend('Rosario Skiron Swan',sprintf('Y = %.2f * X + %.2f',MHs,AHs),'Y = X', 'Location','southeast' );

hold off;
end

% to do the 0 deg !!!
%%%Sorting the data by ANGLE RANGE
idx3a = find(Direction >= 337.5);
idx4 = find(Direction < 22.5);
idx5 = union(idx3a, idx4);
lenIdx0 = length(idx5);

    dDirA0 = BuoyDataSK(idx5,:); % : means all columns
    % BuoyDATASWAN - 1. date    2. Buoy Hs    3. Buoy Tz    4. Swan Hs
    % 5. Swan dir    6. Swan Tm
    HsModel0 = dDirA0(:,4); % Hs MODEL sorted by angle -  X AXIS
    HsBuoy0 = dDirA0(:,2);  % Hs BUOY sorted by angle - Y AXIS

        figure(N);
        plot (HsModel0, HsBuoy0,'b*');  % Plotting Data Hs Model / Buoy
        title(['Correlation Hs - Skiron Swan/ Buoy - Dir. 0
deg.'],'FontSize',14,'FontWeight','bold');
        hold on
        ax = gca;
        ax.FontSize = 12;
        xlabel('Hs Model Skiron Swan /
m','FontSize',12,'FontWeight','bold');
        ylabel('Hs Buoy / m','FontSize',12,'FontWeight','bold');

       axis equal
        box on
        % axis([ min(swanDate) max(swanDate) 0 max(swanTp)]);
        % axis([0 2 0 3]);

% Gradient and intercept
AHs0 = IntSk(8,2);
MHs0 = MSk(8,2);

YEstimate0= [];
% TO PLOT THE ESTIMATE USING Y = MX + INTERCEPT
    for L=1:1:numel(HsModel0(:,1));
        YEstimate0(L,1) = (MHs0* HsModel0(L,1))+AHs0; %  y estimated least
Square method
    end
   plot(HsModel0,YEstimate0(:,1)','R-','LineWidth',1.25);

% to plot the 45 deg line
plot(XTheor,YTheor,'m-','LineWidth',1.25);
axis([0 (max(HsModel0) +1) 0 (max(HsBuoy0)+1) ]);
legend('Rosario Skiron Swan',sprintf('Y = %.2f * X + %.2f',MHs0,AHs0),'Y = X', 'Location','southeast' );
```

```matlab
hold off;

% TO PLOT Tz values of MODEL(X) and ACTAUL (BUOY) (Y)

for N=1:1:length(Degrees)-1;
    % angle = Degrees(N);
    % The central value of the range
    angleC = DegreeC(N);

    %%%Sorting the data by ANGLE RANGE
    idxT1 = find(Direction >= Degrees(N));
    idxT2 = find(Direction < Degrees(N+1));
    idxT = intersect(idxT1, idxT2);

    dDirT =  BuoyDataSK(idxT,:); % : means all columns
    % BuoyDATASWAN - 1. date    2. Buoy Hs    3. Buoy Tz    4. Swan Hs
    % 5. Swan dir    6. Swan Tm

    TmModelSk = dDirT(:,6); % Tm MODEL sorted by angle -  X AXIS
    TzBuoySk = dDirT(:,3);  % Tz BUOY sorted by angle - Y AXIS

        figure(N);
        plot (TmModelSk, TzBuoySk,'b*');  % Plotting Data hs
        title(['Correlation Tz - Skiron Swan/ Buoy - Dir.
',num2str(angleC),' deg.'],'FontSize',14,'FontWeight','bold');
        hold on

        ax = gca;
        ax.FontSize = 12;

        xlabel('Tm Skiron Swan / s','FontSize',12,'FontWeight','bold');
        ylabel('Tz Buoy / s','FontSize',12,'FontWeight','bold');

     axis equal
      box on
      % axis([ min(swanDate) max(swanDate) 0 max(swanTp)]);
      % axis([0 2 0 3]);

    % TO PLOT THE ESTIMATE USING Y = MX + INTERCEPT
    % Gradient and intercept
    ATz = IntTSk(N,2);  % Intersect evaluated from previous program
    MTz = MTSk(N,2);  % Gradient evaluated from previous program

    YEstT= []
    for L=1:1:numel(TmModelSk(:,1));
       YEstT(L,1) = (MTz* TmModelSk(L,1))+ATz; %  y estimated
    end

    plot(TmModelSk, YEstT','R-','LineWidth',1.25);

    % TO PLOT THE 45 DEG LINE
      XTheor = [0 1 2 3 4  5 6 7 8];
      YTheor = [0 1 2 3 4  5 6 7 8];
      plot(XTheor,YTheor,'m-','LineWidth',1.25);

    axis([(min(TmModelSk)-1) (max(TmModelSk)+1) (min(TzBuoySk)-1)
(max(TzBuoySk) +1)]);
```

```matlab
    legend('Rosario Skiron Swan',sprintf('Y = %.2f * X + %.2f',MTz,ATz),'Y
= X', 'Location','southeast' );


hold off;


end


% to do 0 degree
dDirT0 =  BuoyDataSK(idx5,:); % : means all columns
    % BuoyDATASWAN - 1. date    2. Buoy Hs    3. Buoy Tz    4. Swan Hs
    % 5. Swan dir    6. Swan Tm

    TmModelSk0 = dDirT0(:,6); % Tm MODEL sorted by angle -  X AXIS
    TzBuoySk0 = dDirT0(:,3);  % Tz BUOY sorted by angle - Y AXIS

        figure;
        plot (TmModelSk0, TzBuoySk0,'b*');  % Plotting Data hs
        title(['Correlation Tz - Skiron Swan/ Buoy - Dir. 0
deg.'],'FontSize',14,'FontWeight','bold');
        hold on

        ax = gca;
        ax.FontSize = 12;
        xlabel('Tm Skiron Swan / s','FontSize',12,'FontWeight','bold');
        ylabel('Tz Buoy / s','FontSize',12,'FontWeight','bold');

       axis equal
        box on
        % axis([ min(swanDate) max(swanDate) 0 max(swanTp)]);
        % axis([0 2 0 3]);

    % TO PLOT THE ESTIMATE USING Y = MX + INTERCEPT
    % Gradient and intercept
    ATz0 = IntTSk(8,2);  % Intersect evaluated from previous program
    MTz0 = MTSk(8,2);  % Gradient evaluated from previous program

    YEstT0= [];

    for L=1:1:numel(TmModelSk0(:,1));
       YEstT0(L,1) = (MTz0* TmModelSk0(L,1))+ATz0; %  y estimated
    end

    plot(TmModelSk0, YEstT0','R-','LineWidth',1.25);

    % TO PLOT THE 45 DEG LINE
      XTheor = [0 1 2 3 4  5 6 7 8];
      YTheor = [0 1 2 3 4  5 6 7 8];
      plot(XTheor,YTheor,'m-','LineWidth',1.25);
    axis([(min(TmModelSk0)-1) (max(TmModelSk0)+1) (min(TzBuoySk0)-1)
(max(TzBuoySk0) +1)]);
    legend('Rosario Skiron Swan',sprintf('Y = %.2f * X +
%.2f',MTz0,ATz0),'Y = X', 'Location','southeast' );


hold off;
end
```

Appendix H: Programs - Wave Forecasts

Program 1 – To Calculate Wave Length and wave steepness
Program Name: wavelength_steepness.m

```matlab
% Calculate wavelength and wave steepness
% IT HAS ZONE CHARATECERISATION AND PLOTS OF WAVE STEEPNESS

inf = ncinfo('output0101.nc'); % info on the file
Hs = ncread('output0101.nc','hs'); % significant wave height
lat = ncread ('output0101.nc','latitude'); % Latitude of data
lon = ncread('output0101.nc','longitude'); % Longitude of data
MPer = ncread('output0101.nc','tm01'); % mean wave period
PPer = ncread('output0101.nc','tp'); % peak wave period
depth = ncread('output0101.nc','depth'); % depth
Time = ncread('output0101.nc', 'time'); % time
Dir = ncread('output0101.nc','theta0'); %sea surface FROM direction

% converts time vector at the start date
defDate = datenum([1970 1 1 0 0 0]);

% numel - number of elements in array time; 61
% Time is a matrix in seconds from initial date defDate-01 01 1970
% adding Time(t) in 'seconds' to defDate cause Time is in sec

for t = 1:1:numel(Time)
    TimeSeries(t) = addtodate(defDate, Time(t), 'second');
end

%1. TO DETERMINE WAVELENGTH AT EACH COORDINATE
%   creating matrix to evaluate L0 (wavelength at depth)
%   T = mean wave period
%   Lmax = (9.81* MPer^2) /(2*pi);

lenlon = length(lon);
lenlat = length(lat);
t= length(Time);
Lmax = zeros(lenlon,lenlat,t);

for t = 1:1:numel(Time);
    for i=1:1:lenlon;
        for j = 1:1:lenlat;
            Lmax(i,j,t) = (9.81* MPer(i,j,t)^2) /(2*pi);
        end
    end
end

% to Evaluate the wave length based on iterative equation
% Equation wavelength LFinal = (L0) Tanh (2 x pi x Depth / wavelength)
% Units : tanh angles must be in radians ; depth and wavelength in metres
% Matrix LFinal - 3d matrix with calculated wavelengths

LFinal = zeros(lenlon,lenlat,numel(Time));%
NN= zeros(lenlon,lenlat,numel(Time));%

for t = 1:1:numel(Time);
for i=1:1:lenlon;
for j = 1:1:lenlat;
```

```matlab
% iteration to evaluate estimate value ; indefinite number of iterations N
% to evaluate ratio
    L0 = Lmax(i,j,t);
    ratio(i,j,t) = (depth(i,j,t)/L0);

        if ratio(i,j,t)> 0.5 %ratio for deep water
            LFinal(i,j,t)= L0;
            else
                L1 = 0;
                L2 = 0;
                L1 = L0*tanh(2*pi*depth(i,j,t)/L0);
                L2 = L0*tanh(2*pi*depth(i,j,t)/L1);
                n=1;
                    if (abs(L2-L1) <= 0.1)
                    LFinal(i,j,t) = L2;
                    NN(i,j,t) = 1 ;
                    else
                      while abs(L2-L1) > 0.1
                      L1=L2;
                      L2 = L0*tanh(2*pi*depth(i,j,t)/L1);
                      n=n+1;
                      end
                  LFinal(i,j,t) = L2;
                  NN(i,j,t) = n ;
                end
        end
end
end
end

% to create matrix LFinalT (test) not to effect values in L Final
 LFinalT= LFinal;

% To determine RATIO WAVELENGTH /DEPTH for each Lon , Lat and T
% DIFFERENT FORM ABOVE
ratioLD = (zeros(lenlon,lenlat,numel(Time)));

for t = 1:1:numel(Time);
    for i=1:1:lenlon;
        for j = 1:1:lenlat;
            ratioLD(i,j,t) = LFinal(i,j,t)/depth(i,j,t);
        end
    end
end

RatioLDMin= min(min(min(ratioLD)));
RatioLDMax= max(max(max(ratioLD)));

% to produce figures of Actual Wavelength values with time for each
% coordinate
 LFinalMin= min(min(min(LFinal)));
 LFinalMax= max(max(max(LFinal)));

 for t = 1:10:numel(Time);
    figure
    pcolor(lon, lat,LFinal(:,:,t)');
    shading flat;
    axis equal;
    axis tight;
    maltacoast;
```

```matlab
    colorbar;
    colormap jet;
    % or hsv, summer, autumn, bone, grey...

