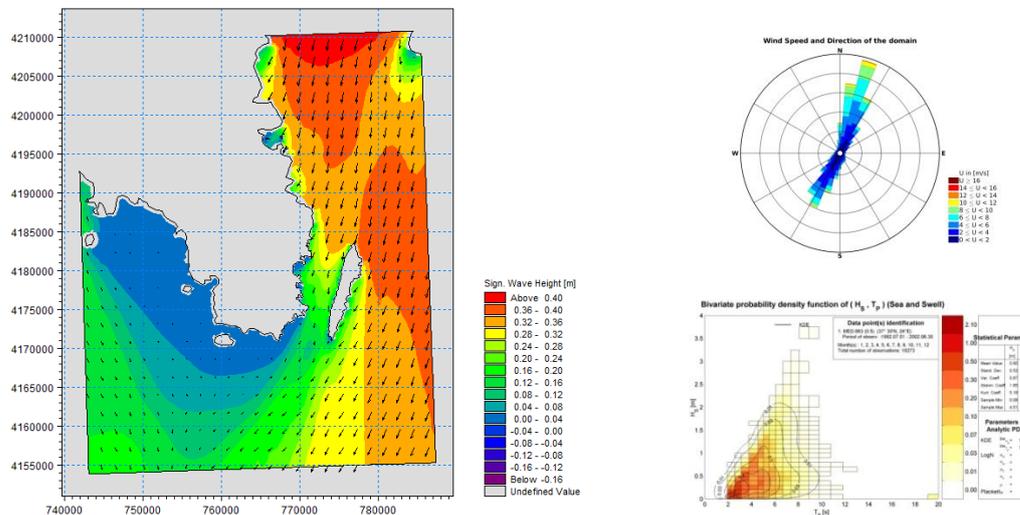




National Technical University of Athens  
Interdisciplinary Postgraduate Program  
Ship and Marine Technology

Wave climatology and extreme value analysis in  
coastal waters of Greek sea areas with application to  
Sounio nearshore region



Author: Korina Mosiou (08119816)

Thesis Committee

Supervisor: Prof. Konstantinos Belibassakis

Dr. Takvor Soukissian (HCMR)

Prof. Helena Flocas (NKUA)

Athens, 2021

## **Acknowledgements**

I would like to sincerely thank my thesis supervisor, Professor Kostas Belibassakis of the School of Naval Architecture and Marine Engineering at NTUA, for his invaluable guidance throughout the process of writing this thesis.

Additionally, I express my sincere appreciation to the rest of the thesis committee, Dr Takvor Soukissian at HCMR and Professor Helena Flocas of the Department of Physics, Section of Environmental Physics - Meteorology, at NKUA.

I would also like to thank Dr. Flora Karathanasi for providing valuable assistance regarding technical matters of MIKE 21, as well as Mr. Elias Mousoulis, sales and marketing manager of MIKE by DHI products in Greece for his guidance and immediate replies.

My deepest gratitude goes to my family and friends for their constant support and encouragement during this project.

## Abstract

The present thesis aims at the simulation of wave conditions in the nearshore area of the Saronic Gulf between Attica and Kea island, with special emphasis/focus at Makronisos channel. The results will be based on a wave transformation making use of long timeseries of wave and wind at the margin of the open sea from ECMWF's ERA5 database, in combination with bathymetry and coastline data of the area in question as well as the systematic usage of MIKE 21 Spectral Waves model (DHI). The climatology of the open sea area will be presented in comparison with that of selected nearshore points utilizing appropriate distributions. An unstructured mesh with finer resolution at the nearshore area of interest and gradually coarser moving to the open sea margin will be employed for the simulation of the wave climate. On a final note, the results of the simulation will be analyzed following extreme value theory in order to assess the possible future extreme events with regards to the significant wave height. According to the results, wave conditions inside Makronisos channel are considered smooth for the period of 1980-2019. Moreover, regarding the extreme value analysis, it is estimated that O2 has the highest significant wave heights for the selected return periods while TP4 has the lowest.

## Περίληψη

Η παρούσα εργασία έχει σκοπό τη προσομοίωση κυματικών συνθηκών στην παράκτια περιοχή του Σαρωνικού κόλπου μεταξύ της Αττικής και της Κέας, με ιδιαίτερη έμφαση στο κανάλι της Μακρονήσου. Τα αποτελέσματα θα βασιθούν σε μετασχηματισμό κυματικών κάνοντας χρήση μακροχρόνιων δεδομένων κύματος και ανέμου στο όριο της ανοικτής θάλασσας από τη βάση ERA5, σε συνδυασμό με δεδομένα βαθυμετρίας και ακτογραμμής της εξεταζόμενης θαλάσσιας-παράκτιας περιοχής και τη συστηματική χρήση του φασματικού κυματικού μοντέλου MIKE SW (DHI). Η κλιματολογία θα παρουσιαστεί συγκριτικά για το όριο της ανοικτής θάλασσας και επιλεγμένων σημείων κοντά στην ακτή κάνοντας χρήση κατάλληλων κατανομών. Θα χρησιμοποιηθεί υπολογιστικό πλέγμα μεταβλητής ανάλυσης αναφορικά με τη προσομοίωση του κυματικού κλίματος, που θα περιλαμβάνει αδρότερη ανάλυση στα ανοικτά και πυκνότερη παράκτια στην περιοχή ενδιαφέροντος. Τέλος θα πραγματοποιηθεί και ανάλυση ακραίων τιμών κυματισμών για το σημαντικό ύψους κύματος. Σύμφωνα με τα αποτελέσματα, οι κυματικές συνθήκες εντός του στενού της Μακρονήσου είναι ήπιες για τη περίοδο 1980-2019. Επιπλέον, αναφορικά με την ανάλυση ακραίων τιμών, υπολογίζεται ότι το σημείο με τα μεγαλύτερα σημαντικά ύψη κύματος για τις επιλεγμένες περιόδους επιστροφής είναι το O2 σε αντίθεση με το TP4 που παρουσιάζει τις χαμηλότερες τιμές σημαντικού ύψους κύματος.

# Table of contents

	Abstract	2
<b>1</b>	Introduction	5
<b>1.1</b>	The importance of wave climatology and extreme value analysis	5
<b>1.2</b>	Description of the study area	5
<b>1.3</b>	Review of climatic data	9
<b>1.4</b>	Wave prediction models and basic wave models for downscaling	10
<b>1.5</b>	Study aim and outline of thesis	13
<b>2</b>	Available Data Sources	14
<b>2.1</b>	Geographic Data	14
<b>2.1.1</b>	GSHHG high-resolution shoreline	14
<b>2.1.2</b>	GEBCO and EMODnet bathymetries	14
<b>2.2</b>	Climatic Data	16
<b>2.2.1</b>	ERA5 Wind and Wave fields	16
<b>3</b>	Statistical analysis of offshore wind and wave data	17
<b>3.1</b>	Univariate distributions of offshore wave data	17
<b>3.2</b>	Bivariate distributions of offshore wave data	23
<b>4</b>	Model Setup	25
<b>4.1</b>	Mike 21	25
<b>4.2</b>	Computational Grid	26
<b>4.3</b>	MIKE 21 Spectral Wave module setup	28
<b>4.3.1</b>	Stationary Runs	29
<b>4.3.2</b>	Three year hindcast (1980-1982)	46
<b>5</b>	Wave climatology inside Makronisos channel and extreme value analysis	49
<b>5.1</b>	Univariate distributions of nearshore data	49
<b>5.2</b>	Bivariate distributions of nearshore data	57
<b>5.3</b>	Wave climatology of the target points in Makronisos channel	60
<b>5.4</b>	Multivariate regression for significant wave height $H_s$	65
<b>5.5</b>	Extreme value analysis using the POT method	70
<b>6</b>	Discussion	79
	Appendix A - Probability distributions of linear and circular variables	82
	Appendix B - Regression analysis	85
	Appendix C - Extreme value analysis methods	88
	References	92



# **1. Introduction**

## **1.1. The importance of wave climatology and extreme value analysis**

The need for high accuracy predictions of the wave climate has been of utmost importance since long time ago. Initially, nineteenth century pioneers of modern theoretical fluid dynamics, namely Lagrange, Airy, Stokes and Rayleigh, studied extensively the properties of surface waves (Phillips, 1977). Later, during the Second World War, wave prediction started to evolve given there was an increased need for precise information regarding the sea state during landing operations. In fact, Sverdrup and Munk (1947) were the first to use empirical laws for wind sea and swell in order to produce operational predictions of sea state (Komen et al., 1994).

Waves are the result of the interaction between atmosphere and ocean, in which momentum and heat are exchanged. There is a significant need for their prediction since they affect maritime transport by means of shipping routes, shape the coastline (sediment transport, whether erosion, resuspension or deposition) which in turn has a significant impact on coastal structures, as well as pose a serious threat to the function of offshore oil and gas industry. Additionally, during the last decades, there has been a rising concern about the effects of climate change on coastal areas. Therefore, it is crucial to examine the occurrence of extreme wave events since they are the main agent of stress on offshore and/or coastal structures. As Hemer et al. (2012) report, there has been an increased interest in investigating wind waves because of their contribution to extreme sea level values at coastal areas and of the fact that they inflict serious damage to infrastructure near the sea.

## **1.2. Description of the study area**

This thesis studies the wave climatology and the occurrence of extreme wave events in the Sounio nearshore region. The selected study area consists of the southeastern part of Attica and the island of Makronisos in terms of land, while the wet/sea part consists of the exit of the Saronic Gulf in the Aegean. The significance of this particular region of Attica lies in the touristic development that has taken place during the last decades as well as its proximity to Athens city and Piraeus port. A high number of resorts can be found following the coastline from cape Sounio up to city of Lavrio, but it is also worth noting that the Olympic Marina, situated south of Lavrio port, is a major attraction for yacht-oriented tourism. Furthermore, the Poseidon temple at Sounio cape is an archeological site with very frequent visits and touristic interest. In addition, a review of the literature on the location and purpose of the present work reveals that the general scientific interest has gravitated more towards the analysis of the wave climate in Greek seas, and only a few examples of work have made high resolution wave simulations for nearshore areas close to

Attica. For example, in the work of Soukisian et al. (2008), data covering the period of 1995-2004 have been analyzed to describe the wind and wave climate of the Ionian and Aegean Seas. Similar to that, Zacharioudaki et al. (2015) have investigated the wave climate of the Hellenic Seas as well as the climate extremes during 1960-2001. A study with great proximity to the area of interest is the work of Belibassakis & Karathanasi (2017), in which the hydronamics and circulation of Varkiza bay are simulated, taking into consideration the impact of high waves. In addition, Lavidas & Venugopal (2017) have employed a fine resolution wave model to simulate the wave characteristics of Greek coastal regions, including a location which is adjacent to the present thesis domain. A case study with similar morphological characteristics is that of Kontoyannis et al. (2015), in which the authors estimated the hydrokinetic renewable energy of Euripus strait. For all the above reasons, five (5) points located at Sounio nearshore area and between Attica and Makronisos island (Makronisos channel) were chosen to simulate the wave conditions (significant wave height  $H_s$ , peak period  $T_p$ , mean wave direction  $\theta$ ), as each point represented proximity to significant touristic locations (e.g. Sounio cape, Olympic Marina and Lavrio city) (Fig. 1).

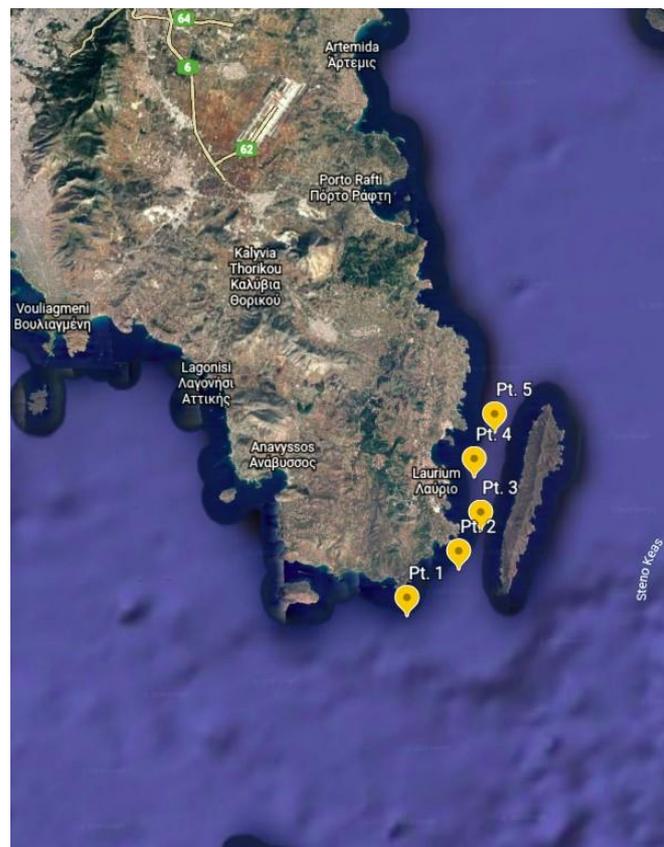


Figure 1 Selected points for wave simulation (picture taken from Google Earth)

Sounio nearshore region is a part of Saronic Gulf and the actual study area is characterised by rather low depths since it involves a part of the continental shelf. For example between Attica and Makronisos island, the maximum depth that occurs is 50 meters. However, on the southeastern corner of the domain the depth increases to 350-400 meters. Nevertheless, this selected region exhibits the unique characteristic of Makronisos channel.