    c.Label.String = 'Wavelength';
    caxis([LFinalMin LFinalMax]);
    axis([min(lon) max(lon) min(lat) max(lat)]);

    xlabel('Longitude (\circE)');
    ylabel('Latitude (\circN)');
    title(['Wave length for ', datestr(TimeSeries(t), 0)]);
    hold on;

    % Defining the contours
    u=(0.1:0.2:0.5);
    v=(0.5:0.5:1);
    w=union(u,v);
    %x=(1:20:100);
    %y= union(x,w);

    %plotting the contours
    [C, H] = contour(lon,lat,ratioLD(:,:,t)',v,'black', 'ShowText','on',
'TextStep',2);
    hh=clabel(C, H);
    set(hh,'Fontsize', 0.1, 'color', 'black');
    set(hh,'FontWeight','bold');
    hold off;
    % w = waitforbuttonpress;
    pause(0.05);
 end

% to plot wavelength extracting time series for a particular point
figure
pcolor(lon, lat, LFinal(:,:,1)');
shading flat;
axis equal;
axis tight;
xlabel('Longitude (\circE)');
ylabel('Latitude (\circN)');
title(['Select point']);

% expect input from mouse
[x y] = ginput(1);
close;

% point chosen  might not be in our data points
% find closest point minimum value and absolute
% it returns a vector - Value of x and its Position

[xMinVal xMinIdx] = min(abs(lon - x));
[yMinVal yMinIdx] = min(abs(lat - y));

% Wave length - timeseries all values of wavelength at that point for all t
timeSeriesWL = LFinal(xMinIdx, yMinIdx, :);
timeSeriesWL = squeeze(timeSeriesWL);
figure
plot(TimeSeries, timeSeriesWL, '-');
dateaxis('x', 0);
title(['Wavelength for Lon: ', num2str(lon(xMinIdx)), '\circE, Lat: ',
num2str(lat(yMinIdx)), '\circN'],'FontSize',12,'FontWeight','bold');
```

```matlab
xlabel('Date','FontWeight','Bold');
ylabel('Wavelength / m','FontWeight','Bold');

% to categorise areas based on wave length
% find max values
x= max(max(max(LFinal(:,:,:))));
range=[0 10 20 30 40 50 60 70 80 90 100 110 120];
i = numel(range);

%%% need to elaborate since the colours are for a range and not a value
for n = 1:1:(numel(range)-1);
  index = find((LFinalT > range(1,n)) & (LFinalT <= range(1,n+1)));
  LFinalT(index) = n*10;
end

 for t = 1:1:3; %numel(Time);
    pcolor(lon, lat,LFinalT(:,:,t)');
    shading flat;
    axis equal;
    axis tight;
    axis([min(lon) max(lon) min(lat) max(lat)]);

    xlabel('Longitude (\circE)');
    ylabel('Latitude (\circN)');

    title(['Wave length zones for ', datestr(TimeSeries(t), 0)]);
    %caxis([min LFinalT max LFinalT]);
    clbar = colorbar;
    set(clbar,'XTick',[0:10:120]);
    colormap winter;
    %colorbar;
    drawnow;
    pause(0.05)
 end

% to categorise areas based on ratio Depth / L final
% definition: Deep water: Depth/Lmax > 0.5 ;
%             Intermediate water : 1/20 < D/L < 0.5
%             Shallow water D/L < 1/20

% to determine Ratio for each lon , lat and T
% to plot different zones for all t we created a matrix ratioT which will
take modified numbers

ratioT = (zeros(lenlon,lenlat,numel(Time)));

for t = 1:1:numel(Time);
    for i=1:1:lenlon;
        for j = 1:1:lenlat;
            ratio(i,j,t) = (depth(i,j,t)/LFinal(i,j,t));
            ratioT(i,j,t) = ratio(i,j,t);
        end
    end
end

% find and change values of ratios as follows
%       Shallow water D/L < 1/20 become 100
%       Intermediate water : 1/20 < D/L < 0.5 become 200
%       Deep Water D/L > 0.5 become 300

index1 = find(ratioT <= 0.05);
```

```matlab
index2 = find((ratioT > 0.05) & (ratioT <= 0.5));
index3 = find(ratioT >= 0.5);

ratioT(index1) = 100;
ratioT(index2) = 200;
ratioT(index3) = 300;

% to plot animation of wave zones with time t

for t = 1:1:numel(Time)
    pcolor(lon, lat,ratioT(:,:,t)');
    shading flat;
    axis equal;
    axis tight;
    axis([min(lon) max(lon) min(lat) max(lat)]);

    xlabel('Longitude (\circE)');
    ylabel('Latitude (\circN)');

    title(['Sea depth categorisation for ', datestr(TimeSeries(t), 0)]);
    %caxis([min ratioT max ratioT]);
    colorbar;
    drawnow;
    pause(0.05);
end

%  Mean Wave Direction
%  angle Dir in degrees. covert to radians
%  magnitude dependant on Significant Height
uvalues = -(cos(deg2rad(Dir)).*Hs);
vvalues = -(sin(deg2rad(Dir)).*Hs);


test = Dir(:,:,1);

sf = 0.05;

lonres2 = (lon(2) - lon(1))/2;
latres2 = (lat(2) - lat(1))/2;

for t = 1:10:61
    figure
    N = 15;
    clf('reset')
    %to plot direction vectors
    pcolor(lon, lat, LFinalT(:,:,t)'); shading flat;
    %caxis([0 maxV]);
    hold on;
    quiver(lon([1:N:end])+lonres2, lat([1:N:end])+latres2,
(uvalues([1:N:end],[1:N:end],t).*sf)',
(vvalues([1:N:end],[1:N:end],t).*sf)', 0, 'k')
    axis equal;
    axis tight;
    title(['Wave length zones with wave direction for ',
datestr(TimeSeries(t), 0)]);
    colorbar;
    colormap winter;
    maltacoast;
    drawnow;
    %pause(1);
end
```

```matlab
% 3.  TO DETERMINE  Wave Steepness - waveheight / wavelength
%        Hs and  LFinal(i,j,t)

WS = (zeros(lenlon,lenlat,numel(Time)));

for t = 1:1:numel(Time);
    for i=1:1:lenlon;
        for j = 1:1:lenlat;
WS(i,j,t) = Hs(i,j,t)/LFinal(i,j,t);
        end
    end
end

% to produce figures of Wave Steepness with time for each
% coordinate
WSMin = min(min(min(WS)));
WSMax = max(max(max(WS)));

 for t = 1:10:numel(Time);
    figure
    pcolor(lon, lat,WS(:,:,t)');
    shading flat;
    axis equal;
    axis tight;
    maltacoast;

    colorbar;
    colormap jet;
   % or hsv, summer, autumn, bone, grey...
    caxis([WSMin WSMax]);
    axis([min(lon) max(lon) min(lat) max(lat)]);

    xlabel('Longitude (\circE)');
    ylabel('Latitude (\circN)');
    title(['Wave Steepness for ', datestr(TimeSeries(t), 0)]);
    drawnow;
    pause(0.05);
 end

% TO SUPERIMPOSE WAVEHEIGHT HS AND WAVE DIRECTION

%  Mean Wave Direction
%  angle Dir in degrees. covert to radians
%  magnitude dependant on Significant Height
uvalues = -(cos(deg2rad(Dir)).*Hs);
vvalues = -(sin(deg2rad(Dir)).*Hs);

sf = 0.05;

lonres2 = (lon(2) - lon(1))/2;
latres2 = (lat(2) - lat(1))/2;
for t = 1:10:60
    figure
    N = 15;
    clf('reset')
    %to plot direction vectors
    pcolor(lon, lat, WS(:,:,t)'); shading flat;

    hold on;
```

```
    quiver(lon([1:N:end])+lonres2, lat([1:N:end])+latres2,
(uvalues([1:N:end],[1:N:end],t).*sf)',
(vvalues([1:N:end],[1:N:end],t).*sf)', 0, 'k')
    axis equal;
    axis tight;
    maltacoast;

    colorbar;
    colormap jet;

    caxis([WSMin WSMax]);
    axis([min(lon) max(lon) min(lat) max(lat)]);
    title(['Wave Steepness with wave direction for ',
datestr(TimeSeries(t), 0)]);

    drawnow;
    %pause(1);
end
```

Program 2 – To Determine the Forecasted 3 hourly highest wave

Program Name : H 3 hours.m

```
Hs = ncread('output0101.nc','hs'); % Significant Wave Height in Metres
Time = ncread('output0101.nc', 'time'); % Time Vector (every hour)
lat = ncread ('output0101.nc','latitude'); % Latitude of data
lon = ncread('output0101.nc','longitude'); % Longitude of data
MPer = ncread('output0101.nc','tm01'); % mean wave period

% converts time vector at the start date
defDate = datenum([1970 1 1 0 0 0]);

for T = 1:1:numel(Time)
    TimeSeries(T) = addtodate(defDate, Time(T), 'second');
end

HsIS = isnan(Hs);
% ISNAN changes Nan entries to 1, other values are changed to 0
index1 = find (HsIS ==1);
Hs(index1) = 0;
MPer(index1) = 0;

T = numel(Time(:,1));
a = numel(lon(:,1));
b = numel(lat(:,1));

HsAv = zeros(a,b,T); % Average value of Hs over the next 3 hours
MPerAv = zeros(a,b,T); % Average Mean wave period over the next 3 hours


 for i= 1:1:numel(lon(:,1));
    for j=1:1:numel(lat(:,1));
        for t=1:10:T-2;
            Hstot(i,j,t) = (Hs(i,j,t)+ Hs(i,j,(t+1)) + Hs(i,j,(t+2)));
            HsAv(i,j,t)= (Hstot(i,j,t)*0.3333); % to divide by 3 to get
average

            MPertot(i,j,t)= (MPer(i,j,t)+ MPer(i,j,(t+1)) +
MPer(i,j,(t+2)));
```

169

```matlab
                MPerAv(i,j,t) = (MPertot(i,j,t)*0.3333); % to divide by 3 to
get average
        end

  HsAv(i,j,60) = (Hs(i,j,60)+ Hs(i,j,61))/2; % to determine the average Hs
close to the end of the data
  HsAv(i,j,61)=  Hs(i,j,61);

  MPerAv(i,j,60) = (MPer(i,j,60)+ MPer(i,j,61))/2;  % to determine the
average mean wave period close to the end of the data
  MPerAv(i,j,61)=  MPer(i,j,61);

    end
 end

 % to determine number of waves N in 3 hours = 3600 / MPer (i,j,k)
 % Maximum WaveHeight  for  a sequence of N waves = Hs * sqrt( 0.5 x ln N)
 % Maximum crest height is 0.6 total wave height

 NrW = zeros(a,b,T); % Number of waves
 Numerator = zeros(a,b,T);
 H3H = zeros(a,b,T);

 for i= 1:1:numel(lon(:,1));
    for j=1:1:numel(lat(:,1));
        for t = 1:10:T; % T;
        NrW(i,j,t)= (3600 / MPerAv(i,j,t));
        Nume(i,j,t) = (log(NrW(i,j,t)) *0.5);
        Numerator(i,j,t) = sqrt(Nume(i,j,t));
        H3H(i,j,t) = HsAv(i,j,t) * Numerator(i,j,t);
        % HCrest(i,j,t) = 0.6* H3H(i,j,t);
    end
 end
 end

 % to plot data every 5 hours for t= 0 to 61

for tt = 1:10:T;
figure
pcolor(lon, lat, H3H(:,:,tt)'); % pcolor - plot data values in colour
colormap('jet');
hold on;
shading flat;
axis equal;
axis tight;

% maltacoast;
colorbar;
caxis([00 5])
title(['Highest Wave 3hour Period Hmax3hr for ', datestr(TimeSeries(tt),
0)]);
axis([min(lon) max(lon) min(lat) max(lat)]);
xlabel('Longitude (\circE)');
ylabel('Latitude (\circN)');
end


%%% to do plots of Hs for same tiem period to compare values
for tt = 1:10:61;
```

```matlab
figure
pcolor(lon, lat, Hs(:,:,tt)'); % pcolor - plot data values in colour
colormap('jet');
hold on;
shading flat;
axis equal;
axis tight;

maltacoast;
colorbar;
caxis([00 5])
title(['SWave height Hs for ', datestr(TimeSeries(tt), 0)]);
axis([min(lon) max(lon) min(lat) max(lat)]);

xlabel('Longitude (\circE)');
ylabel('Latitude (\circN)');

end

% to compare values between H and Hs values for a period in time
% teh graphs we have used are for 02 January 2020 at 16.00
% that is t = 40.

ValHs = Hs(:,:,41);
ValH= H3H(:,:,41);

for i =1:1:numel(lon(:,1));
    for j = 1:1: numel(lat(:,1));
diff(i,j) = ValH(i,j)-ValHs(i,j);
    end
end

figure
pcolor(lon, lat, diff(:,:)'); % pcolor - plot data values in colour
colormap('jet');
hold on;
shading flat;
axis equal;
axis tight;
colorbar;
caxis([00 1.8])

title(['Difference H / Hs 02 January 16.00']);
axis([min(lon) max(lon) min(lat) max(lat)]);

xlabel('Longitude (\circE)');
ylabel('Latitude (\circN)');
```

Program 3 – To Determine the Probability of Exceedance and Percentikle Values – Actual Data
Program Name : *exceedance.m*

```matlab
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PROGRAM Percentiles
% Random values
X=[3.4;3.2;2.8;3.4;3.4;4.1;3.6;4.2;4.3;3.1;3.9;3.4;3.6;3.9;4.0;3.5;3.7;4.5;
5.2;3.4];
```

```matlab
m=max(size(X));
Xsort=sort(X, 'descend');
Xrank=(1:m)';
Xprob=zeros(m,1);
Xprob=Xrank/(m+1); % probability of exccedance

T = table(Xsort,Xrank,Xprob,'VariableNames',{'Sorted
Hs','Rank','Exceedance'});

% plotting exceedance probability curve (in %)
plot(Xprob * 100,Xsort,'r-','LineWidth',2);
xlabel('Exceedance Probability (%)','FontWeight','Bold');
ylabel('Value','FontWeight','Bold');

% to find the percenatge exceedance for a value we want say wavelength
% 3.8m
x=Xprob;
y=Xsort;
yval=3.8;

% interesting function. we do not have a y value for 3.8
% find the x value that gives us y = 3.8
% from the data we have, find the nearest y (Xsort) which is less than yval

xval=x(find(y<=yval,1,'first'))
f=@(a) interp1(x,y,a,'pchip') - yval;
xval=fzero(f,xval);

plot(x,y,'o-',xval,yval,'*');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% histogram , data collected from X put in 100 bins
% b is teh value of each bin, n : number of occurances
[n,b]=hist(X,100);

% plot bin against percentage nr. of occurances . 20 is the total number of
% data x
plot(b,((n/20)*100));

% cumsum = cumulative sum. ( adding the number of occurances)
cV1=cumsum(n);
plot(b,cV1);

% but for exceedance we want the opposite - so we flip n
cV2=fliplr(cumsum(fliplr(n)));
plot(b,cV2)

% To find the wavelength for a given exceedance  say 0.48
x=b; %wavelength
y=(cV2/20); % note that we are putting prob of exceedance as a value from
zero to 1
yval=0.48;
xval=x(find(y<=yval,1,'first'))
f=@(t) interp1(x,y,t,'pchip') - yval;
xval=fzero(f,xval);
```

Program 4 – To Determine probability wave height h exceeds a preset value by using Rayleigh's distribution
Program Name : RayleighDist_1.m

```matlab
% VIMP RAYLEIGH DISTRIBUTION
% FOR EACH LON AND LAT To determine the probability wave height h exceeds a
% preset value by using Rayleigh's number

% Use the Rayleigh distribution to forecast the probability of waves
% exceeding pre set wave heights
% after the model has calculated Hs

Hs = ncread('output0101.nc','hs'); % Significant Wave Height in Metres
Time = ncread('output0101.nc', 'time'); % Time Vector (every hour)
lat = ncread ('output0101.nc','latitude'); % Latitude of data
lon = ncread('output0101.nc','longitude'); % Longitude of data
MPer = ncread('output0101.nc','tm01'); % mean wave period

% converts time vector at the start date
defDate = datenum([1970 1 1 0 0 0]);

for t = 1:1:numel(Time)
    TimeSeries(t) = addtodate(defDate, Time(t), 'second');
end

% PDF based on Aldo's equation waveheight will reach height H
% we will find the PDF for each lon / lat based on the forecasted Hs
% we can allow operator to click on area and will give details as to
maximum forcasetd waveheight etc


HsIS = isnan(Hs);
% ISNAN changes Nan entries to 1, other values are changed to 0
index1 = find (HsIS ==1);
Hs(index1) = 0;


for t=1:10:61
    HsT = Hs(:,:,t);  % Significant wave height at t= t


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% TO FIND THE PDF (Probability Distribution Function)- see Aldo's notes
% for each lon and lat we will find the PDF for H 0.5,1,1.5,2,2.5,3

H=[0.5 1 1.5 2 2.5 3 3.3 4 4.5 5 5.5 6];
PDF = zeros(851,876,numel(H(1,:)));   % PDF
PDFPer= zeros(851,876,numel(H(1,:))); % PDF written as percentage

for i= 1:1:numel(lon(:,1));
    for j=1:1:numel(lat(:,1));
        HsP = HsT(i,j);
        for h=1:1:numel(H(1,:))
            H=[0.5 1 1.5 2 2.5 3 3.3 4 4.5 5 5.5 6];
            % a = 4*H/Hs^2 ; b = -2*H^2/Hs^2
            a= ((4*H(1,h)) / (HsP^2));
```

```matlab
                b= ((-2*H(1,h)^2)/(HsP^2));

                % to collect the data in a more clear manner, PDF(i,j,1) =
                % 0.5m etc
                PDF(i,j,h) = a*exp(b);

                % to convert to percentage
                PDFPer(i,j,h) = (PDF(i,j,h)*100);
            end
        end
end


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% TO FIND THE  PROBABILITY OF EXCEEDANCE RAYLEIGH DISTR. - ALDOS NOTES
% for each lon and lat we will find the PE for H 0.5 - 6
% HsT is as per above: t = 1
% CP = exp(-2*H^2/Hs^2)

% t= 10;
PE = zeros(851,876,numel(H(1,:)));        % PE
PEPer= zeros(851,876,numel(H(1,:)));      % PE as a percentage

for i= 1:1:numel(lon(:,1));
    for j=1:1:numel(lat(:,1));
        HsP = HsT(i,j);
        for h=1:1:numel(H(1,:))
            H=[0.5 1 1.5 2 2.5 3 3.3 4 4.5 5 5.5 6];

            % b = -2*H^2/Hs^2 - used above
            b= ((-2*H(1,h)^2)/(HsP^2));

            VAL = exp(b); % check with Aldo the equation

            PE(i,j,h) =VAL;
            % to convert to percentage
            PEPer(i,j,h) = (PE(i,j,h)*100);
        end
    end
end



%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% TO CATEGORISE AREAS BASED ON PROBABILITY OF EXCEEDANCE WAVEHEIGHT OF
SPECIFIED WAVEHEIGHT
% SAY 1.5M
% range: percentage range.

% PE for wave height 1.5m = PE(i,j,3) ;
% PE Percenatge for wave height 1.5m = PePer(i,j,3) ;

test = PEPer(:,:,3); % percentage

figure;
pcolor(lon, lat,test');
    shading flat;
    axis equal;
    axis tight;
    axis([min(lon) max(lon) min(lat) max(lat)]);
    maltacoast
```

```matlab
    xlabel('Longitude (\circE)');
    ylabel('Latitude (\circN)');

    title(['Probability H exceeds 1.5 m ', datestr(TimeSeries(t), 0)]);
    colormap jet
    clbar = colorbar;
    caxis([00 35])
    drawnow;
end

%%% to group in ranges depending on the percentages OF EXCEEDANCE
range=[0 5 10 15 20 25];
range2 = [2.5 7.5 12.5 17.5 22.5];
i = numel(range);

%%% need to elaborate since the colours are for a range and not a value
PEPerT=zeros(851,876);

for n = 1:1:(numel(range)-1);
  index = find((PEPer(:,:,3) > range(1,n)) & (PEPer(:,:,3)<=
range(1,n+1)));
  % those that fall in the percentage, will get value range (n+1)
  PEPerT(index) = range2(n);
end

figure
 pcolor(lon, lat,PEPerT(:,:)');
    shading flat;
    axis equal;
    axis tight;
    axis([min(lon) max(lon) min(lat) max(lat)]);
    maltacoast

    xlabel('Longitude (\circE)');
    ylabel('Latitude (\circN)');

    title(['Percentage Exceedance Grouped ', datestr(TimeSeries(t), 0)]);
    colormap jet
    clbar = colorbar;
    set(clbar,'XTick',[0:2.5:20]);
    %%% write the clear bar with text of range = see Perntile TM

    drawnow;
```

Program 5 – To Determine h for a given probability of Exceedance
Program Name : RayleighDist_2.m

```matlab
%%% THIS PROGRAM FINDS H GIVEN A PROBABILITY OF EXCEEDANCE
% TEH PROBABILITY OF EXCEEDANCE OF EACH LON / LAT AT TIME T.
% THE PROBABILITY OF EXCEEDANCE IS SET BY US- AT MOMENT ITS IS SET AT 20%
% DATA USED Hs
% plots the h that have a pr. of exceedance 20 for each lon and lat

Hs = ncread('output.nc','hs'); % Significant Wave Height in Metres
```

```matlab
Time = ncread('output.nc', 'time'); % Time Vector
lat = ncread ('output.nc','latitude'); % Latitude of data
lon = ncread('output.nc','longitude'); % Longitude of data

% converts time vector at the start date
defDate = datenum([1970 1 1 0 0 0]);


for t = 1:1:numel(Time)
    TimeSeries(t) = addtodate(defDate, Time(t), 'second');
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % to plot the graph of probability of exceedance for each lon / lat
    % lets set t= 1
    % Taking H from 0 to 6m in steps of 0.1m

  for t=5:1:5; % t =1 to 61 in acual fact
    HsT = Hs(:,:,t);

    HN= 0.0:0.2:6;  % Range of wave heights Hs from 0 to 6 in steps of 0.2
    PEG = zeros(851,876,numel(HN));
    PEGPer= zeros(851,876,numel(HN));

    for i= 1:1:numel(lon);
        for j=1:1:numel(lat);
            HsP = HsT(i,j); % Hs at a point

            for h= 1:1:numel(HN)

                % Equation : b = -2*H^2/Hs^2
                b= ((-2*HN(1,h)^2)/(HsP^2));
                VAL = exp(b); % check with Aldo the equation

                PEG(i,j,h) = VAL; % Pr. Exceedance at a point i,j for all h
                % Pr. Exccedance (percentage) at a point i,j for all h
                % (wave height)
                PEGPer(i,j,h) = (PEG(i,j,h)*100);
            end
        end
    end

xtest = squeeze(PEGPer(60,1,:));
plot(HN,xtest,'r-','LineWidth',2);