Regarding the wave climatology of the area, according to Medatlas Project (Wind and Wave Atlas of the Mediterranean Sea, 2004), wave heights (significant wave height  $H_s$ ) rarely exceed 1 meter (mean value at 0.6 m), peak period ( $T_p$ ) has high probability around 3-6 seconds and the most probable wave directions are from  $0^\circ$ - $30^\circ$  and from  $210^\circ$ - $230^\circ$ . The abovementioned features can be viewed in Fig. 2-5. Medatlas results are produced by statistical analysis (with either univariate or bivariate distributions) over a ten year time series at a specific offshore point [ $37^\circ 30'N$ ,  $24^\circ E$ ]. As seen in Fig. 3, the highest occurrence of waves in this ten year period lies between approximately 0.3 m  $H_s$  and 3 s  $T_p$ .

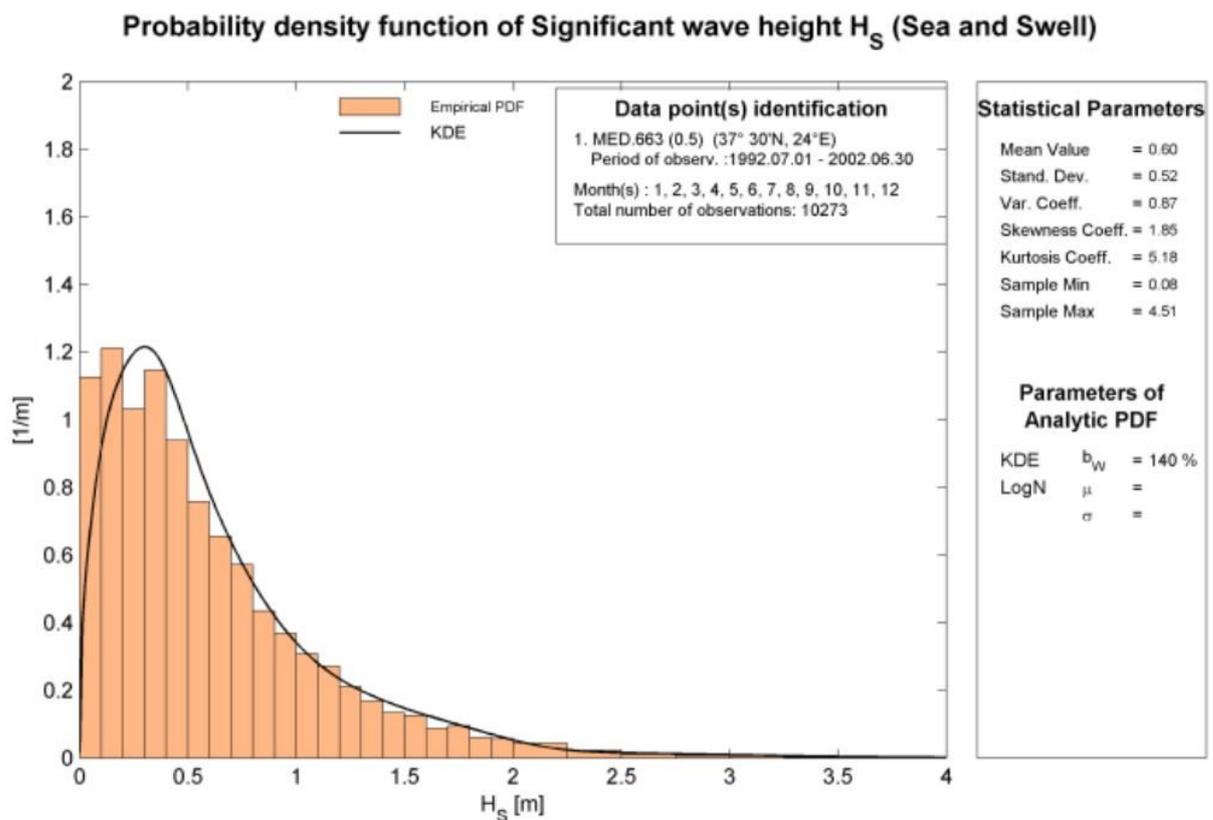


Figure 2 Probability density function of Significant wave Height (Medatlas)

### Probability density function of Peak period of 1d spectra

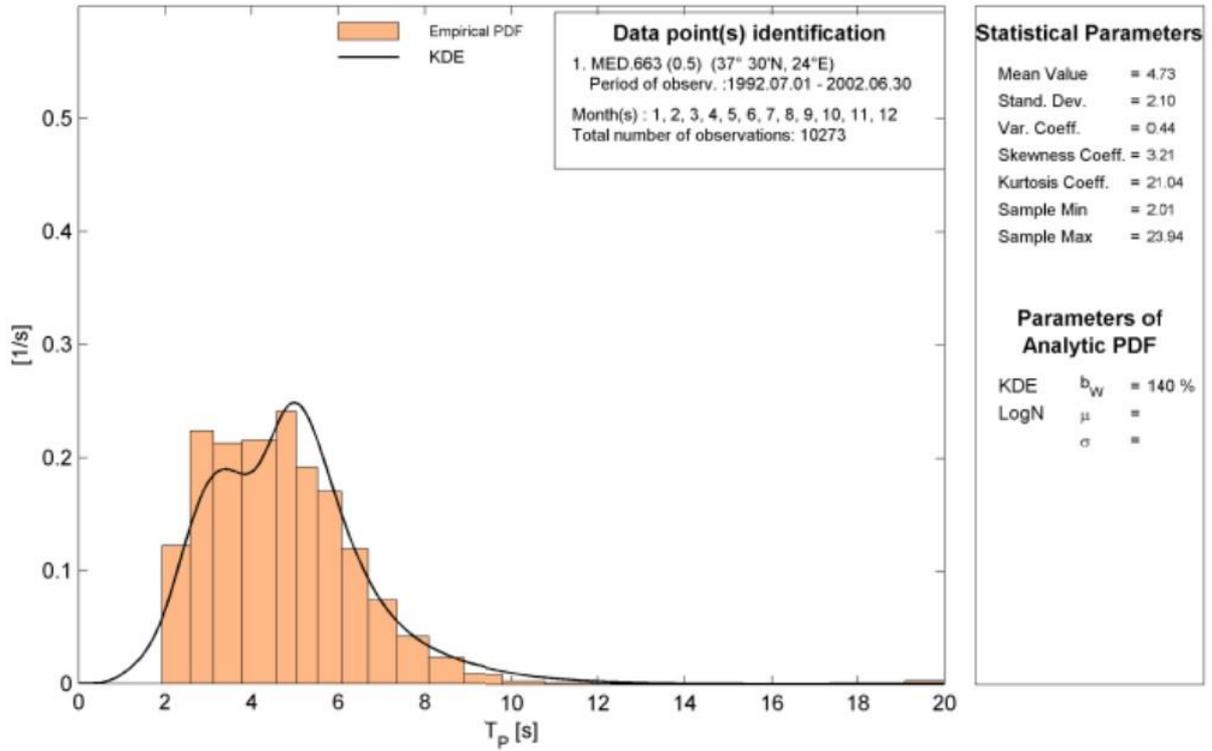


Figure 3 Probability density function of Peak Period (Medatlas)

### Bivariate probability density function of ( $H_S$ , $T_P$ ) (Sea and Swell)

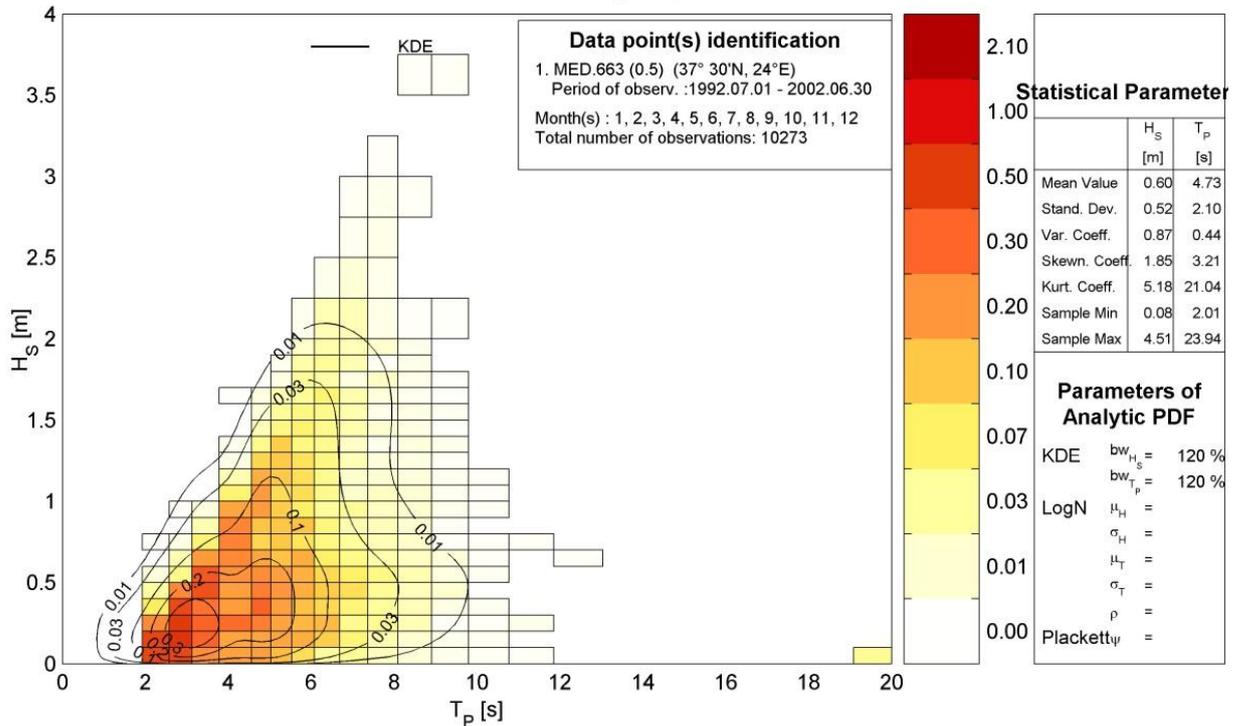


Figure 4 Bivariate probability density function of (Hs, Tp) (Medatlas)

## Polar histogram (%) of Mean wave direction $\theta$ (Sea and Swell)

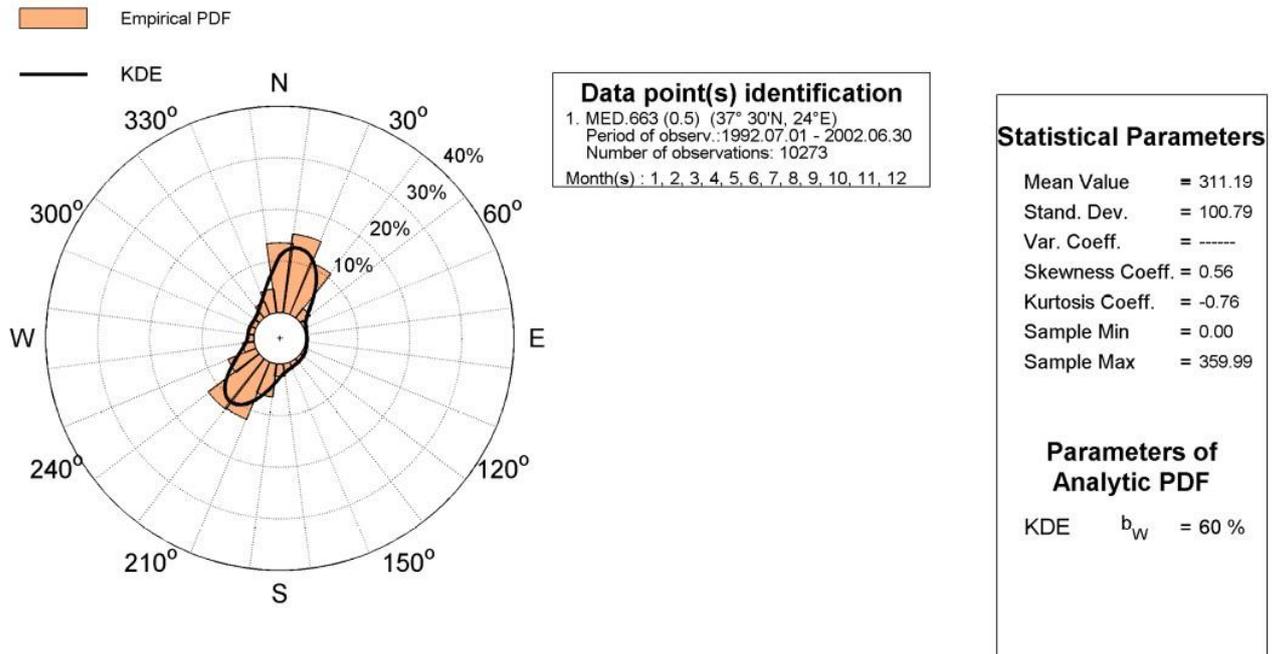


Figure 5 Polar histogram (%) of Mean wave direction  $\theta$ (Medatlas)

### 1.3. Review of climatic data

Meteorological and oceanographic measurements for wind and waves may originate from four different sources:

- wave observations from voluntary observing ships
- in-situ data from buoys and platforms
- data from remote sensing including scatterometers and altimeters
- data from numerical model simulations

Each and one of these categories has their consequent advantages and disadvantages. However, in this study only the numerical model data which correspond to the study area will be discussed.