% %plotting slices of PEGPer
% for t = 1:1:size(PEGPer, 3)
%     pcolor(PEGPer(:,:,1)'); shading flat;
%     title(['t = ', num2str(t)]); axis equal; axis tight;
%     waitforbuttonpress;
% end

% TO FIND WAVEHEIGHT H THAT WILL GIVE US A PROBABILITY OF EXCEEDANCE 20%
xvalT = nan(numel(lon), numel(lat)); % creates matrix of all NAN values

for i= 1:1:numel(lon)
    disp(['Processing: ', num2str((i/numel(lon))*100), '%']); % Display
progress

    for j=1:1:numel(lat)
```

```matlab
        if (~isnan(PEGPer(i,j,1))) %

            x = HN';% range of wavelengths

            a = squeeze(PEGPer(i,j,:)); % Probabiity of excedance graph at
point i,j)
            y = a;

            % find the x value that gives us y (probability of exceedance)=
20%
            yval= 20;
            xval=x(find(y<=yval,1,'first')) ;
            f=@(u)interp1(x,y,u,'pchip') - yval;
            xval=fzero(f,xval);

            % to visually check what we are doing
            %plot(x,y,'-',xval,yval,'*');

            xvalT(i,j,t) = xval;
        end
    end
end
end
xvalTMin = min(min(min(xvalT(:,:,:))));
xvalTMax = max(max(max(xvalT(:,:,:))));

figure
pcolor(lon, lat, xvalT(:,:,5)');
title(['Waveheight with Prob. of Exceedance of 20%  at
',datestr(TimeSeries(5),0)]);
maltacoast;
title(['Probability H exceeds 1.5 m ', datestr(TimeSeries(t), 0)]);
shading flat;
axis equal;
axis tight;
colorbar;
colormap jet;

caxis([xvalTMin xvalTMax]);
axis([min(lon) max(lon) min(lat) max(lat)]);

    xlabel('Longitude (\circE)');
    ylabel('Latitude (\circN)');

    title(['H with probability of exceedance 20% ', datestr(TimeSeries(t),
0)]);
    drawnow;
```

Appendix I: Programs - Climatology EWA

Program 1 - *POT Analyses to Determine Peaks*
Program Name *: POT Wave Peaks.m*

```matlab
% TO FIND PEAKS USING POT METHOD  of Hs per month and per period we have
decided to use
% period    : 01 November - 30 April
% Data used : Maria Swan Model Data

%TO READ ROSARIO MARIA SWAN DATA COORDINATES :  FROM JOEL MAIL LON : 14.168
LAT: 36.0812
dataSwanM = dlmread('timeseries_RosarioMariaSwan 30.11.txt');
MariaDate = datenum([dataSwanM(:,1) dataSwanM(:,2) dataSwanM(:,3)
dataSwanM(:,4) zeros(size(dataSwanM, 1), 1) zeros(size(dataSwanM, 1), 1)]);

MariaH = dataSwanM(:,5);% Sea surface Wave significant height
MariaHs = MariaH*1.5; % putting in the correction factor from validation

MariaDir = dataSwanM(:,6); % Sea surface wave FROM direction

MariaT = dataSwanM(:,8) ; % Wave Mean Period
MariaTm = (MariaT*0.8)+1.25; % putting in the correction factor from
validation

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% CREATE MATRIX MARIADATA -CONTAINS - Datenum , HS ,Tm and calculated wave
steepness and wavelength
a=numel(dataSwanM(:,1));
MariaDATA = zeros(a,7);

for i = 1:1:a
MariaDATA(i,1) = MariaDate(i,1); % Wave Datenum
MariaDATA(i,2) = MariaHs(i,1);    % Hs in M
MariaDATA(i,3) = MariaTm(i,1);   % Tm

% WAVE STEEPNESS  - 2  pi Hs / g  Tm^2  saved in BuoyDATA (i,4);
MariaDATA(i,4)= (2*pi*MariaDATA(i,2))/(MariaDATA(i,3)^2*9.81);

% WAVELENGTH= g*T^2/ 2*pi saved in BuoyDATA (i,6);
MariaDATA(i,6)= (9.81 * MariaDATA(i,3)^2)/(2 * 3.142);

% WAVE DIRECTION - Sea waves  FROM direction
MariaDATA(i,7)= MariaDir(i,1);
end


% To plot entire Time series
figure;
plot(MariaDATA(:,1), MariaDATA(:,2),'b'); %Full time series
hold on
title(['Hs Time Series Rosario Maria Swan - Coordinate Lon : 14.168 Lat:
36.0812 '],'FontSize',14,'FontWeight','bold');

ax = gca;
ax.FontSize = 12;
xlabel('Date','FontWeight','bold');
```

```matlab
ylabel('Hs / m','FontWeight','bold');
grid on;
grid minor;
axis([ min(MariaDate) max(MariaDate) 0 max(MariaHs)+0.5]);
dateaxis('x',1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% To sort data by Chosen Period 01 November - 31 March
% derived fro previous program Peaks full time series
indStart= find((dataSwanM(:,2) >= 11));
indEnd = find((dataSwanM(:,2) < 4));
indRange= union(indStart,indEnd);

DATARange= MariaDATA(indRange,:);

figure;
hold on;
plot(DATARange(:,1),DATARange(:,2));

title(['Hs Time Series 01 November to 31 March; Period:  2011 -
2019'],'FontSize',14,'FontWeight','bold');
ax = gca;
ax.FontSize = 12;
xlabel('Date','FontWeight','bold');
ylabel('Hs / m','FontWeight','bold');

axis([ min(MariaDate) max(MariaDate) 0 max(MariaHs)+0.5]);
grid on;
grid minor;
dateaxis('x',1);
axis tight;
box on;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% TO DETERMINE THE PERCENTILE OF THE REDUCED DATA PERIOD 01 November to 30
April
% 80 to 95 in steps of 5
% to use Drago's function and the testing period we set
% as can be seen there is not much difference to the result attained
% before - Percentile 95 - TO CONFIRM

m= length(DATARange);
Xsort=sort(DATARange(:,2),'descend');

Xrank=(1:m)';
Xprob=zeros(m,1);
Xprob=Xrank/(m+1);

T = table(Xsort,Xrank,Xprob,'VariableNames',{'Sorted
Hs','Rank','Exceedance'});

% plotting exceedance probability curve (in %)
plot(Xprob * 100,Xsort,'r-','LineWidth',2);
xlabel('Exceedance Probability (%)','FontWeight','Bold');
ylabel('Value','FontWeight','Bold');

% histogram , data collected from X put in 100 bins
% b is the value of each bin, n : number of occurances
[n,b]=hist(DATARange(:,2),100);
```

```matlab
% plot bin against percentage nr. of occurance . c is the numel
BuoyDATA12(:,1)
m= numel(DATARange(:,1));
plot(b,((n/m)*100));

% cumsum = cumulative sum. (adding the percentage occurances)
cV1=cumsum(n);
plot(b,cV1);

% but for exceedance we want the opposite - so we flip n
cV2=fliplr(cumsum(fliplr(n)));
plot(b,cV2)

% To find the wavelength for a given exceedance say 0.90 exceedance
x=b;
y=(cV2/m); % note that we are putting prob of exceedance as a value from
zero to 1

% put probability of exceedance range 0.05 to 0.15 in steps of 0.05 and 0.4
% giving percentile 95,90,85 and 60 percentile for the lower end of value
figure
for i=1:1:4
    E= [ 0.05 0.1 0.15 0.40];
% yval(i,1) = the set exceedance ; yval(i,2) = wavelength value for that
% exceedance
yval(i,1)=E(i);
xval=x(find(y<=yval(i,1),1,'first'))
f=@(t) interp1(x,y,t,'pchip') - yval(i,1);
xval=fzero(f,xval);
yval(i,2) = xval


plot(x,y,'-',yval(i,2),yval(i,1),'*');
hold on
end

Percentile95= yval(1,2); % 95 Percentile = 3.5m
Percentile90= yval(2,2); % 90 Percentile = 2.8m
Percentile85= yval(3,2); % 85 Percentile = 2.4m
Percentile60= yval(4,2); % 60 Percentile = 1.5

title(['Probability of Exceedance / Corrected Hs Full Time
Series'],'FontSize',14,'FontWeight','bold');
grid on;
grid minor;
ax = gca;
ax.FontSize = 12;
xlabel('Significant Wave Height / m','FontWeight','bold');
ylabel('Pr. Exceedance','FontWeight','bold');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% FOR THE SORTED DATA SERIES - ALL YEARS
% Period 01 November - 31 March
% To evaluate the minimum points of the series which go below the 60
percentile.
% testing period : SORTED DATA SET
```

```matlab
for i=1:1:length(DATARange(:,1));
    DATAInv60(i,1)= DATARange(i,1);
    DATAInv60(i,2) = ((2*Percentile60)- DATARange(i,2));
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PLOTS: FOR Sorted data BETWEEN November and April
figure;
plot(DATARange(:,1),DATARange(:,2),'B');
hold on

plot(DATAInv60(:,1),DATAInv60(:,2),':r');
% axis([min(DataInvB(:,1)) max(DataInvB(:,1)) min(DataInvB(:,2))
max(DataInvB(:,2))]);

yline(Percentile95,'black'); %  95 percentile line
yline(Percentile60,'black'); %  60 percentile line

aaa= min(DATARange(:,1));
bbb= max(DATARange(:,1));
xticks([aaa:300:bbb]);
axis([ min(DATARange(:,1)) max(DATARange(:,1)) 0 max(DATARange(:,2))+0.5]);

dateaxis('x', 1);
title(['Peaks of Hs and Inverted Hs Period: November - April ; Year: 2011-
2019'],'FontSize',14,'FontWeight','bold');
ax = gca;
ax.FontSize = 10;
xlabel('Time','FontSize',12,'FontWeight','Bold');
ylabel('Sig. Wave Height / m','FontWeight','Bold');

legend ('Hs Time Series', 'Hs time series Inverted', '95-Percentile
Value','60-Percentile Value');

% to adjust limits on plot
Nov15= datenum([2015 11 01 00 00 01]);
April16 = datenum([2016 4 30 23 59 59]);
Year = April16 - Nov15;

axis([ Nov15 April16 0 max(DATARange(:,2))+0.5]);
title(['Peaks of Hs and Inverted Hs Period: November 2015- April 2016']);
xticks([Nov15:50:April16]);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% FIND PEAKS in TIME SERIES -  use FINDPEAKS - to find peaks in dataseries
% locs - gives points of peaks;
% lets set MinPeakDistance between successive peaks as 1 DAY
% use the time series 2012-2019

diff =  (datenum([2011 12 02 00 00 00])-(datenum([2011 12 01 00 00 00])));
HST= DATARange(:,2); % Hs
xT= DATARange(:,1); % Datenum

% finding the peaks with 95 percentile (0.05 exceedance)FOR THE CHOSEN DATA
% SERIES
% location of Peak is given in matrix locs95

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Setting the threshold values at different values
```

```matlab
Percentile95 = Percentile95;
[pksTD,locs95,widths] =
findpeaks(HST,xT,'Annotate','extents','MinPeakHeight',Percentile95,'MinPeak
Distance',diff);
plot(locs95,pksTD,'*','MarkerSize',10,'color','r');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% FIND PEAKS in INVERTED TIME SERIES -  use FINDPEAKS - to find peaks in
dataseries
% locs - gives points of peaks; Keep same percentile value
% Change MinPeakDistance between successive peaks as 12hours

%%% TO ADJUST THAT THE HIGHEST PEAK IS CHOSEN BETWEEN TWO WHEN ONE NOT
%%% CHOSEN

HSTI=DATAInv60(:,2);
xTI= DATAInv60(:,1);
diff2 =  (datenum([2011 12 01 12 00 00])-(datenum([2011 12 01 00 00 00])));