First, on a global scale, the European Centre for Mid-Range Weather Forecast (ECMWF), <https://www.ecmwf.int>, produces meteorological as well as oceanographic data which have been routinely used. There have been quite a few reanalysis datasets delivered, the latest being ERA-40 (Uppala et al., 2005), ERA-Interim (Dee et al., 2011) and the most recent ERA5 (Hersbach et al., 2020). To begin with, ERA-40 was a global reanalysis dataset containing 6-hour fields of atmosphere, land and ocean-wave state for the period of December 1957 to August 2002 with  $2.5^\circ \times 2.5^\circ$  resolution for atmosphere products and  $1.5^\circ \times 1.5^\circ$  for wave products. In

2005, ERA-Interim was launched to create a link between ERA-40. ERA-Interim was the transition from ERA-40, and it provided 3-hour fields for atmosphere land and ocean covering the period since 1979 until August 2019 with  $\sim 0.75^\circ \times 0.75^\circ$  resolution for atmosphere and  $1.0^\circ \times 1.0^\circ$  for wave results. The latest improved reanalysis launched from ECMWF is ERA5, which delivers hourly estimates for a large number of atmospheric, land and ocean-wave parameters. The fields given cover the period of 1979 to present, with atmospheric resolution at  $0.25^\circ \times 0.25^\circ$  and  $0.5^\circ \times 0.5^\circ$  for ocean waves.

Second, Copernicus Marine Environment Monitoring Service (CMEMS) [marine.copernicus.eu](http://marine.copernicus.eu) provides basin scale wave data for the Mediterranean sea, produced by a high-resolution operational wave forecasting system, Med-waves, which is implemented by the Hellenic Centre for Marine Research (HCMR) (Ravdas et al., 2018). Med-waves uses WAM Cycle 4.5.4 (Günther and Behrens, 2012) and produces hourly estimates for wave parameters with a  $1/24^\circ$  horizontal resolution starting from 1st February 2006 until today.

Additionally, the POSEIDON system of HCMR provides in situ measurements of marine parameters like significant wave height and wind speed at 10 meters measured from buoys at several locations in the Aegean and Ionian seas. Apart from that, POSEIDON also produces atmospheric, hydrodynamic and wave daily forecasts covering the whole Mediterranean.

#### **1.4. Wave prediction models and basic wave models for downscaling**

As stated in the beginning of this chapter, operational wave prediction started in the late 40's, prior to D-Day invasion in Normandy in June 1944. Following that first attempt to predict the sea state, scientists achieved to develop wave prediction models which are classified as first, second and third generation. During the late 50's, Gelci et al. (1957) operated the first numerical spectral wave model. Notwithstanding, significant work regarding wave generation (Philips, 1957; Miles, 1957) and non-linear energy transfer due to wave-wave interactions (Hasselmann, 1962) has contributed to a better understanding of wave generation modeling.

Examples of first generation prediction models are VENICE by Cavaleri and Rizzoli (1981) and the model developed by the Meteorological Research Institute (MRI) in Japan (SWAMP Group, 1985). This class of models did not take into account non-linear interactions (only partly in the growth rate  $\mu$ ), and therefore there was some uncertainty whether the physics of wave generation were expressed correctly (Massel, 2017), which consequently led to the development of second generation wave models. Examples of second generation wave models are the Scripps Institution of Oceanography Model (DNS) which is a discrete spectral model, a

parametric wave model by Hasselmann et al. (1976), and Hybrid Parametrical model (coupled hybrid model) by Günther et al. (1979a,b).

Eventually, third generation wave models started to develop during the mid 80's in order to overcome the limitations that first and second generation models posed, such as the fact that the later did not respond well to all types of wind fields (e.g. extreme winds), as stated in the SWAMP (1985) wave model intercomparison study. An example of a widely used third-generation wave model is WAM (WAMDI Group, 1988) which contained specific improvements, namely a parameterization of the exact non-linear transfer source function which included the same number of degrees of freedom as the spectrum itself and the specification of the unknown dissipation source function, as proposed by Komen et al. (1984), to close the energy balance. Furthermore, Simulating WAVes Nearshore (SWAN) is another example of third generation wave model, developed by Delft University of Technology (Booij et al., 1999), which was designed specifically for nearshore wave modeling. SWAN integrates all the necessary physics to simulate effectively wave phenomena in shallow waters.

Both WAM and SWAN models are very useful tools in terms of producing regional solutions via dynamic downscaling. Dynamic downscaling is a method to obtain high-resolution numerical results from a coarser resolution setup which is very important in coastal research and engineering studies.

The basis for all numerical prediction models is the radiative transfer equation(or spectral energy balance equation) with the corresponding source-sink terms, is presented in Cartesian geometry below (Massel, 2017):

$$\begin{aligned}
& \frac{\partial}{\partial t}(C_p C_g \hat{S}) + C_g \cos\theta \frac{\partial}{\partial x}(C_p C_g \hat{S}) + C_g \sin\theta \frac{\partial}{\partial y}(C_p C_g \hat{S}) \\
& + \frac{C_g}{C_p} \left[ \sin\theta \frac{\partial C_p}{\partial x} - \cos\theta \frac{\partial C_p}{\partial y} \right] \frac{\partial}{\partial \theta}(C_p C_g \hat{S}) \\
& = C_p C_g Q(\omega, \theta, \mathbf{x}, t)
\end{aligned} \tag{1}$$

where  $\frac{\partial}{\partial t}(C_p C_g \hat{S})$  is the local evolution of the spectrum  $\hat{S}$  in time,  $C_g \cos\theta \frac{\partial}{\partial x}(C_p C_g \hat{S})$  and  $C_g \sin\theta \frac{\partial}{\partial y}(C_p C_g \hat{S})$  represent the propagation, while  $\frac{C_g}{C_p} \left[ \sin\theta \frac{\partial C_p}{\partial x} - \cos\theta \frac{\partial C_p}{\partial y} \right] \frac{\partial}{\partial \theta}(C_p C_g \hat{S})$  describes the combination of refraction and shoaling.  $C_p$  is the wave phase velocity and  $C_g$  is the wave group velocity. On the

right hand side of eq. (1) the term  $Q(\omega, \theta, \mathbf{x}, t)$  denotes the net transfer of energy to, from or within the spectrum such as

$$Q = Q_{in} + Q_{nl} + Q_{dis} + Q_{bf} + Q_{brk} \quad (2)$$

where  $Q_{in}$  is the atmospheric input,  $Q_{nl}$  are the non-linear interactions within the spectrum,  $Q_{dis}$  is the dissipation due to white-capping,  $Q_{bf}$  is the shallow water decay due to bottom friction and  $Q_{brk}$  is the depth limited wave breaking.

Additionally, the radiative transfer equation which describes the development of the two-dimensional ocean wave spectrum  $F(f, \vartheta, \varphi, \lambda, t)$  in spherical geometry as written in WAMDI(1988) is:

$$\frac{\partial S}{\partial t} + (\cos\varphi)^{-1} \frac{\partial}{\partial \varphi} (\dot{\varphi} \cos\varphi S) + \frac{\partial}{\partial \lambda} (\dot{\lambda} S) + \frac{\partial}{\partial \theta} (\dot{\theta} S) = Q \quad (3)$$

where  $f$  is the frequency,  $\vartheta$  is the direction,  $\varphi$  is the latitude,  $\lambda$  is the longitude,  $Q$  is the net source function expressing the change of energy of a propagating group of waves and

$$\dot{\varphi} = \frac{d\varphi}{dt} = c_g R^{-1} \cos\theta \quad (3.1)$$

$$\dot{\lambda} = \frac{d\lambda}{dt} = c_g \sin\theta (R \cos\varphi)^{-1} \quad (3.2)$$

$$\dot{\theta} = \frac{d\theta}{dt} = c_g \sin\theta \tan\varphi R^{-1} \quad (3.3)$$

The above three terms describe the rate of change of the position and propagation direction of a wave packet traveling along a great circle.  $R$  is the earth radius and  $c_g = g/4\pi f$  is the group velocity for finite depth. Since eq. (3), (3.1),..., (3.3) are referring to deep water propagation, the source function  $Q$  is the sum of  $Q_{in}$  (wind input)  $Q_{nl}$  (non-linear interactions) and  $Q_{dis}$  (dissipation). The extension of eq. (3) to include wave propagation in shallow waters is described in WAMDI (1988).

However, the radiative transfer equation (eq. (1), (3)) cannot be used in the presence of ambient currents. In such a case, the energy density spectrum  $E(\sigma, \vartheta)$  cannot be conserved and as a consequence the action density spectrum  $N(\sigma, \vartheta)$  is preferred due to its conservation property (Booij et al., 1999). The action density spectrum  $N(\sigma, \vartheta)$  is connected with the energy density spectrum  $E(\sigma, \vartheta)$  via the following expression

$$N(\sigma, \theta) = \frac{E(\sigma, \theta)}{\sigma} \quad (3.4)$$

where  $\sigma$  is the intrinsic frequency [ $\sigma^2 = gk \tanh(kh)$ ]. For this reason, the spectral action balance equation is preferred to render the evolution of the wave spectrum under conditions of a non-stationary current

$$\frac{\partial N}{\partial t} + \nabla \cdot (\bar{v}N) = \frac{S}{\sigma} \quad (3.5)$$

which is described more thoroughly in chapter 4.

## 1.5. Study aim and outline of thesis

In this thesis, wave climatology and the occurrence of extreme wave events in the Sounio nearshore region are investigated during a forty year period (1980-2019). The structure of the rest of this document is as follows. In chapter 2, an extensive description of the available data sources is being listed. In chapter 3, the statistical analysis of wind forcing and wave offshore data is being presented. To continue, chapter 4 includes a detailed report regarding the transformation of offshore to nearshore wave fields using MIKE 21 Spectral Waves. Moreover, the climatology inside Makronisos channel along with the analysis of extreme events are presented in chapter 5. Additionally, the conclusions regarding the present work as well as suggestions for future research are documented in chapter 6. Lastly, a theoretical background concerning the probability distributions of linear and circular variables, regression analysis and extreme value analysis methods are found in appendices A, B and C respectively.

## 2. Available data sources

In order to successfully run a wave simulation, there are some important input data that need to be collected and processed. The inputs consist of the geographical details of the selected region (bathymetry and coastline), and the initial and boundary conditions e.g. the wind and wave fields. The first, bathymetry and coastline, are essential for the making of the computational grid in which the simulation will take place. For studies including the simulation of wave conditions in shallow water areas with distinctive geography, high-resolution shoreline and bathymetry data are indispensable in order to produce accurate results.

### 2.1. Geographic Data

#### 2.1.1. GSHHG coastline

The Global Self-Consistency, Hierarchical, High-Resolution Geographic Data Base (GSHHG) was created by Paul Wessel and Walter H.F. Smith (Wessel P & Smith WHF, 1996). GSHHG high-resolution shoreline data set is a combination of two public domain data sets, World Data Bank II (WDB II) and World Vector Shoreline (WVS). WDB II is a set which contains information about coastlines, lakes, political boundaries and rivers with a scale of 1:3.000.000. WVS set solely includes shorelines along the ocean/land interface (but does not include lakes), with a scale of 1:100.000 which indicates that WVS is better in terms of resolution in comparison with WDB II. The data set has been carefully processed as a means to be accurate and without inconsistencies. GSHHG can be found in <https://www.soest.hawaii.edu/pwessel/gshhg/> and is available in netCDF 4, ESRI shapefile and native binary formats, and come in five resolutions: a) full (original), b) high, c) intermediate, d) low and e) crude resolution.

#### 2.1.2. EMODnet and GEBCO bathymetries

Two bathymetric datasets are presented in this section, European Marine Observation and Data Network (EMODnet) Digital Terrain Model bathymetry and General Bathymetric Chart of the Oceans (GEBCO) 2020 Grid.

First, as the name suggests, EMODnet is a network of organizations which is supported by the EU's integrated maritime policy. EMODnet bathymetric data extend from 36°W,15°N to 43°E,90°N. There are two available versions of the Digital Terrain Models, the 2016 version with 1/8 x 1/8 arc minutes resolution, and the 2018 version with 1/16 x 1/16 arc minutes resolution. To produce the DTM, an array of data sources are employed including satellite derived bathymetry products from Landsat 8 images, more than 9400 bathymetric survey data sets and other DTMs from 24 countries which border on European seas (EMODnet Bathymetry

Consortium, 2018).EMODnet 2018 DTM is available in an array of formats including netCDF, xyz, ASCII and GeoTiff.

Second, GEBCO\_2020 Grid is a continuous global terrain model for land and ocean, as well as it is the latest global bathymetric product released by GEBCO and has been developed through the Nippon Foundation-GEBCO Seabed 2030 Project (GEBCO Compilation Group 2020, GEBCO 2020 Grid). The grid resolution is 15 arc seconds of 43200 rows x 86400 columns, giving 3,732,480,000 data points. The gridded bathymetric datasets which derive from multibeam data to a great extent, have been created by the four Seabed 2030 Regional Centers. GEBCO bathymetry ([https://www.gebco.net/data\\_and\\_products/](https://www.gebco.net/data_and_products/)) are available in netCDF, Data GeoTiff and ESRI ASCII raster formats.

The comparison between EMODnet and GEBCO bathymetries of the study domain can be viewed in Fig. 6, in which the GSHHG coastline is also visualised.

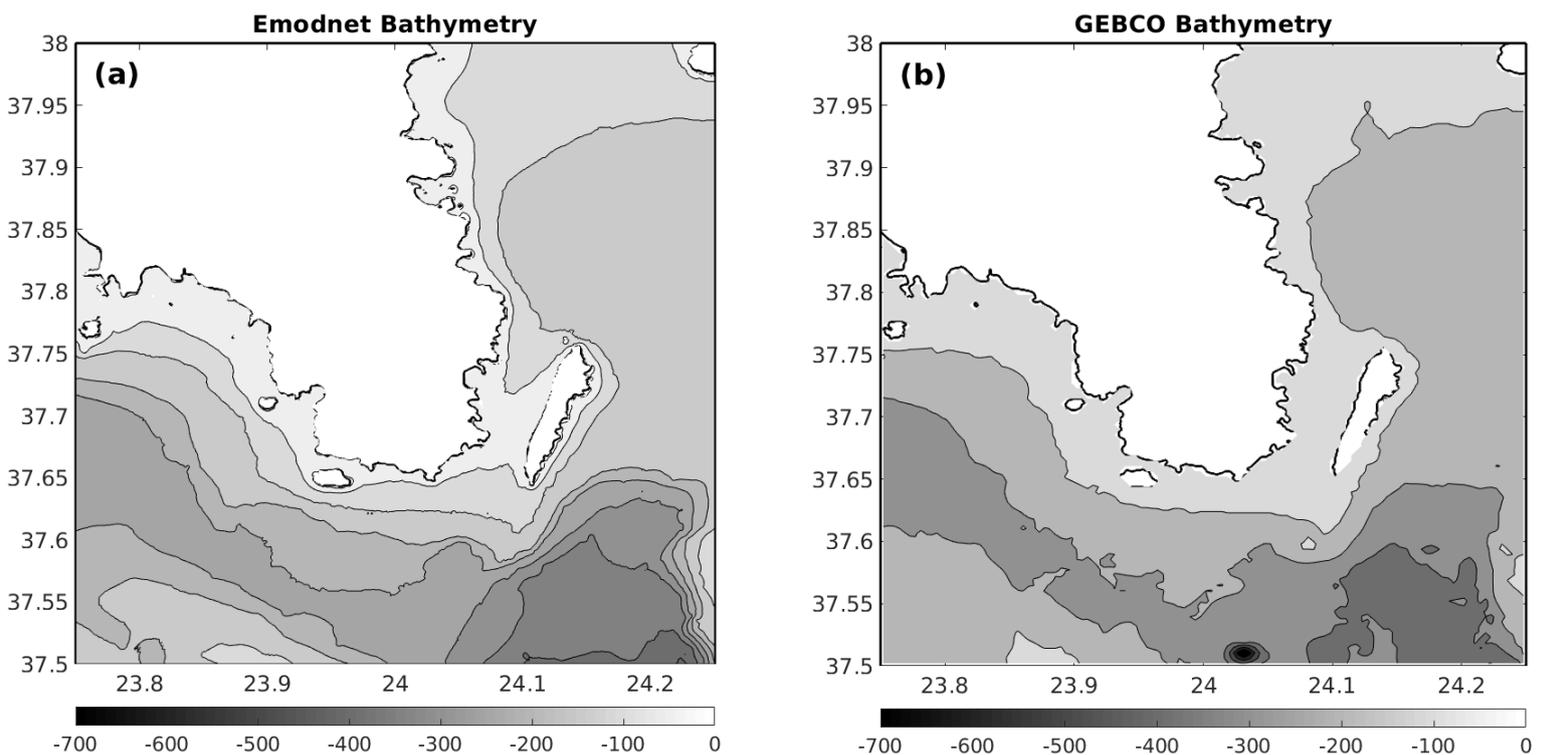


Figure 6 (a) EMODnet and (b) GEBCO bathymetries with coastline from GSHHG for the study area.

## **2.2. Climatic Data**

### **2.2.1. ERA5 Wind and Wave fields**

The second type of data input, as mentioned before at the start of chapter 2, is the boundary conditions of the computational domain, meaning the wind and wave fields. For the purposes of this thesis, the ERA5 global atmospheric reanalysis set was employed which is a product of the European Centre for Medium-Weather Forecasts (ECMWF) within the Copernicus Climate Change Service (C3S). ERA5 is replacing ERA-Interim dataset which stopped production at 2019. The advancements of ERA5 are the improvement of the horizontal resolution from 80 km (ERA-Interim) to 31 km, the hourly outputs in comparison of the three-hour outputs of ERA-Interim, and the utilization of the Integrated Forecasting System (IFS) Cy41r2 (Hersbach et al., 2020). ERA5 uses a 4D-Var assimilation scheme and produces 205 surface, 46 wave and 30 upper air parameters (Hersbach et al., 2020). It is available from 1979 to 2-3 months behind real time. The wave model of ERA5 has a resolution of 55.5 km and uses wave spectra with 24 directions and 30 frequencies. The parameters used for the present thesis simulation were 10m u and v hourly wind speed components, as well as hourly significant height of wind waves, peak wave period and mean wave direction.

### 3. Statistical analysis of offshore wind and wave data

The main parameters, wind forcing/direction and boundary conditions ( $H_s$ ,  $T_p$ ,  $\vartheta$ ), of the 40 year period 1980-2019 which were used in MIKE 21 SW in order to simulate the wave climate in the particular study area are described in a probabilistic manner by means of univariate and bivariate distributions.

The analysis included the division of the parameters range of values into bins (24 bins for  $H_s$ ,  $T_p$ , and  $\vartheta$ , 9 bins for  $u_w$ ) for the calculation of the probability density function and consequently the best fitting kernel density function. The theoretical background of the kernel density function estimation for linear and circular variables both for univariate and bivariate analysis is listed in Appendix A. More details regarding kernel density models for the description of metocean parameters are found in Athanassoulis & Belibassakis (2002).

#### 3.1. Univariate distributions of offshore wave data

The wind forcing input histogram - consisting of wind speed and direction - for the period of 1980-2019 is shown in Fig. 7. Let us note that the wind direction used is the same as the mean wave direction at offshore point [24° E, 37.5° N]. The derivation of the wind speed for the whole computational domain is described in chapter 4 (§4.3.2). As it can be seen, the most frequent directions of wind are approximately 0°-30° N-NE and 200°-220° S-SW. However, the most strong - in terms of speed - winds come from the 0°-30° N-NE direction, which can be explained very well by the annual occurrence of Etesian winds during the summer months.

The univariate distributions of offshore wave data ( $H_s$ ,  $T_p$ ,  $\vartheta$ ) for the 40-year period at offshore points O1 [24° E, 37.5° N], O2 [24.5° E, 37.5° N], O3 [24.5° E, 38° N] are presented in Fig 8-16. The wave statistics at the southernmost offshore point O1 (Fig. 8-10) show that in terms of significant wave height the highest probability is at 0-0.5 m (~67%) with ~5% probability of  $H_s$  exceeding 1.5 m, the most probable peak period lies between 4-5.5 s and the two main mean wave directions are 0°-30° (35%) and 195°-230° (23%). In a similar way, at the offshore point O2 (Fig. 11-13) the significant wave height has a 55% probability of falling into the 0-0.5 m interval with ~ 15% probability of exceeding 1.5 m, the most probable peak period is between 3-5.5 s and the two main directions from which waves come from are 0°-30° (46%) and 195°-230° (24%). Lastly, 67% of the significant wave height is between 0-0.5 m at offshore point O3 while there is a ~5% probability that it will exceed 1.5 m