[pksI60,locsInv60,widths] =
findpeaks(HSTI,xTI,'Annotate','extents','MinPeakHeight',Percentile60,'MinPe
akDistance',diff2);
plot(locsInv60,pksI60,'*','MarkerSize',10,'color','g');

aT = numel(locs95(:,1));%%%%
bT = numel(locsInv60(:,1));
test = zeros(aT ,1);

for i= 1:1:numel(locs95);
dataPeaks(i,1)=  locs95(i,1);
dataPeaks(i,2)= pksTD(i,1);
end

 test=zeros(aT,2);
for v=2:1:aT;
  Peaks(1,:)= dataPeaks(1,:);
  indexT1 = find((locsInv60(:,1) > dataPeaks((v-1),1)));
  indexT2 = find((locsInv60(:,1) < dataPeaks(v,1)));
  indexT3 = intersect(indexT1,indexT2);

  if indexT3>0
      Peaks(v,:)= dataPeaks(v,:);
  else
      Peaks(v,1)= 0;
      test(v,:) = dataPeaks(v,:);
  end

end

% to compare data of points and choose the highest Hs point
for i=2:1:length(Peaks);
    if Peaks(i-1,2)<test(i,2)
     Peaks(i-1,1)=0
     Peaks(i,1)= test(i,1)
end
end

indexx= find(Peaks(:,1)>0);
peak2D(:,1)= Peaks(indexx,1);
```

```matlab
for i=1:1:numel(indexx);
idx2t(i) = find (DATARange(:,1)== peak2D(i,1)); % find teh data from the
data series which has same datenum as peak2D
end

% WE HAVE THE DATES OF WHEN THE PEAKS OCCUR
% COLUMNS ARE MADE UP
% 1. date 2. Hs 3. Tz 4. Wave steepness 5. Wave Steepbessx100 6. Wavelength
% 7. wave direction
waw2 = DATARange(idx2t,:,:);
plot(waw2(:,1),waw2(:,2),'o','MarkerSize',10,'color','black');

legend('Hs Time Series', 'Hs Inverted', '95 Percentile Value', '60
Percentile Value', 'Peaks Time Series','Peaks Inverted Time
Series','Selected Peaks');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Setting the number of peaks
nrx = (numel(waw2(:,1)))/8;
nr = 8;

%%%%%% TO EXTRACT THE DATA for these peaks
for i=1:1:numel(waw2(:,1));
DataPeaks(i,:)= datevec(waw2(i,1));
end

for i=1:1:numel(waw2(:,1));
DataPeaks(i,7)= waw2(i,1); % DATENUM
DataPeaks(i,8)= waw2(i,2); % HS in M
DataPeaks(i,9)= waw2(i,3); % Tm
DataPeaks(i,10)=waw2(i,4); % STEEPNESS
DataPeaks(i,11)=waw2(i,6); % WAVELENGTH
DataPeaks(i,12)=waw2(i,7); % WAVE DIRECTION
end

% HISTOGRAM NUMBER OF PEAKS PER MONTH
figure
histBins = [1:1:12];
month=DataPeaks(:,2)
histCnts = histc(month, [1:1:12]);
bar(histBins, histCnts);

total= numel(DataPeaks(:,7));

bar(histBins, histCnts*100/total);

title(['Percentage Nr. of Exceedances/ Month (Exc. Peripheral Months)
'],'FontSize',14,'FontWeight','bold');

xlabel('Month','FontSize',12','fontweight','bold');
ylabel('Percentage/ %','FontSize',12,'fontweight','bold');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% HISTOGRAM VALUES of PEAKS HS AND NR. OF OCCURANCES
Counter1 =0;
WH = (0:0.5:7);

    for i=1:1:15; %wave height classes
```

183

```matlab
    for I = 1:1:length(DataPeaks(:,1));
        a = WH(1,i)+0.25;  %%%%%% CHECK THESE
        b = WH(1,i)-0.25;     %%%% IF THEY ARE CENTRED AROUND WH
        if DataPeaks(I,8)<=a && DataPeaks(I,8)>b
        Counter1 = Counter1 +1;
        end
    end
% Percent = Counter1/(length(Tcross)-N)*100;
WH(2,i) =Counter1;
Counter1 =0;
end

% To plot histogram Hs  / percentage occurrance
bar(WH(1,:),WH(2,:)*100/total);

title(['Range Hs Storms /  Percentage Nr. of Occurances
'],'FontSize',14,'FontWeight','bold');
ax = gca;
ax.FontSize = 12;
xlabel('Sig. Wave Height / m ','FontSize',12,'fontweight','bold');
ylabel('Percentage Nr. of Occurances / %','FontSize',12,
'fontweight','bold');

axis([ 3 8 0 40]);

% To plot histogram Hs  / nr of occurances

bar(WH(1,:),WH(2,:));

title(['Range Hs Storms /  Nr. of Occurances
'],'FontSize',14,'FontWeight','bold');
ax = gca;
ax.FontSize = 12;
xlabel('Sig. Wave Height / m ','FontSize',12,'fontweight','bold');
ylabel('Nr. of Occurances  ','FontSize',12, 'fontweight','bold');

axis([ 3 8 0 20]);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% HISTOGRAM OF WAVELENGTH AND NR. OF OCCURANCES
WLmin = min(DataPeaks(:,11));
WLmax = max(DataPeaks(:,11));

Counter1 =0;
WL = (50:10:150);

    for i=1:1:numel(WL(1,:)); %wave height classes
        for I = 1:1:length(DataPeaks(:,1));
            a = WL(1,i)+5;
            b = WL(1,i)-5;
            if DataPeaks(I,11)<=a && DataPeaks(I,11)>b
            Counter1 = Counter1 +1;
            end
        end
    % Percent = Counter1/(length(Tcross)-N)*100;
    WL(2,i) =Counter1;
    Counter1 =0;
    end
```

```matlab
% To plot histogram Wavelength
bar(WL(1,:),WL(2,:)*100/total);

title(['Wavelength Peak Events / Percentage Nr. of Occurances
'],'FontSize',14,'FontWeight','bold');
ax = gca;
ax.FontSize = 14;
xlabel('Wavelength / m ','FontSize',12,'fontweight','bold');
ylabel('Percentage Nr. of Occurances /
%','FontSize',12,'fontweight','bold');

axis([ 40 100 0 45]);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% HISTOGRAM OF Tm UPCROSSING PERIOD
TcrossMin = min(DataPeaks(:,9));
TcrossMax = max(DataPeaks(:,9));
Counter1 =0;
TC= (5:0.5:10);
    for i=1:1:numel(TC(1,:)); %wave height classes
        for I = 1:1:length(DataPeaks(:,1));
            a = TC(1,i)+0.25;
            b = TC(1,i)-0.25;
            if DataPeaks(I,9)<=a && DataPeaks(I,9)>b
            Counter1 = Counter1 +1;
            end
        end
    %  Percent = Counter1/(length(Tcross)-N)*100;
    TC(2,i) =Counter1;
    Counter1 =0;
    end

% To plot histogram T upcross period
bar(TC(1,:),TC(2,:)*100/total);

title(['T upcross Period Extreme Events / Percentage Nr. of Occurances
'],'FontSize',14,'FontWeight','bold');

ax = gca;
ax.FontSize = 14;

xlabel('T upcross Period / sec ','FontSize',12,'fontweight','bold' );
ylabel('Percentage Nr. of Occurances / %','FontSize',12,'fontweight','bold'
);

axis([ 5 8 0 40]);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% HISTOGRAM OF WAVE STEEPNESS AND NR. OF OCCURANCES
WSMin = min(DataPeaks(:,10))
WSMax = max(DataPeaks(:,10))

Counter1 =0;
WS = [];
WS(1,:) = (0.0:0.005:0.1);

    for i=1:1:numel(WS(1,:)); %wave height classes
```

```matlab
    for I = 1:1:length(DataPeaks(:,1));
        a = WS(1,i)+0.0025;
        b = WS(1,i)-0.0025;
        if DataPeaks(I,10)<=a && DataPeaks(I,10)>b
        Counter1 = Counter1 +1;
        end
    end
% Percent = Counter1/(length(Tcross)-N)*100;
WS(2,i) =Counter1;
Counter1 =0;
end

% To plot histogram WaveSTEEPNESS
bar(WS(1,:),WS(2,:)*100/total);
title(['Wave Steepness Extreme Events/ Per. Nr. of Occurances
'],'FontSize',14,'FontWeight','bold');

ax = gca;
ax.FontSize = 14;
xlabel('Wave Steepness','FontSize',14,'fontweight','bold' );
ylabel('Percentage Nr. of Occurances','FontSize',14,'fontweight','bold');

axis([ 0.055 0.095 0 40]);

% HISTOGRAM VALUES of WAVE DIRECTION AND NR. OF OCCURANCES
Counter1 =0;
Dir = (0:45:315);

    for i=2:1:length(Dir); %wave direction classes
        for I = 1:1:length(DataPeaks(:,1));
            a = Dir(1,i)+22.5;  %%%%% CHECK THESE
            b = Dir(1,i)-22.5;    %%%% IF THEY ARE CENTRED AROUND WH
            if DataPeaks(I,12)<=a && DataPeaks(I,12)>b
            Counter1 = Counter1 +1;
            end
        end

        Dir(2,i) =Counter1;
    Counter1 =0;
    end

    indexa = find(DataPeaks(:,12)>=0);
    indexb = find(DataPeaks(:,12)<22.5);
    index= intersect(indexa,indexb); %%%% to check intersection

    indexc= find(DataPeaks(I,12)>=347.5);%%%% to check
    indexT= union(index , indexc);%%%% to check

    Dir(2,1) = length(indexT);

figure
% To plot histogram Hs
bar(Dir(1,:),Dir(2,:)*100/total);

title(['Wave Direction Peak Events / Per.  Nr. of Occurances
'],'FontSize',14,'FontWeight','bold');
xlabel('Direction / Deg.','FontSize',12 );
ylabel('Percentage Nr. of Occurances / % ','FontSize',12);
```

```matlab
% to find wave height for direction 270

inda = find(DataPeaks(:,12)>=247.5);
    indb = find(DataPeaks(:,12)<292.6);
    ind= intersect(inda,indb); %%%% to check intersection

HSDir = zeros(numel(ind),12);
HSDir(:,:)=DataPeaks(ind,:)

Counter1 =0;
WH = (0:0.5:7);
totalE=length(HSDir(:,1));

    for i=1:1:15; %wave height classes
        for I = 1:1:length(HSDir(:,1));
            a = WH(1,i)+0.25;   %%%%%% CHECK THESE
            b = WH(1,i)-0.25;    %%%% IF THEY ARE CENTRED AROUND WH
            if HSDir(I,8)<=a && HSDir(I,8)>b
            Counter1 = Counter1 +1;
            end
        end
    %  Percent = Counter1/(length(Tcross)-N)*100;
    WHE(2,i) =Counter1;
    Counter1 =0;
    end

% To plot histogram Hs
bar(WH(1,:),WHE(2,:));

title(['Range Hs 270 deg /  Nr. of Occurances
'],'FontSize',14,'FontWeight','bold');
ax = gca;
ax.FontSize = 12;
xlabel('Sig. Wave Height / m ','FontSize',12,'fontweight','bold');
ylabel(' Nr. of Occurances ','FontSize',12, 'fontweight','bold');

axis([ 3 7.5 0 15]);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% EXTRACTING DATA PER PERIOD
% MONTH - PERIOD 01 February 2012 - 01 March 2012
dated1 = datenum([2012 02 01 00 00 00]);
dated2 = datenum([2012 03 01 00 00 00]);

indexd= find((MariaDATA(:,1) > dated1) & (MariaDATA(:,1)<= dated2 ));
BuoyDATAd= MariaDATA(indexd,:,:);
d= numel(indexd(:,1));