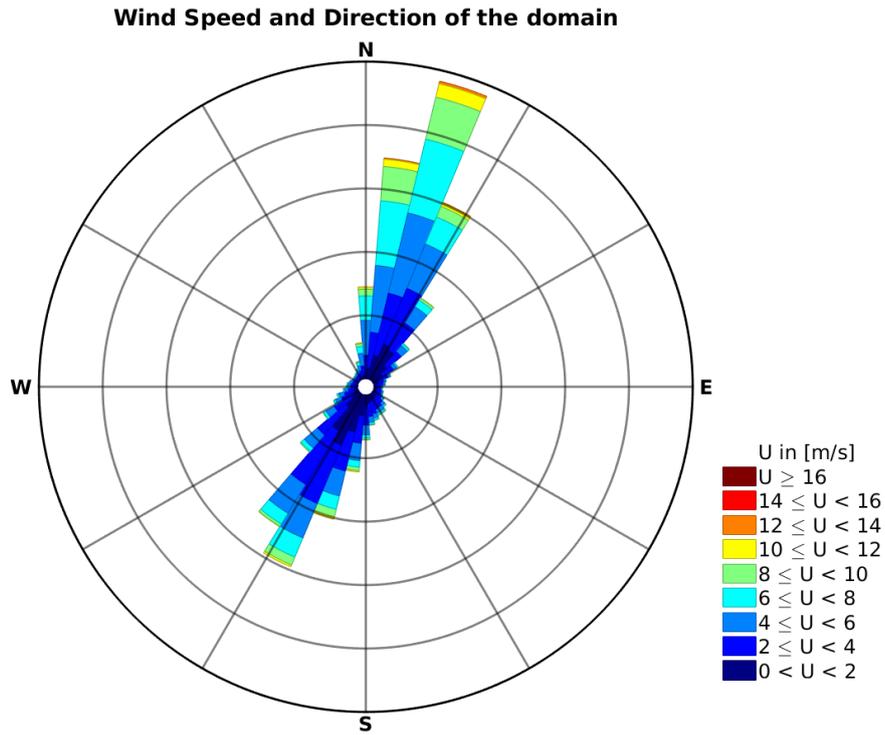


Figure 7 Histogram of wind speed and direction for the period of 1980-2019

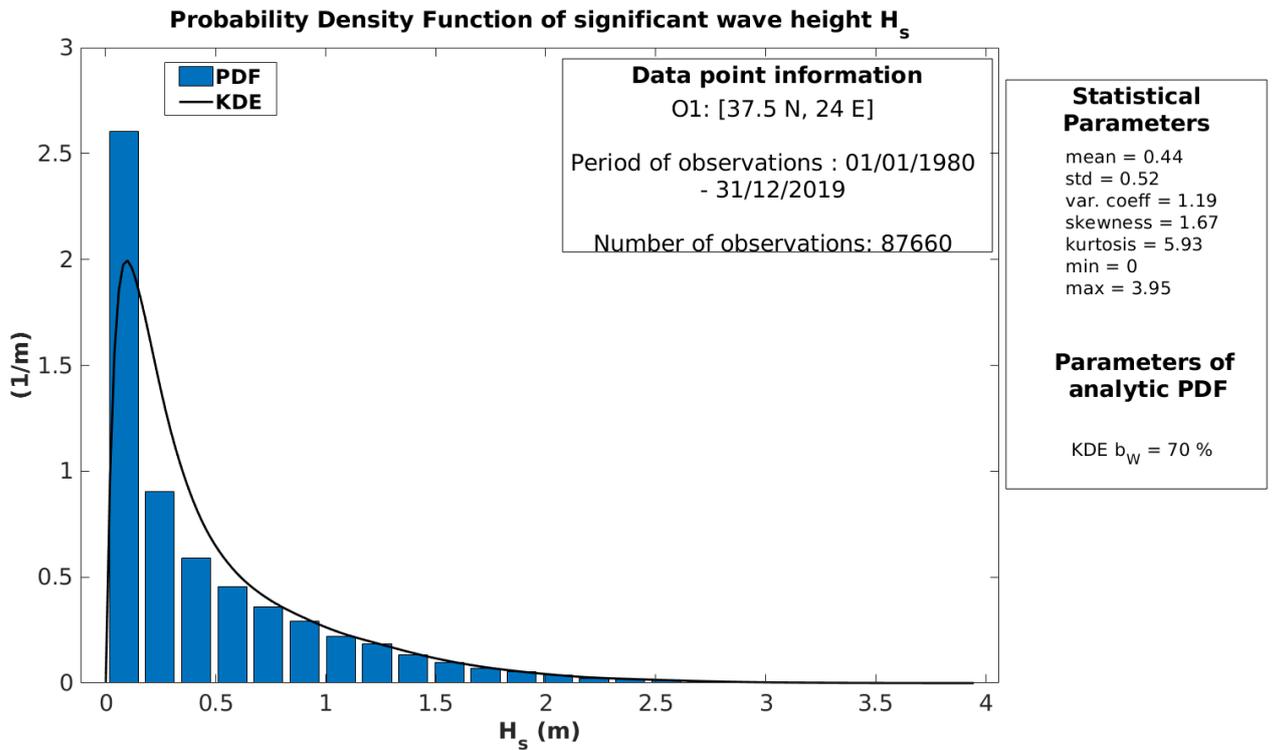


Figure 8 Probability Density Function of  $H_s$  at O1

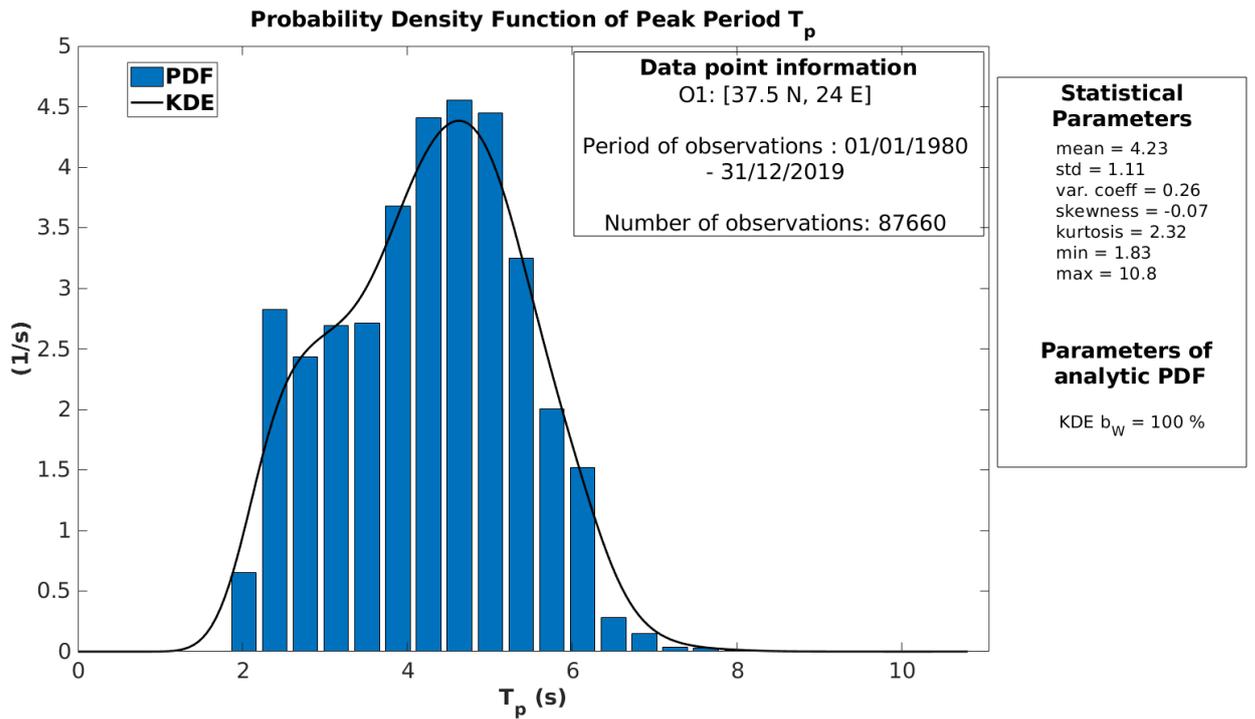


Figure 9 Probability Density Function of  $T_p$  at O1

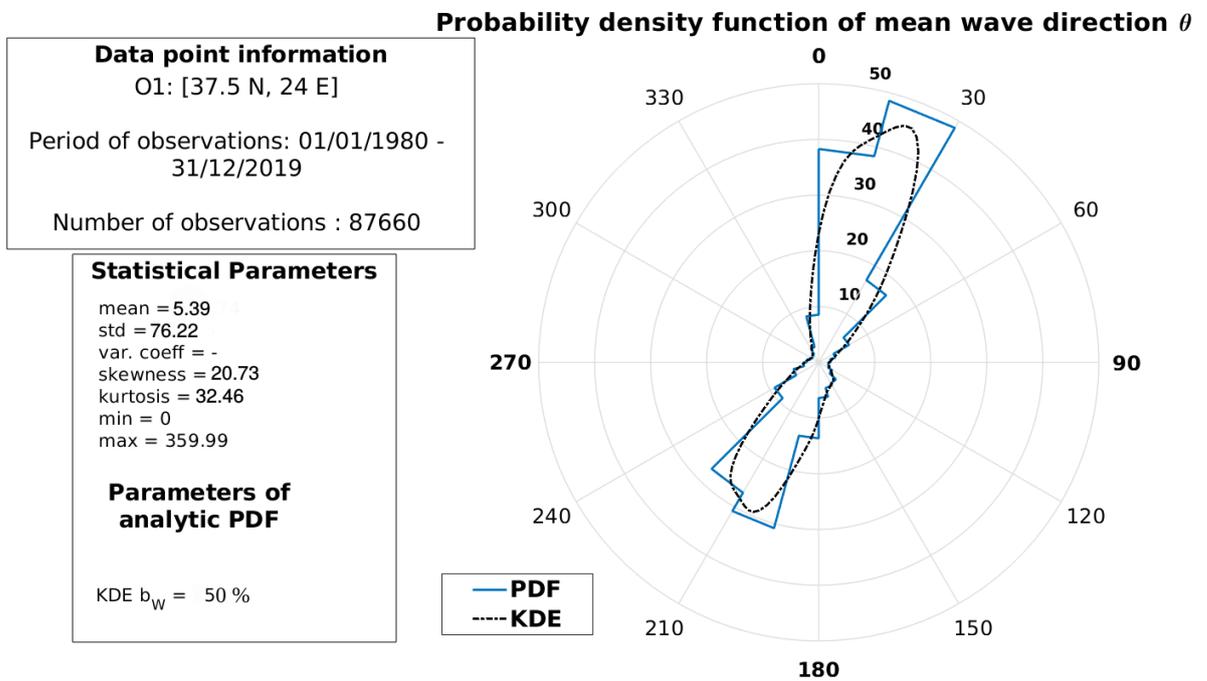


Figure 10 Probability Density Function of  $\theta$  at O1

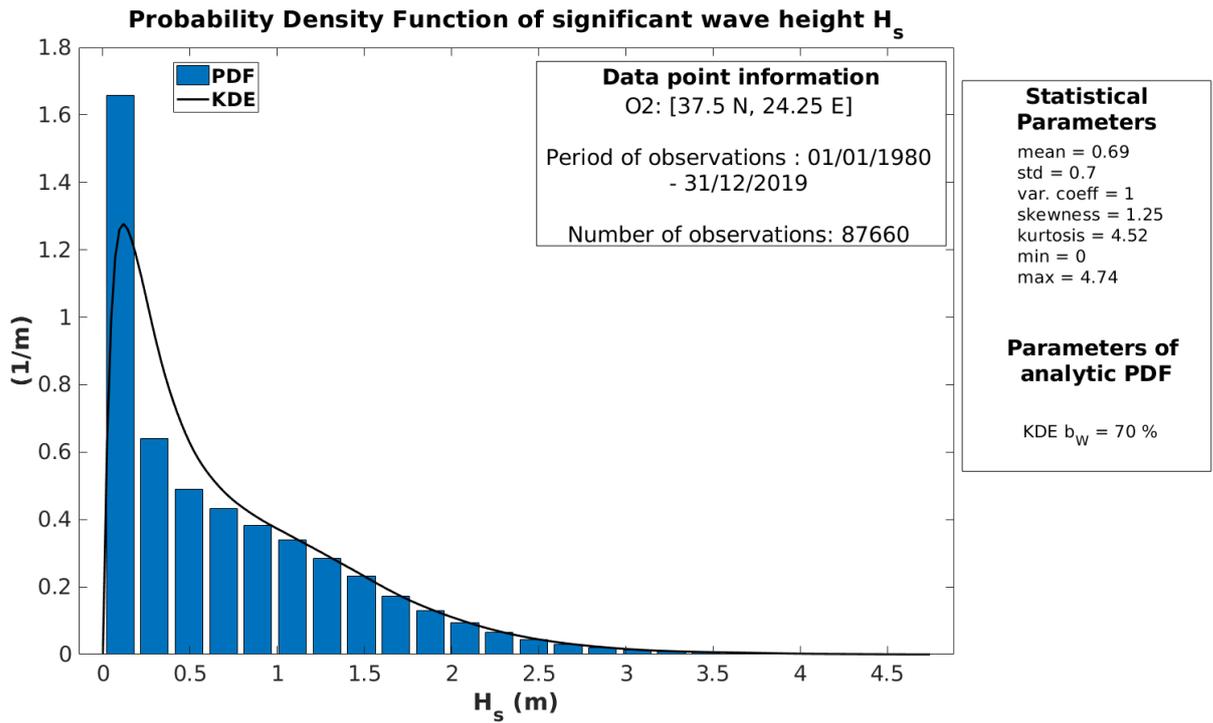


Figure 11 Probability Density Function of  $H_s$  at O2

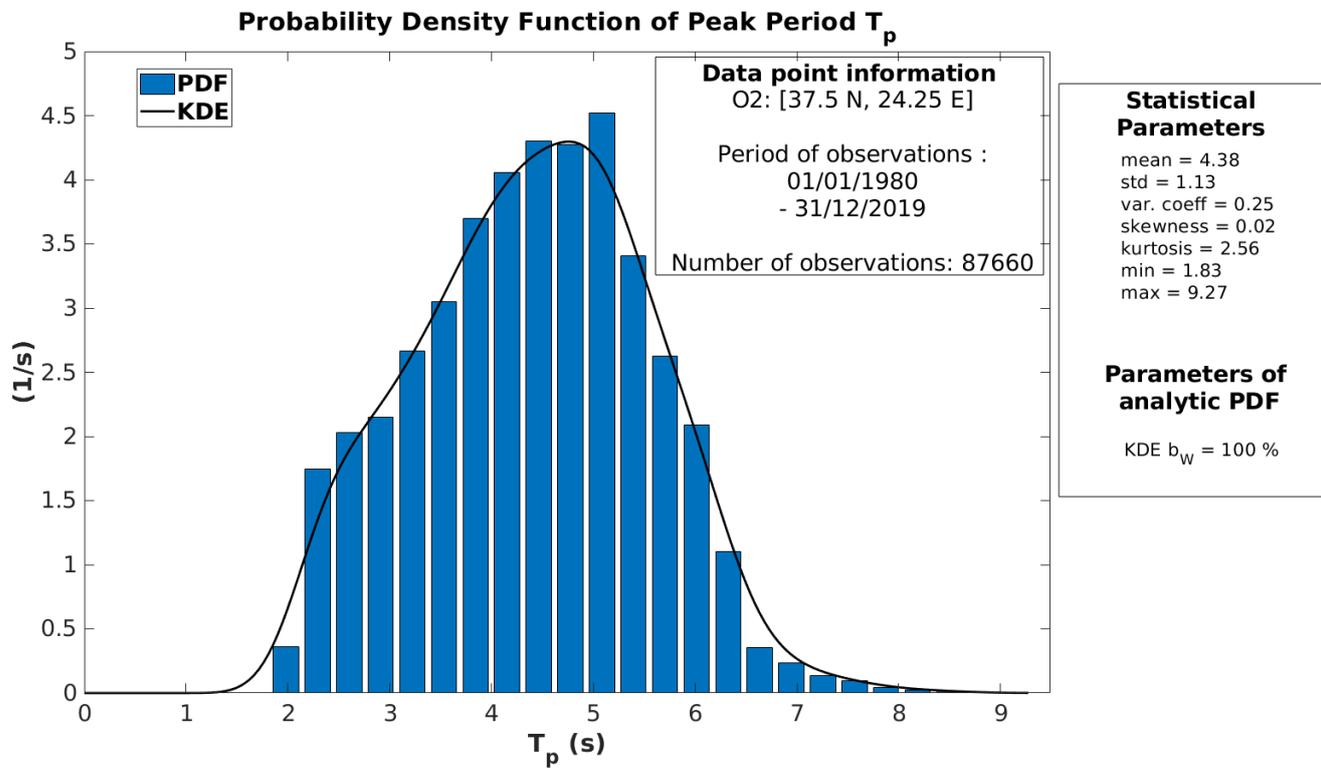


Figure 12 Probability Density Function of  $T_p$  at O2

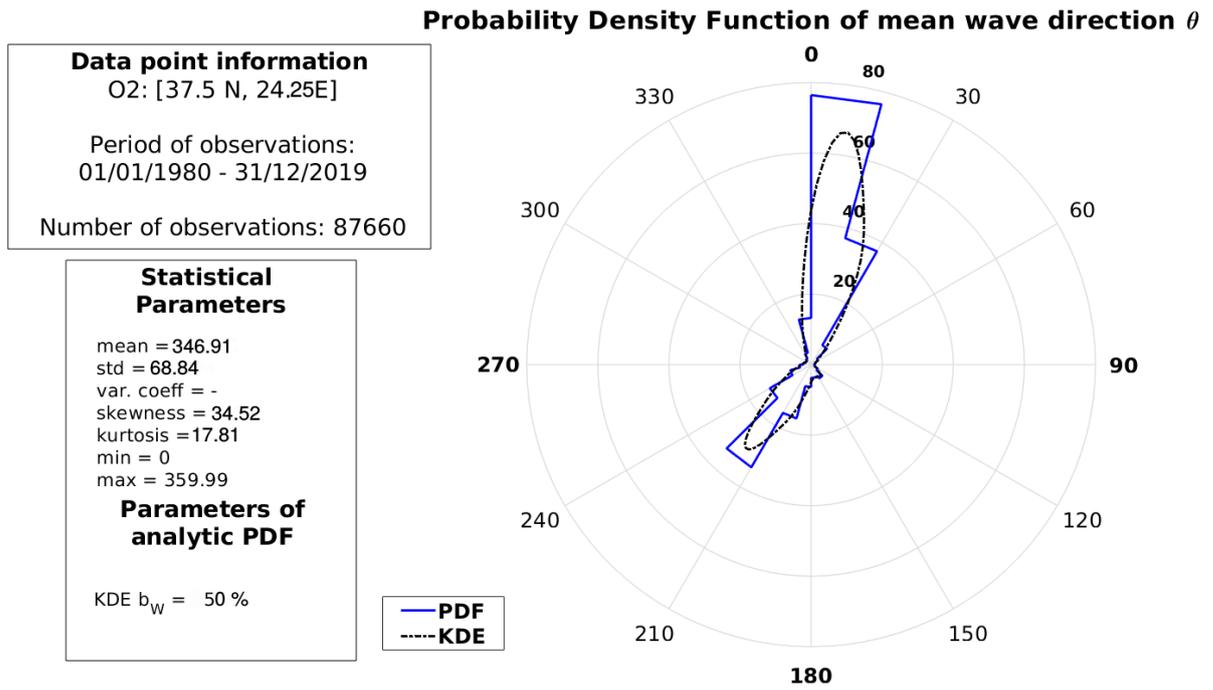


Figure 13 Probability Density Function of  $\theta$  at O2

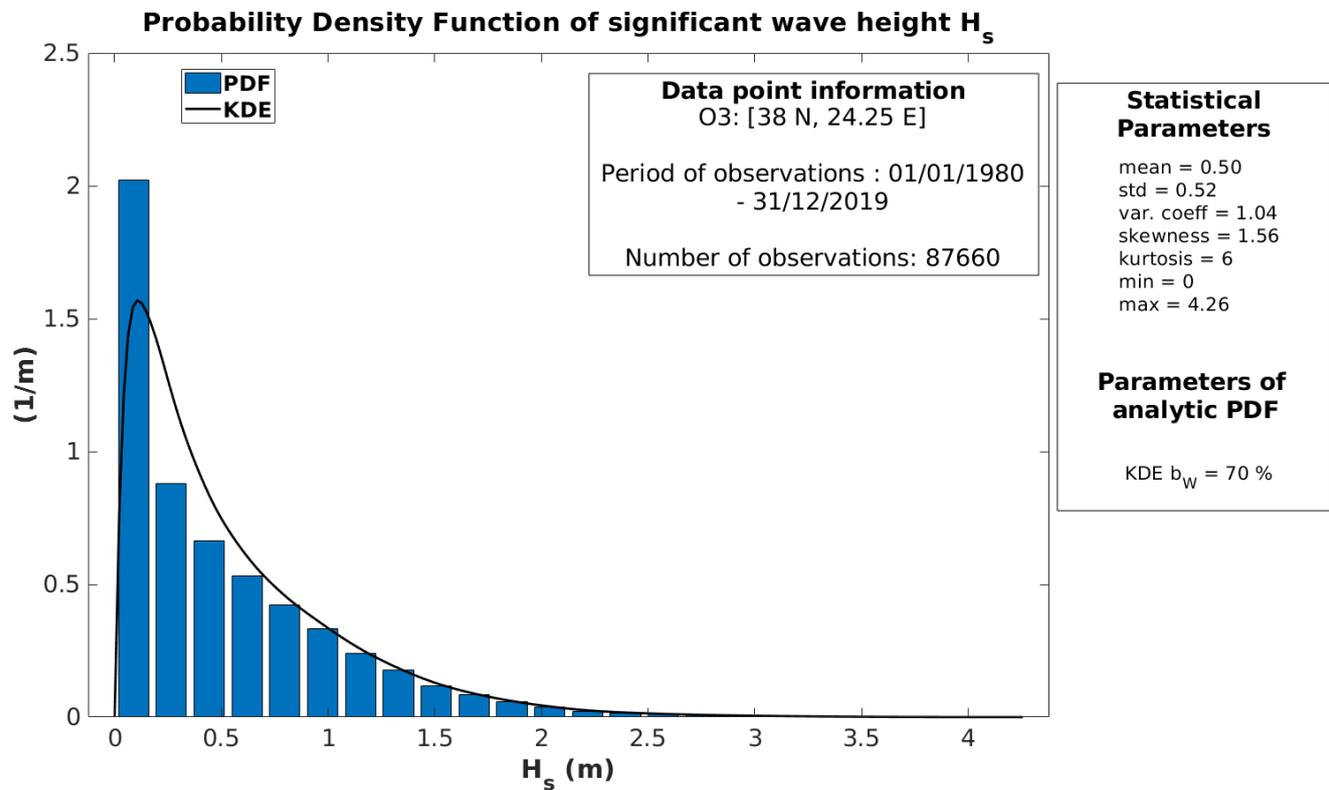


Figure 14 Probability Density Function of  $H_s$  at O3

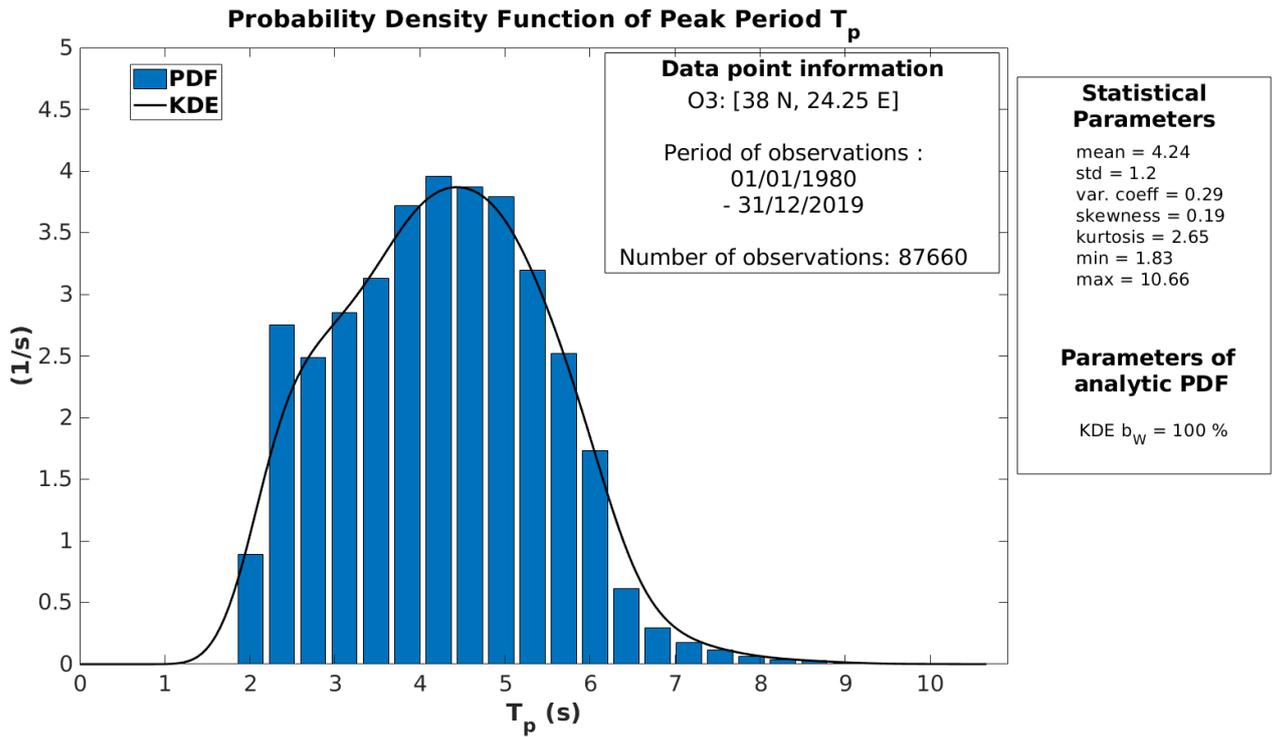


Figure 15 Probability Density Function of  $T_p$  at O3

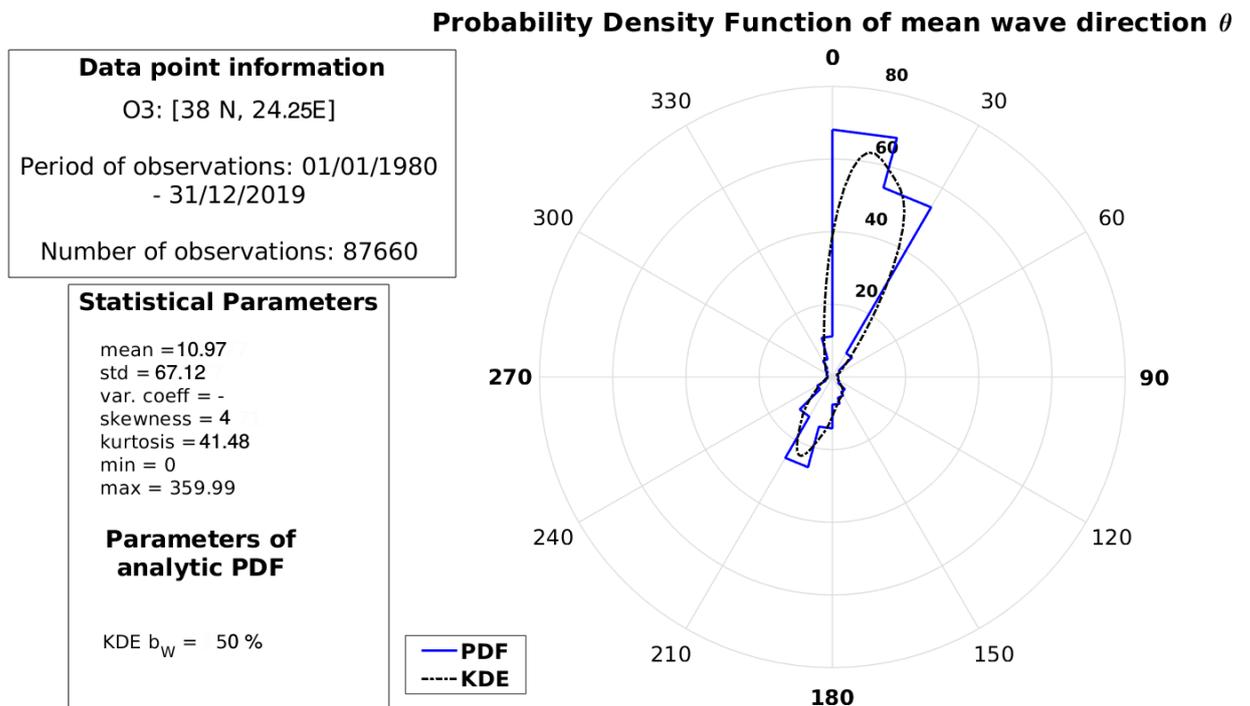


Figure 16 Probability Density Function of  $\theta$  at O3

### 3.2. Bivariate distributions of offshore wave data

As for the bivariate histograms, Fig. 17-19 represent the offshore points O1, O2 and O3. The statistical characteristics of all offshore points in this analysis do not differ substantially, at least for the significant wave height, meaning that the results show a tendency (or high probability) of small waves (less than 0.5 m  $H_s$ ) with rather short periods ( $T_p$  between 2-5 s). More specifically, at O1 (Fig. 17) the most probable combinations of  $H_s$  and  $T_p$  lie between  $\{H_s < 0.5 \text{ m}, 2 < T_p < 3 \text{ s}\}$  and  $\{H_s < 0.5 \text{ m}, 3.5 < T_p < 5 \text{ s}\}$ . Moreover, at O2 (Fig. 18) the situation is similar with O1, depicting correlation between low  $H_s$  and  $T_p$  values  $\{H_s < 0.5 \text{ m}, 2 < T_p < 5 \text{ s}\}$ . The same pattern applies for O3 (Fig. 19), in which there is a specific dense area for  $H_s < 0.2 \text{ m}$  and  $T_p \cong 2.3 \text{ s}$ .

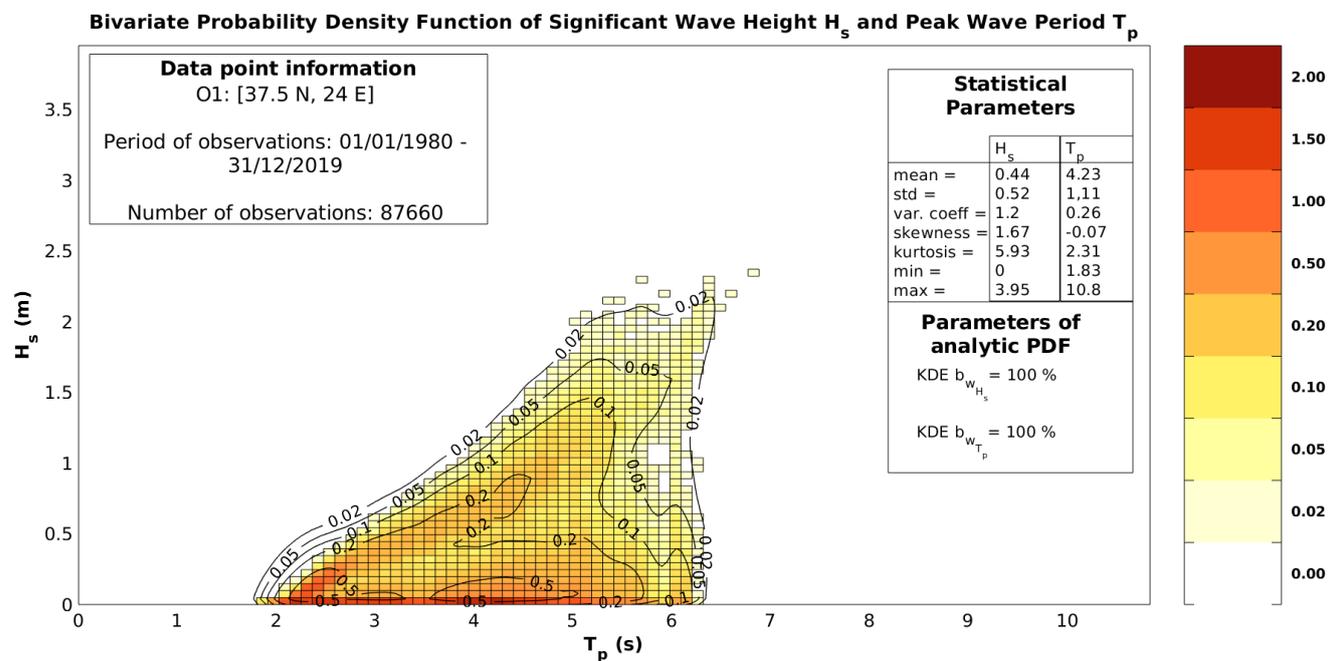


Figure 17 Bivariate distribution of  $H_s$  &  $T_p$  at O1

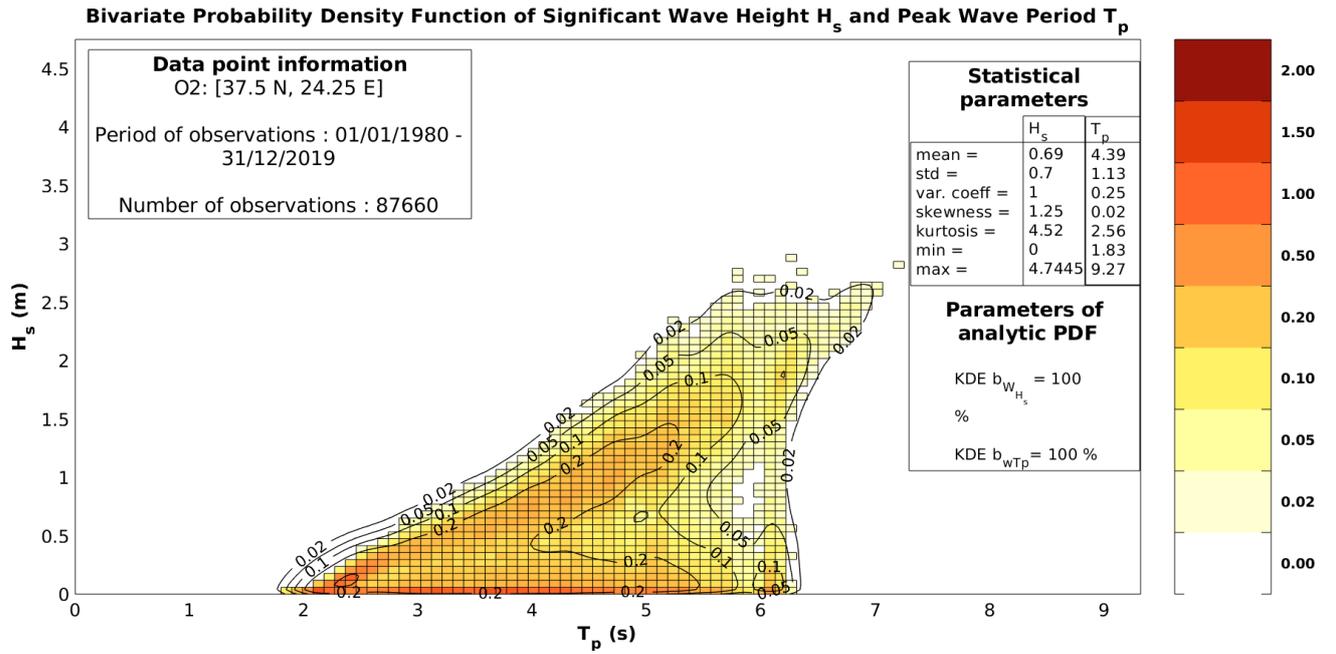


Figure 18 Bivariate Distribution of  $H_s$  &  $T_p$  at O2

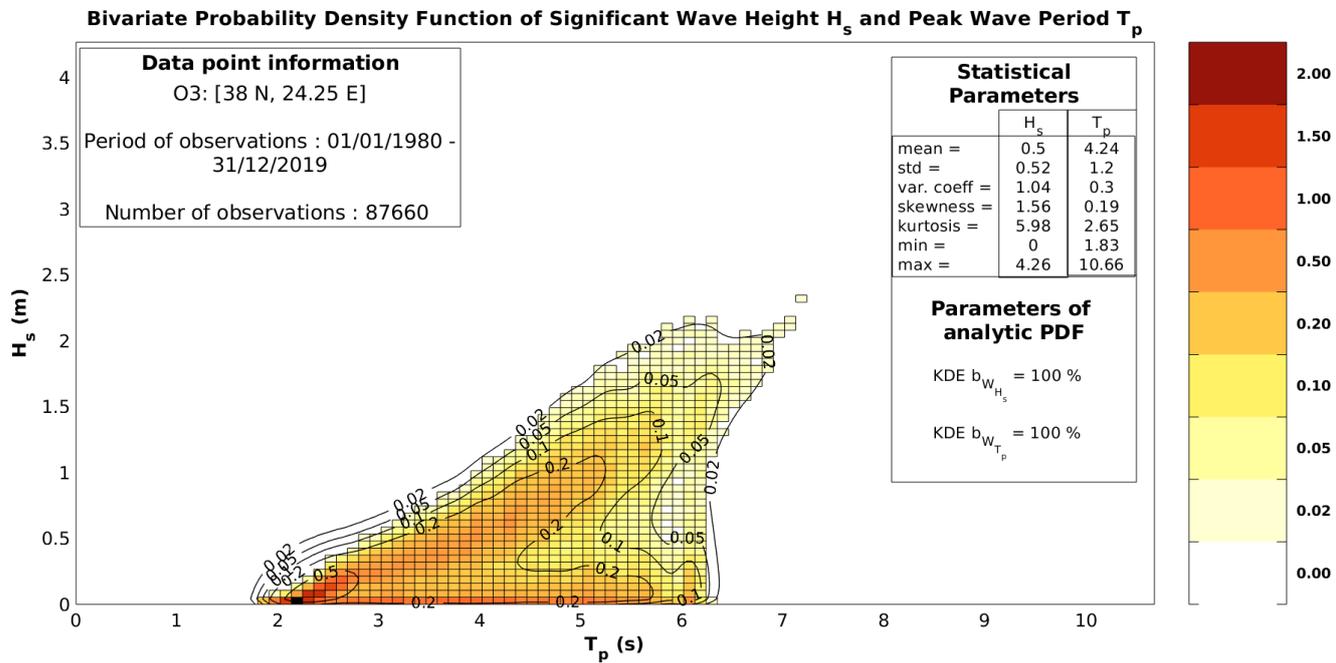


Figure 19 Bivariate Distribution of  $H_s$  &  $T_p$  at O3

## 4. Model Setup

### 4.1. About MIKE 21

The aim of this study is the modeling/assessment of wave conditions/climate of a selected region in Saronic Gulf with emphasis on the Makronisos canal and the analysis of extreme values for a given period (1980-2019). MIKE 21 Spectral Waves (SW) module was employed to achieve the first part of the thesis. MIKE 21 is a product of DHI Group. In particular, the SW module is a third generation spectral wind-wave model that simulates the growth, decay and transformation of wind-generated waves and swells in offshore and coastal areas. There are two distinct formulations of MIKE 21 SW that can be chosen, which are (a) the fully spectral formulation and (b) the directional decoupled parametric formulation. In this thesis, the fully spectral formulation was used, which is based on the wave action conservation equation (Komen et al., 1994 & Young, 1999). The reason behind this choice was the plethora of the physical phenomena it includes, which are:

- wave growth by action of wind
- non-linear wave-wave interaction
- dissipation due to white-capping
- dissipation due to bottom friction
- dissipation due to depth-induced wave breaking
- refraction and shoaling due to depth variations
- wave-current interaction
- effect of time-varying water depth and
- effect of ice coverage on the wave field.

The main equation in MIKE 21 SW is the wave action balance equation, which in horizontal Cartesian coordinates is

$$\frac{\partial N}{\partial t} + \nabla \cdot (\vec{v}N) = \frac{S}{\sigma} \quad (4)$$

$$\frac{\partial}{\partial t}N + \frac{\partial}{\partial x}c_xN + \frac{\partial}{\partial y}c_yN + \frac{\partial}{\partial \sigma}c_\sigma N + \frac{\partial}{\partial \theta}c_\theta N = \frac{S}{\sigma} \quad (4.1)$$

where  $N(\vec{x}, \sigma, \theta, t)$  is the action density,  $t$  is the time,  $\vec{x}=(x, y)$  is the Cartesian coordinates,  $\vec{v}=(c_x, c_y, c_\sigma, c_\theta)$  is the propagation velocity of a wave group in the four-dimensional phase space  $\vec{x}, \sigma$  and  $\theta$ .  $\nabla$  is the four-dimensional differential operator in the  $\vec{x}, \sigma$  and  $\theta$  space and  $S$  is the source term for energy balance equation. The characteristic propagation speeds are given by the following relationships

$$(c_x, c_y) = \frac{d\vec{x}}{dt} = \vec{c}_g + \vec{U} = \frac{1}{2} \left( 1 + \frac{2kd}{\sinh(2kd)} \right) \frac{\sigma}{k} + \vec{U} \quad (4.2)$$

$$c_\sigma = \frac{d\sigma}{dt} = \frac{\partial\sigma}{\partial d} \left[ \frac{\partial d}{\partial t} + \vec{U} \cdot \nabla_{\vec{x}} d \right] - c_g \vec{k} \cdot \frac{\partial \vec{U}}{\partial s} \quad (4.3)$$

$$c_\theta = \frac{d\theta}{dt} = -\frac{1}{k} \left[ \frac{\partial\sigma}{\partial d} \frac{\partial d}{\partial m} + \vec{k} \cdot \frac{\partial \vec{U}}{\partial m} \right] \quad (4.4)$$

Here  $s$  is the space coordinate in wave direction  $\vartheta$  and  $m$  is a coordinate perpendicular to  $s$ .  $\nabla_{\vec{x}}$  is the two-dimensional differential operator in the  $\vec{x}$  space.

Regarding the computational features, MIKE 21 SW uses cell-centered finite volume method for the discretization of the wave action equation in geographical and spectral space and a fractional step approach for time integration.

## 4.2. Computational Grid

In order to build a mesh in MIKE 21 SW, the requirements are bathymetry and coastline inputs. The region which was selected to model is defined by a box with coordinates on the bottom left corner [37.5°N, 23.75°E] and on the top right corner [38°N, 24.25°E]. The bathymetry which was chosen for this study is from EMODnet, specifically the 2018 DTM version with an approximately 91 x 91 m resolution.

Initially the bathymetry was downloaded from <https://www.emodnet-bathymetry.eu> as a .dtm file, which contained a greater region than the one required for the study. For this reason, the .dtm file was resized using CDO operators and then manipulated through Matlab numerical computing environment in order to be written in the proper form as a .xyz file. The bathymetry of the region can be viewed in Fig. 20.

Regarding the coastline, a map from the Hellenic Navy Hydrographic Service (HNHS) was used in order to digitize the desired length of the study area coastline in MIKE environment. The HNHS's map scale was 1:100.000, with a Mercator projection. The sources of the data on the map were from Hellenic surveys up to 1985. Both the bathymetry and coastline files were first created in accordance with World Geodetic System (WGS) but then converted to Universal Transverse Mercator (UTM) 34 coordinate system, using the datum convert feature of MIKE Zero toolbox.

The goal in this part of the study was to produce accurate wave results in 5 points between Attica and Makronisos island. Therefore, it was best suited to create a variable resolution unstructured mesh, highlighting Makronisos channel while having a standard resolution for the rest of the computational area. The mesh created in MIKE can be viewed in Fig. 21 without the bathymetry and Fig. 22 with the interpolated bathymetry.