% TO PLOT THE TIMESERIES PERIOD 01 February 2012 - 01 March 2012
figure
plot(BuoyDATAd(:,1), BuoyDATAd(:,2)); % February 2012
hold on
title(['Hs Rosario Maria Swan - Feb.
2012'],'FontSize',14,'FontWeight','bold');

ax = gca;
ax.FontSize = 12;
xlabel('Date','FontWeight','bold');
ylabel('Hs / m','FontWeight','bold');
```

```matlab
dateaxis('x',1);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% To evaluate the minimum points of the series which go below the 60
percentile.
% To use y = -(H - H60) + H60 where H 60 = 60 percerntile
% y = 2* H60 - H
% Period : 01 February 2012 - 01 March 2012

for i=1:1:length(BuoyDATAd(:,1));
    InvD60(i,1)= BuoyDATAd(i,1);
    InvD60(i,2) = ((2*Percentile60)- BuoyDATAd(i,2));
end

% PLOTS  Period:  01 February 2012 - 01 March 2012
figure
plot(BuoyDATAd(:,1),BuoyDATAd(:,2),'b');
hold on
plot(InvD60(:,1),InvD60(:,2),':r');
yline(Percentile95,'-r');
yline(Percentile60,'-r');
dateaxis('x', 1);
legend ('TimeSeries', 'TS Inverted 90Percentile','TS Inverted
60Percentile','90 percentile', '60 percentile');

title(['Peaks of Hs and Inverted Hs Period: February 2012']);
xlabel('Time','FontWeight','Bold');
ylabel('Sig. Wave Height / m','FontWeight','Bold');

axis([ min(BuoyDATAd(:,1)) max(BuoyDATAd(:,1)) 0 max(BuoyDATAd(:,2))+0.5]);
grid minor;
grid on;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% FIND PEAKS in TIME SERIES -  use FINDPEAKS - to find peaks in dataseries
% locs - gives points of peaks;
% lets set MinPeakDistance between successive peaks as 1 day
% use the time series February 2012

diff =  (datenum([2011 12 02 00 00 00])-(datenum([2011 12 01 00 00 00])));
HS= BuoyDATAd(:,2);
x= BuoyDATAd(:,1);

% finding the peaks with 95 percentile ( 0.05 exceedance)
% location of Peak is given in matrix locs95

[pks,locs95M,widths] =
findpeaks(HS,x,'Annotate','extents','MinPeakHeight',Percentile95,'MinPeakDi
stance',diff);
plot(locs95M,pks,'*','MarkerSize',10,'color','r');

% finding the peaks with 90 percentile ( 0.1 exceedance)
% [pks,locs90,widths] =
findpeaks(HS,x,'MinPeakHeight',Percentile90,'MinPeakDistance',diff);
% plot(locs90,pks,'o','MarkerSize',10,'color','black');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% FIND PEAKS in INVERTED TIME SERIES -  use FINDPEAKS - to find peaks in
dataseries
% locs - gives points of peaks; Keep same percentile value
```

```matlab
% Change MinPeakDistance between successive peaks as 12hours

HSI=InvD60(:,2);
xI= InvD60(:,1);
diff2 =  (datenum([2011 12 01 12 00 00])-(datenum([2011 12 01 00 00 00])));

[pks,locsI60,widths] =
findpeaks(HSI,xI,'Annotate','extents','MinPeakHeight',Percentile60,'MinPeak
Distance',diff2);
plot(locsI60,pks,'*','MarkerSize',10,'color','g');

%%%%%%%  TO CONFIRM PEAKS
% A PEAK IS VALID IF BEFORE PEAK, HS IS LOWER THAN 60 PERCENTILE
% THAT IS WE WE HAVE PEAK ON 60INVERTED

a = numel(locs95M(:,1)); % location of teh peaks exceeding 95 percentile
b = numel(locsI60(:,1)); % location of teh inverted peaks exceeding 60
percentile

for v= 2:1:a;
  peak(1,1)= locs95M(1,1) % First peak is always a peak
  idx1 = find((locsI60(:,1) > locs95M((v-1),1)));
  idx2 = find((locsI60(:,1) < locs95M(v,1)));
  idx3 = intersect(idx1,idx2);

  if idx3>0
      peak(v,1)= locs95M(v,1);
  else
      peak(v,1)= 0;
  end

  % if idx6>0  % PART OF FINDING MINIMUM PEAK AFTER
    % locs95M(v,1) = locs95(v,1);
  % else
    % peak(v,1)=0 ;
end

idx= find(peak(:,1)>0);
peak2(:,1)= peak(idx,1);

% To extract the data from the data series for the points which match peak2
% datenum
for i=1:1: numel(peak2(:,1));
idx2(i) = find (DATARange(:,1)== peak2(i,1));
end

% to plot
waw = DATARange(idx2,:,:);
plot(waw(:,1),waw(:,2),'o','MarkerSize',10,'color','black');


%%% TO USE ALDO METHOD TO DETERMINE CUMULATIVE DISTRIBUTION FUNCTION OF THE
%%% PEAKS
X = waw2(:,2);

m = max(size(X));
Xsort = sort(X, 'descend');
Xrank = (1:m)';
Xprob = zeros(m,1);
```

```matlab
Xprob = Xrank/(m+1); %Probability of exceedance
Xcdf= 1-Xprob; % see paper "Uncertainties " page 4. Gives different
equation

T = table(Xsort,Xrank,Xprob,Xcdf,'VariableNames',{'Sorted
Hs','Rank','Exceedance','Cumulative Distribution Data'});

% plotting CDF curve (in %)
plot(Xcdf * 100,Xsort,'r-','LineWidth',2);
xlabel('CDF (%)','FontWeight','Bold');
ylabel('Value','FontWeight','Bold');

% plotting exceedance probability curve (in %)
plot( Xprob* 100,Xsort,'r-','LineWidth',2);
xlabel('Exceedance Probability (%)','FontWeight','Bold');
ylabel('Value','FontWeight','Bold');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% finding the PDF from data - histogram-
% X put in 10 bins ; b is teh value of each bin, n : number of occurances
[n,b]=hist(X,10);

% plot bin against perc nr. of occurances. 42 is the total number of data x
bar(b, (n/42));
% plot(b,(n/42));
% cumsum = cumulative sum. (adding the number of occurances)= CDF
cV1=cumsum(n);
plot(b,(cV1/42));
```

Program 2 - *To fit distributions and determine return period*
Program Name : Return Period.m

```matlab
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% DATA OF PEAKS TAKEN FORM PREVIOUS PROGRAM 'POT Wave Peaks)
% WE HAVE THE DATA OF THE PEAKS from waw2 and datapeaks
% COLUMNS ARE MADE UP
%1. date 2. Hs 3. Tz 4. Wave steepness 6. Wavelength 7. wave direction

% we need to correct the correction factor from 1.5 to 1.6
% se we will multiply the Hs column by 1.6 / 1.5 = 1.07

waw3 = waw2;
waw3(:,2)= waw2(:,2)* 1.067;

X=waw3(:,2); % peak Hs sorted based on our constraints


m= max(size(X));
Xsort=sort(X, 'descend');
Xrank=(1:m)';
Q =zeros(m,1);

% to ensure there is only one value of Xsort. If two are the same it adds
% 0.001 to one of the values
for i=2:1:m
```

```matlab
    for j=1:1:i-1
      if Xsort(i)== Xsort(j);
         Xsort(i) = Xsort(i)+0.001;
      end
    end
end

% TO DETERMINE PROBABILITY OF EXCEEDANCE Q USING GRINGORTEN EQUATION-
% REAL DATA. ** katalinic
a = 0.44; % Constant Gringorten equation
Q = (Xrank - a) /(m + 1 - (2*a)); % Q  Probabiity of Exceedance

% mean value of X ; Peak Hs
Xmean = mean(X);

% TR is the RETURN PERIOD = 1 / (N Q) based on assumption of N peaks per
year.
% From actual data we will determie the average nr. of peaks per year = 7
% CHECK PAPER"Uncertainties of Estimating etc etc
% 7 PEAKS PER YEAR Average

% Estimating Tr and P for the real data

for i= 1:1: m
    TR(i,1) = 1 / (nr * Q(i)); %%%%%%%%%%%% NR YEARS TR Return period -
see above
    P(i,1) = 1- Q(i); % P = Non Exceedance Probability P, CDF, = 1 - Q
    Su(i) = (X(i) - Xmean)^2;
end

Summation = sum(Su);
% Standard Deviation of distribution - (St. dev)^2 = Variance
StDv2 = Summation / (m-1);  % Variance
StDv = sqrt(StDv2); % St. deviation

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Plotting the PDF from the data X - CHOSEN PEAKS
% X put in 10 bins ; b is teh value of each bin, n : number of occurances
[n,b]=hist(X,100);
% plot bin against perc nr. of occurances. 'm' is the number of occurance
bar(b,(n/m)); % Probability Distribution Function
% plot(b,(n/m));
% cumsum = cumulative sum. (adding the number of occurances)= CDF
cV1=cumsum(n);

figure;
plot(b,(cV1/m));
hold on
% plotting CDf attained from Gringorten equation against bin method
plot(Xsort,P);
title(['Cumulative Distribution Function / Gringorten method and bars']);
xlabel('Hs / m','FontWeight','Bold');
ylabel('Probability ','FontWeight','Bold');

hold off

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Gumbell Distribution - from Aldo Paper Return Period
% paper Uncetainties of estimating extreme sig wave heiht
```

```matlab
% Gumbell dist is used for the ANNUAL MAXIMUM approach rather than the POT
% so lets assume we do not need to remove the threshold from x
% the x values are Xsort

alpha = sqrt(6)* StDv / 3.142 ;
U = (Xmean - (0.5772*alpha));

% CDF = exp [-exp[-(xi - u/alpha)] ;

for i = 1:1:m
const(i)= (-(Xsort(i) - U) / alpha);
val(i)= exp(const(i));
%val(i) = exp(-(Xsort(i) - U) / alpha);

PDFGum(i,1)= exp(const(i)-val(i))/alpha; % PDF Gumbel Distribution - Check

CDFGum(i,1) = exp(-val(i)); % CDFGum = CDF , Non Exceedance Probability P =
1 - Q
QGum(i,1) = 1 - CDFGum(i,1); % QGum = Pr. of exceedance
TRGum(i,1) = 1 / (nr*QGum(i,1)); % Return Period = 1 / Pr. of exceedance
%%%% NR YEARs
end


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% POT- For POT method x = x(i) - h0 with h0 = threshold value
% THRESHOLD VALUE SHOULDBE  THE 95 PERCENTILE

h0= Percentile95;

%X Excess = XE = X - X threshold (h0)
for i=1:1:m;
    XE(i,1) = X(i) - h0;
end

XEMean = mean(XE);
XESort = sort(XE, 'descend');

for i= 1:1: m
    SumE(i) = (XE(i) - XEMean)^2;
end
SummationE = sum(SumE);

% Standard Deviation of distribution - (St. dev)^2 = Variance
StDvExp2 = SummationE / (m-1);  % Variance
StDvExp = sqrt(StDvExp2); % St. deviation

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% METHOD OF FITTING WEIBULL DISTRIBUTION
% Finding the parameters - Weibull distribution
% Fitting distribution. Maximum Likelihood fitting mechanism is used
% Data used is XE that is x excess

Wpd = fitdist(XE,'Weibull');
A = Wpd.A;
B = Wpd.B;
```

```matlab
%%%%%%% Plotting the Weibull Probability Distribution function
x_values = XESort;
PDFwei = pdf(Wpd,x_values);

% FOR PLOTTING WE ARE KEEPING THE ORIGINAL X VALUE WITHOUT REDUCUNG THE
% THRESHOLD
figure
plot(Xsort,PDFwei);
hold on;
plot(Xsort,P);% P - probability Distr. Gringorten
title(['Probability Distribution Function - Weibull
Distribution'],'FontSize',14,'FontWeight','bold');
xlabel('Waveheight','FontSize',12,'FontWeight','Bold');
ylabel('Probability Disrtibution
Function','FontSize',12,'FontWeight','Bold');

legend('Weibull distribution','Actual Values');

% CREATE THE CDF BASED ON Weibull AND USING A AND B FROM ABOVE
CDFWei = wblcdf(XESort,A,B);

% CREATE THE PROBABILITY OF EXCEEDANCE Weibull Qwei = 1- CDFWei
% AND RETURN PERIOD TR
for i = 1:1:m
QWei(i,1) = 1- CDFWei(i);
TRWei(i,1) = 1/(nr*QWei(i,1)); %%%%%%%% NR OF YEARS
end

%%%%%%%%%%% GENERALISED PARETO DISTRIBUTION
GPDpd = fitdist(XE,'GeneralizedPareto');

k = GPDpd.k; % Shape parameter
sigma = GPDpd.sigma; %Scale Parameter
theta = GPDpd.theta; % Location Parameter

%%%% Generating the GPD PDFunction
PDFgpd= pdf(GPDpd,XESort);
plot(Xsort,PDFgpd);
hold on

%%% CREATING THE CDF GPD
CDFgpd = gpcdf(XESort,k,sigma,theta);

% CREATE THE PROBABILITY OF EXCEEDANCE GPD = 1- CDFgpd
% and RETURN PERIOD
for i = 1:1:m
Qgpd(i,1) = 1- CDFgpd(i); % Prob. of exceedance = 1 - Cum Distr Function
TRgpd(i,1) = 1/(nr*Qgpd(i,1)); %%%%%%%% NR OF YEARS
end


T2 = table(Xsort,Xrank,Q,P,QGum,CDFGum,QWei,CDFWei,Qgpd,
CDFgpd,'VariableNames',{'Sorted Hs','Rank','Exceedance','Cum.
Dist.Data','Exceedance Gum','Cum. Dist.Gum','Exceedance Wei','Cum.
Dist.Wei','Exceedance GPD','Cum. Dist.GPD'});
% To extract values from a table use the{} brackets i.e. T3= T2{:,1};

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% TABLES AND PLOTTING - PROBABILITY DISTRIBUTION FUNCTION
```

```matlab
Tpdf = table(Xsort,Xrank,PDFGum,PDFwei,PDFgpd,'VariableNames',{'Sorted
Hs','Rank','PDF Gum','PDF Weibull','PDF GPD'});

% plotting the PDF against real data
figure
plot(Xsort,PDFGum, 'r','linewidth',2);

title(['Probability Distribution Function - Fitted Distributions; Sig. Wave
height '],'FontSize',14,'FontWeight','bold');
xlabel('Significant Wave Height Hs / m','FontSize',12,'FontWeight','Bold');
ylabel('Probability / % ','FontSize',12,'FontWeight','Bold');

hold on
plot(Xsort,PDFwei,'b','linewidth',2);
plot(Xsort,PDFgpd,'g','linewidth',2);

ax = gca;
ax.FontSize = 12;

legend('Gumbell distribution','Weibull distribution','GPD');
hold off;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PLOT THE CDF DATA / GUMBELL / WEIBULL / GEN. PARETO DISTRIBUTION

Tcdf = table(P,CDFGum,CDFWei,CDFgpd,'VariableNames',{'CDF data','CDF
Gum','CDF Wei','CDF GPD'});

figure
plot(Xsort,P,'black','linewidth',2);% plot the CDF based on Aldo's method
title(['Cumulative Distribution
Function'],'FontSize',14,'FontWeight','bold');
xlabel('Waveheight / m','FontSize',12,'FontWeight','Bold');
ylabel('Cum. Dist. Function','FontSize',12,'FontWeight','Bold');

hold on
ax = gca;
ax.FontSize = 12;

plot(Xsort,CDFGum,'r','linewidth',2);
plot(Xsort,CDFWei,'b','linewidth',2); % plot the CDF based on Weibull's
method
plot(Xsort,CDFgpd,'g','linewidth',2);

legend('Gringorten method','Weibull distribution','Gumbell
Dist','GPD','Location','southeast' );
hold off;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% RETURN PERIOD -
% Return Period = 1/ Av nr. of peak * pr. of exceedance q

Texceedance = table(Xsort,Q,QGum,QWei,Qgpd,'VariableNames',{'X value','Pr.
Exceedance','Pr. Exceedance Gum','Pr. Exceedance Wei','Pr. Exceedance
GPD'});
Tretpd = table(Xsort,TR,TRGum,TRWei,TRgpd,'VariableNames',{'Hs','Return
Period','Ret. Per. Gum','Ret. Per. Wei','Ret. Per. GPD'});
```

```matlab
figure
semilogx(TR,Xsort, 'black','linewidth',2);
title(['RETURN PERIOD Threshold 3.0 m
'],'FontSize',14,'FontWeight','bold');
xlabel('Return Period / Years','FontSize',12,'FontWeight','Bold');
ylabel('Waveheight / m','FontSize',12,'FontWeight','Bold');

hold on
plot(TRGum,Xsort,'r','linewidth',2);
plot(TRWei,Xsort,'b','linewidth',2);
plot(TRgpd,Xsort,'g','linewidth',2);

ax = gca;
ax.FontSize = 12;

grid on;
grid minor;
legend('Gringorten method','Gumbell distribution','Weibull
distribution','GPD','Location','southeast');

hold on

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% To determine the correlation coefficient R
% to determine the waveheight for EACH DISTRIBUTION that gives a
probability of exceedance 0.1 - 1.0

Texceedance = table(Xsort,Q,QGum,QWei,Qgpd,'VariableNames',{'X value','Pr.
Exceedance','Pr. Exceedance Gum','Pr. Exceedance Wei','Pr. Exceedance
GPD'});

figure;
plot(Q,Xsort);
hold on
plot(QGum,Xsort);
plot(QWei,Xsort);
plot(Qgpd,Xsort);
hold off

x = Xsort;
YAct = Q;
YGum = QGum;
YWei = QWei;
YGpd = Qgpd;

yval1 = 0.05:0.05:0.95; % probability of exceedance
n=numel(yval1);

%< or > ?
for i=1:1:n
    yval= yval1(1,i);

    XvalAct = x(find(YAct>=yval,1,'first')); %< or > ?
    f=@(a) interp1(x,YAct,a,'pchip') - yval;
    XvalAct=fzero(f,XvalAct);

    XvalGum = Xsort(find(YGum>=yval,1,'first'));
    f=@(a) interp1(x,YGum,a,'pchip') - yval;
```

```matlab
    XvalGum=fzero(f,XvalGum);

    XvalWei = Xsort(find(YWei>=yval,1,'first'));
    f=@(a) interp1(x,YWei,a,'pchip') - yval;
    XvalWei = fzero(f,XvalWei);

    XAct(i,1) = XvalAct;
    XGum(i,1) = XvalGum;
    XWei(i,1) = XvalWei;
    % XGpd(i,1) = XvalGpd;
end


%%%% for GPD we had to reduce teh range because the points are less than
%%%% 0.85
yval2 = 0.05:0.05:0.85; % probability of exceedance
n2=numel(yval2);

for i=1:1:n2
    yval= yval2(1,i);

    XvalGpd = Xsort(find(YGpd>=yval,1,'first'));
    f=@(a) interp1(x,YGpd,a,'pchip') - yval;
    XvalGpd = fzero(f,XvalGpd);
    XGpd(i,1) = XvalGpd;
end

% Actual Values
figure
plot(x,YAct,'black','linewidth',1.5);
hold on
plot(XAct,yval1,'black*');
%hold off

%  Weibull Distribution
plot(x,YWei,'b','linewidth',1.5);
hold on
plot(XWei,yval1,'b*');
%hold off

% GumbellDistribution
plot(x,YGum,'r','linewidth',1.5);
hold on
plot(XGum,yval1,'r*');
%hold off

% GPD Distribution
plot(x,YGpd,'g','linewidth',1.5);
hold on
plot(XGpd,yval2,'g*');
%hold off

title(['Determining Hs of Distributions for same Pr. of Exceedance
'],'FontSize',14,'FontWeight','bold');
ylabel('Probability of Exceedance','FontSize',12,'FontWeight','Bold');
xlabel('Wave height / m','FontSize',12,'FontWeight','Bold');
```

```matlab
legend('Actual
data','Intersect','Weibull','Intersect','Gumbell','Intersect','GPD','Inters
ect');

 % Put the wavelengths for given Pr. Exceedance in a table
 % had to remove GPD because teh size of table is now different
 TExc = table((yval1)',XAct,XGum,XWei,'VariableNames',{'Set Pr. of
Exc','Act. Wave length','Gum Wavelength','Wei Wavelength'});

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Determining fit Actual / Gumbell
% Plotting Actual wavelength / wave length estiamted from Gumbell for the
% SAME Probability of Exceedance
figure
plot(XAct,XGum,'r*');
hold on
coeffsGM = polyfit(XAct,XGum,1);
fittedXGM= linspace(min(XAct), max(XAct),200); % generates 200 equally
spaced points between the minimum and teh maximum
fittedYGM= polyval(coeffsGM,fittedXGM); % generates the points on tehse
generated lines

plot(fittedXGM,fittedYGM,'r-','LineWidth',2);

% plot 45 deg line
XTheor = [3 4 5 6 7];
YTheor = [3 4 5 6 7];
plot(XTheor,YTheor,'black-','LineWidth',0.5);

title(['Linear Regression - Actual Hs and Gumbell
Distribution'],'FontSize',14,'FontWeight','bold');
xlabel('Hs Actual Values / m','FontSize',12,'FontWeight','Bold');
ylabel('Hs Gumbell / m','FontSize',12,'FontWeight','Bold');
axis equal

axis([ 3 7 3 7]);


% Gumbell distribution : Finding the Pearson correlation coefficient
RGum = corrcoef(XAct,XGum);
RGumValue = RGum(1,2);
legend('Values','Regression Line', 'y=x');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Determining fit Actual / Weibull - same as above
figure
plot(XAct,XWei,'b*');
hold on
coeffsW = polyfit(XAct,XWei,1);
fittedXW=linspace(min(XAct), max(XAct),200);
fittedYW= polyval(coeffsW,fittedXW);
plot(fittedXW,fittedYW,'b-', 'LineWidth',2);

% plot 45 deg line
plot(XTheor,YTheor,'black-','LineWidth',0.5);

title(['Linear Regression - Actual Hs and Weibull
Distribution'],'FontSize',14,'FontWeight','bold');
axis equal
```

```matlab
xlabel('Hs Actual Values / m','FontSize',12,'FontWeight','Bold');
ylabel('Hs Weibull Distribution / m','FontSize',12,'FontWeight','Bold');
legend('Values','Regression Line', 'y=x');

% Weibull distribution : Finding the Pearson correlation coefficient
RWei = corrcoef(XAct,XWei); % Pearson correlation coefficient
RWeiValue = RWei(1,2);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Determining fit Actual / GPD- same as above
% had to change teh range since for GPD the max pr. of exceedance recorded
% was 0.80

for i=1:1:n2
XActG(i,1) = XAct(i,1);
end

figure
plot(XActG,XGpd,'gr*');
hold on
coeffsG = polyfit(XActG,XGpd,1);
fittedXG=linspace(min(XActG), max(XActG),200);
fittedYG= polyval(coeffsG,fittedXG);
plot(fittedXG,fittedYG,'gr-', 'LineWidth',2);

% plot 45 deg line
plot(XTheor,YTheor,'black-','LineWidth',0.5);

title(['Linear Regression - Actual Hs and Gen.
Par.Distribution'],'FontSize',14,'FontWeight','bold');
xlabel('Hs Actual Values / m','FontSize',12,'FontWeight','Bold');
ylabel('Hs GPD Distribution / m','FontSize',12,'FontWeight','Bold');
axis equal
axis([ 3 7 3 7]);
legend('Values','Regression Line', 'y=x');

% GPD distribution : Finding the Pearson correlation coefficient
RGPD = corrcoef(XActG,XGpd); % Pearson correlation coefficient
RGPDValue = RGPD(1,2);

 % TExc = table(XAct,XGum,XWei,'VariableNames',{'Act. Wave length','Gum
Wavelength','Wei Wavelength'});

 RPCC= table(RGumValue,RWeiValue,RGPDValue,'VariableNames',{'R Gum','R
Wei','R GPD'});

 %%% Plotting all 3 regression on teh same graph
 figure;
 plot(XAct,XGum,'r*');
 hold on
 plot(fittedXGM,fittedYGM,'r-','LineWidth',1);

 plot(XAct,XWei,'b*');
 plot(fittedXW,fittedYW,'b-', 'LineWidth',1);

 plot(XActG,XGpd,'gr*');
```

```matlab
 plot(fittedXG,fittedYG,'gr-', 'LineWidth',1);

title(['Linear Regression Lines - Actual Hs and
Distributions'],'FontSize',14,'FontWeight','bold');
xlabel('Hs Actual Values / m','FontSize',12,'FontWeight','Bold');
ylabel('Hst Distribution / m','FontSize',12,'FontWeight','Bold');

axis equal
axis([ 3 7.5 3 7.5]);
% to plot the 45 deg line

plot(XTheor,YTheor,'black-','LineWidth',0.5);

legend('Gumbell distribution','Regression. Gumbell','Weibull
distribution','Regression Weibull','GPD','Regression GPD',' Y = X
','Location','southeast');
```

***Program 3** - 100 year return period*
Program Name : Extending Prediction to 100.m

```matlab
%%% to extend the CDF to find 100 year return period
% extend range of X to reach Waveheights 9m

% AA = max(Xsort); % to start where we left off frm other plots
AA = h0;
testHs = (AA:0.1:10);

% X Excess = XE = X - X threshold (h0)
for i=1:1:length(testHs);
    TestE(i,1) = testHs(1,i) - h0;
end

% WEIBULL %%%%%%%%%%%%%%%%%%%%
% EXTEND THE CDF BASED ON Weibull AND USING A AND B FOUND PREVIOUSLY
CDFWeiT = wblcdf(TestE,A,B);

% CREATE THE PROBABILITY OF EXCEEDANCE Weibull Qwei = 1- CDFWei
% AND RETURN PERIOD TR

for i = 1:1:length(testHs)
QWeiT(i,1) = 1- CDFWeiT(i);
TRWeiT(i,1) = 1/(nr*QWeiT(i,1)); %%%%%% NUMBER OF YEARS
end

hold on
plot(TRWeiT,testHs,'*b');
plot(TRWeiT,testHs,'b','linewidth',0.25);

% GPD %%%%%%%%%%%%%%%%%%%%%%%%
CDFgpdT = gpcdf(TestE,k,sigma,theta);

% CREATE THE PROBABILITY OF EXCEEDANCE GPD = 1- CDFgpd
% and RETURN PERIOD
for i = 1:1:length(testHs)%
QgpdT(i,1) = 1- CDFgpdT(i); % Prob. of exceedance = 1 - Cum Distr Function
TRgpdt(i,1) = 1/(nr*QgpdT(i,1));  %%%%%% NUMBER OF YEARS
```

```matlab
end
plot(TRgpdt,testHs,'g','linewidth',0.25);


%%%% Gumbell
% plot(TRGumT,testHs,'-r');
% Change the range of X from 1 - 10m
XExt=(1:0.5:10); % we have etended XE to cover range. Compare plots with
"return Perriod 2"
MExt= length(XExt);


alpha = sqrt(6)* StDv / 3.142 ;
U = (Xmean - (0.5772*alpha));


XExtSort=sort(XExt, 'descend');
% CDF = exp [-exp[-(xi - u/alpha)] ;


for i = 1:1:MExt
constE(i)= (-(XExtSort(i) - U) / alpha);
%valE(i) = exp(-(XExtSort(i) - U) / alpha);
valE(i)= exp(constE(i));


PDFGum(i,1)= exp(constE(i)-valE(i))/alpha; % PDF Gumbel Distribution -
Check


CDFGumE(i,1) = exp(-valE(i)); % CDFGum = CDF , Non Exceedance Probability P
= 1 - Q
QGumE(i,1) = 1 - CDFGumE(i,1); % QGum = Pr. of exceedance
TRGumE(i,1) = 1 / (nr *QGumE(i,1)); % Return Period = 1 / Pr. of exceedance
end

plot(TRGumE,XExtSort,'r*');
plot(TRGumE,XExtSort,'r','linewidth',0.25);


grid on;
xline =100;
```

***Program 4** - Battjes method*
Program Name : Evaluating Hmax for n years.m

```matlab
% TO EVALUATE Hmax for n years Using Battjes method
% Using Maria Swan Data to do scatter plots of Hs / Zero Upcross period
%%%% EXTRACT DATA FOR ONE YEAR

%TO READ ROSARIO MARIA SWAN DATA COORDINATES :  FROM JOEL MAIL LON : 14.168
LAT: 36.0812
dataSwanM = dlmread('timeseries_RosarioMariaSwan 30.11.txt');

% Extracting data for 1 year - year chosen 2018
index=find(dataSwanM(:,1)==2018);
dataSwanS= zeros(numel(index),8);
dataSwanS = dataSwanM(index,:);

% transferring the data to individual vectors
MariaDate = datenum([dataSwanS(:,1) dataSwanS(:,2) dataSwanS(:,3)
dataSwanS(:,4) zeros(size(dataSwanS, 1), 1) zeros(size(dataSwanS, 1), 1)]);
```

```matlab
MariaH = dataSwanS(:,5); % Sea surface Wave significant height
MariaHs = MariaH*1.5; % Corrected Value

MariaDir = dataSwanS(:,6); % Sea surface wave FROM direction

MariaT = dataSwanS(:,8) ; % Sea surface wave MEAN period from variance
spectral density first frequency moment
MariaTm = (MariaT*0.8)+1.25; % putting in the correction factor from
validation


%removing dummy numbers
index = find(MariaTm <=0);
MariaTm(index)=NaN;
N=length(index);

%SWH and peak wave period
Table = [];

Counter1 = 0;
Counter2 = 1; % Hs index (row index)
Counter3 = 1; % T cross period (column index)

% interetsing how it works.
% starts off with first values of Tm and Wh , and it goes through I
% counter adds 1 - and saved in table
% and it moves to the next WH and goes through the list

% TABLE has the data separated based on Hs / Tm
for Tm = 0.25:0.5:8.75; % wave mean period classes
    for WH = 0.25:0.5:7.75; % wave height classes
        for I = 1:1:length(MariaTm)
            if MariaHs(I,1)<=WH+0.25 && MariaHs(I,1)>WH-0.25 &&
MariaTm(I,1)<=Tm+0.25 && MariaTm(I,1)>Tm-0.25
            Counter1 = Counter1 +1;
            end
        end

    Table(Counter2,Counter3)= Counter1; % number of entries that fall in
box
    Counter1 =0;
    Counter2=Counter2+1;
    end
Counter3 = Counter3+1;
Counter2 = 1;
end

Tm = (0.25:0.5:8.75); %T cross period %6.75
WH = (0.25:0.5:7.75); %Hs %5.75

%%%%%% Probability of exceedance using Rayleigh distribution
% we need to find teh probability of exceedance for every h for every Hs
% WH =  Significant waveheight
% HVal = value of waveheight h
% find the probability a waveheight is greater than Hval for sig. wave WH
% height HVal  = exp (-2(HVal/ WH)^2)

%%% Value of waveheight we want to find Pr. Exceedance
```

```matlab
HVal = 0.25:0.5:9.75; % Value H 7.75
WH = 0.25:0.5:7.75; % Hs Significant Wave Height 5.75

% Do to table of the Pr. of exceedance
% HS= x and Hval = y

for i = 1:1: length(WH);
    for j= 1:1: length(HVal);
        HValT = HVal(1,j);
        HST = WH(1,i);
            btop = 2*(HValT^2);
            bbottom = (HST^2);
            b = (-btop/bbottom)
            Prob = exp(b);
 PE(i,j) = Prob; % probability of exceedance h = H axis Hs Y axis
    end
end

%%%%%%%%%% TO DETERMINE THE NUMBER OF WAVES THAT OCCUR IN 3 HOURS
% Number of waves in 3 hours = 3 hrs x 60 x 60 / Tz
% Range Tz  = 1 : 1 : 13 sec

Tm = 0.25:0.5:8.75;
for i= 1:1:numel(Tm(1,:));
    Tz(i)= Tm(1,i);
Nr(i,1)= 60*60*3 / Tz(i);
end

T1 = table(Tz',Nr,'VariableNames',{'Tz','Nr. of waves in 3 hrs'});

% Nr. of Waves in Hs/Tp window in 1 year
% = number of 3 hour periods falling in that pocket x total nr of waves in
that period

% PRW = table HsTp x correpsonding number of waves occuring in 3 hours
table Nr

Tm = (0.25:0.5:8.75); %T cross period %6.75
WH = (0.25:0.5:7.75); %Hs %5.75

for i = 1:1:numel(Tm(1,:));
    for j = 1:1:numel(WH(1,:));
    PRW (j,i) = Table(j,i)* Nr(i,1);
end
end

% adding the values in each row will give the Nr of waves occuring in one
% year for each Hs

for i = 1:1:numel(WH(1,:))
PRWSum(i,1) = sum(PRW(i,:));
end

T2 = table(WH',PRWSum,'VariableNames',{'Hs','Nr. of waves Value Hs'});

% therefore total number Hs (3 hour packest) in 1 year is teh sum of the
above
totalwaves = sum(PRWSum(:,1));
```

```matlab
%%%%%%%%%% TO FIND THE NUMBER OF WAVES IN 1 YEAR THAT EXECEED H
% We have found the total amount of waves that fall within Hs
% We have found the probability waveheight is greater than H for each Hs
% (Table PE)
% we multiply nr. of waves in each Hs with probability of exceedance and
% add

% m = Number of waves exceeding h
% Recall HVal = 0.25:0.5:7.75; that is h

for j =1:1:numel(HVal(1,:));
    for i = 1:1:numel(WH(1,:));
        m(i,j) = PE(i,j)*PRWSum(i,1); % m = number of
    end
        totalm(1,j) = sum(m(:,j)); % probable number of waves exceeding
Hval
end

T3 = table(HVal',totalm','VariableNames',{'h','Nr. of waves exceeding h'});

% Plotting the values of pr. nr of waves execcedance Hval and Hval
plot(HVal,totalm','b');

title(['Probable Nr. of Waves Exceeding h / Battjes
method'],'FontSize',14,'FontWeight','bold');
xlabel('Waveheight / m');
ylabel('Probable Nr. of Waves Exceeding h ');
box on;
ax = gca;
ax.FontSize = 12;
```