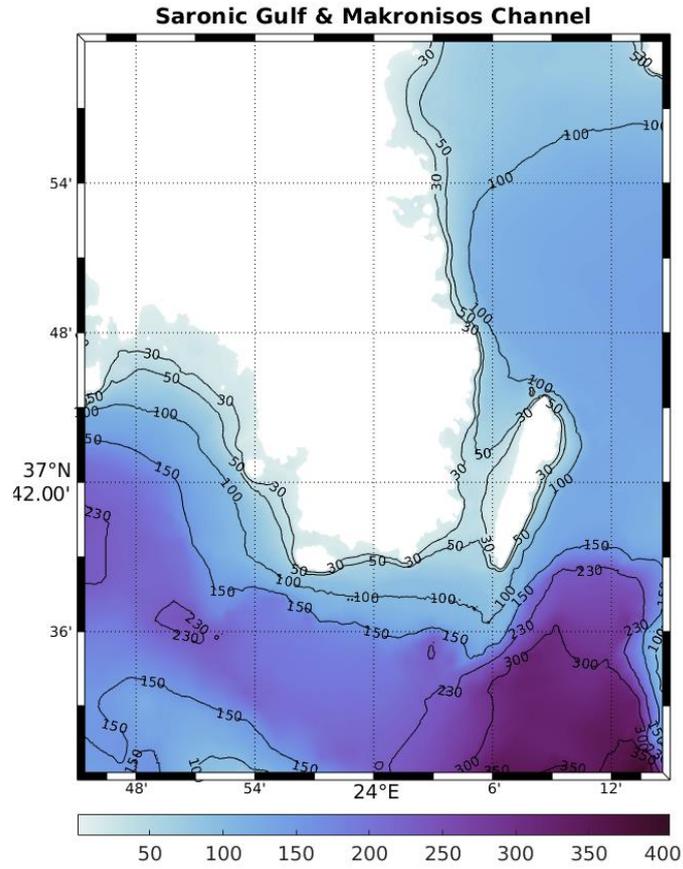


Figure 20 Study area bathymetry from EMODnet

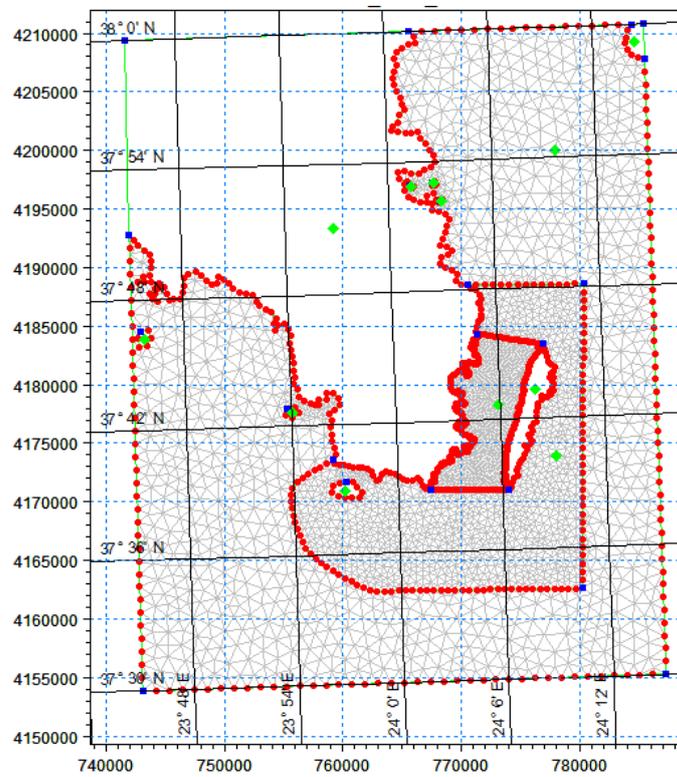


Figure 21 Computational grid without bathymetry

As seen in Fig. 21-22, the domain is split in three different regions. The outer region, named Level 1, has 1.000.000 elements, the intermediate region, Level 2, has 250.000 elements and lastly the region between Attica and Makronisos, Level 3, has 75.000 elements.

In accordance with the desired analysis of each Level mentioned above, the respective boundaries (including the coastline) of the regions had to be edited in terms of node spacing. Thus, as shown in Fig. 21, the boundaries of the whole study area have scarcer nodes in comparison with the boundaries of the finer region of Makronisos channel.

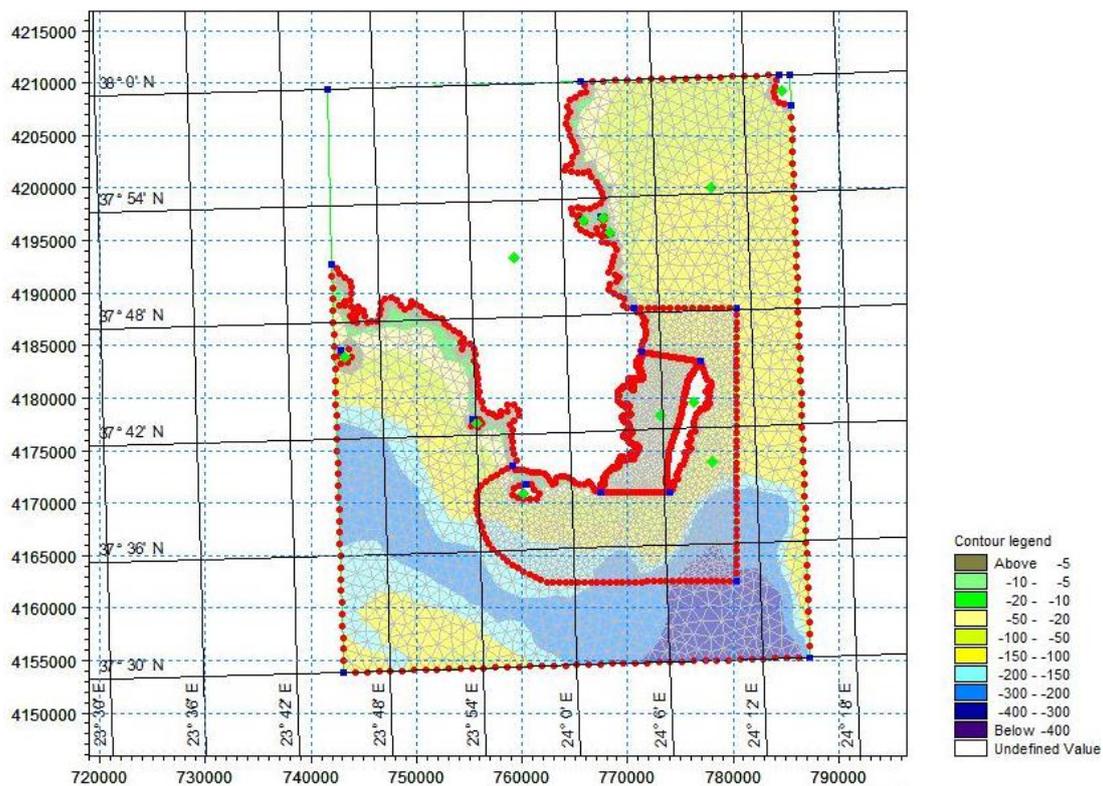


Figure 22 Computational grid with interpolated bathymetry

### 4.3. Spectral Wave module setup

The simulation of both the stationary runs and the two year period wave hindcast was completed with MIKE 21 SW. Except wind forcing, boundary conditions and the output file (area or point) all runs had the same features regarding the model setup.

As mentioned before, fully spectral formulation was chosen instead of directional decoupled parametric, and as for time formulation the option was instationary. The type of spectral discretization was logarithmic with 25 frequencies (minimum 0.055 Hz) and frequency factor 1.1. Moreover, the type of the directional discretization

was 360 degree rose with 16 distinct directions. The solution technique regarding geographical space discretization was low order/fast algorithm, with maximum 32 levels in transport calculation, 0.01 sec min and 30 sex max time steps. The energy transfer included quadruplet-wave interaction.

Wave breaking was described by specified gamma with the constant value of 0.8, bottom friction was expressed via Nikuradse roughness (Kn) using 0.04 m constant value, white capping with 4.5 constant value and dissipation coefficient was set to 0.5. Both runs initialized with zero spectra. The simulations did not include the separation of wind sea and swell, diffraction, water level conditions, current conditions, ice coverage and structures.

### 4.3.1. Stationary Runs

In order to validate the efficiency of the generated mesh, several stationary runs (with constant wind forcing and boundary conditions) were performed using indicative values for the input parameters (significant wave height  $H_s$ , peak wave period  $T_p$ , mean wave direction  $\vartheta$ , wind speed  $u_w$  and wind direction  $\vartheta_w$  as seen in Table 1) from Medatlas Climatology (Wind and Wave Atlas of the Mediterranean Sea, 2004). This particular climatology is the product of Medatlas project which took place between 1999-2004, aiming for the production of long term wind and wave statistics at selected points of the Mediterranean Sea. The time series were created by the statistical analysis of wind and wave data of a ten year period. The origin of the data analyzed was from visual observations from ships, buoys and/or platforms (Spain, Cyprus, Northern Crete, Italian coasts), remote sensing (ERS1, ERS2, TOPEX/Poseidon) and met-ocean operational models (ECMWF WAM). The model was set in order to produce results for the whole area.

	$H_s$ (m)	$T_p$ (s)	$\Theta$ (°)	$u_w$ (m/s)	$\theta_w$ (°)
1	0.6	4.75	15°	3.7	15°
2	0.6	4.75	230°	3.7	230°
3	0.4	4.5	15°	2.5	15°
4	0.4	4.5	230°	2.5	230°
5	0.5	10	15°	3.1	15°
6	0.5	10	230°	3.1	230°
7	2	7	15°	12.3	15°
8	2	7	230°	12.3	230°

Table 1 Medatlas Climatology

As the results suggest (seen in Fig. 23-27), the mesh did correspond particularly well with the given input parameters. In every case where the wind direction was from 15° N, significant wave height magnitude is higher in the northeast and lower in southwest. The opposite, as expected, happened with wind direction from 230° S-

SW, with the highest  $H_s$  coming from southwest. Mean wave direction follows the trend of wind direction and coincides with  $H_s$  vectors.

More in depth, starting from the first stationary run, the highest waves ( $> 0.6$  m) appear on the north and east boundary (Fig. 23). Since the wind blows from  $15^\circ$  N and its magnitude is quite low, the area southwest of Attica is governed by rather low waves, a fact which could be possibly explained because of shading. The peak period appears to have the same magnitude as the input period ( $\sim 4.75$  s) and the only changes are observed in the Sounio nearshore region where the peak periods magnitude drops significantly and a very small region close to Thoriko where the period exceeds 6 s (Fig. 24). Concerning the mean wave direction, in the north and east area of the domain it coincides with the direction of the wind but on the west-southwest region it changes to east-northeast ( $300^\circ$ - $360^\circ$ ) (Fig. 25).

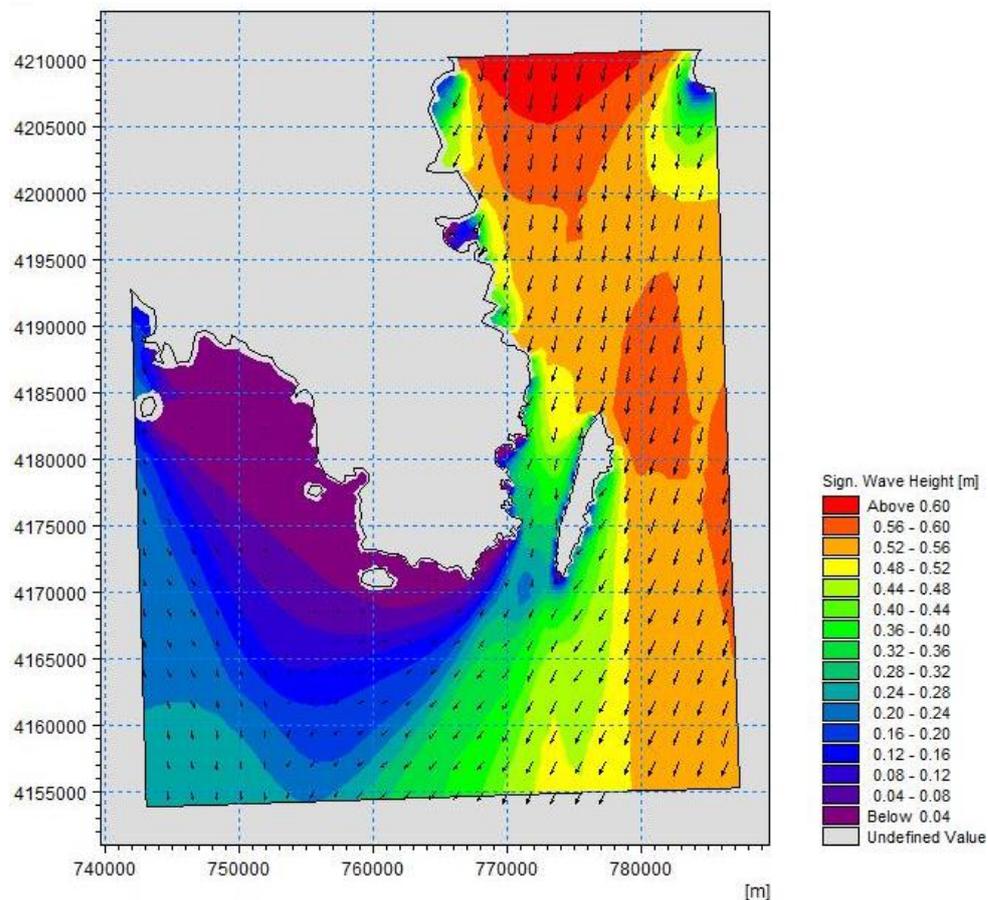


Figure 23 1st stationary run result for  $H_s$

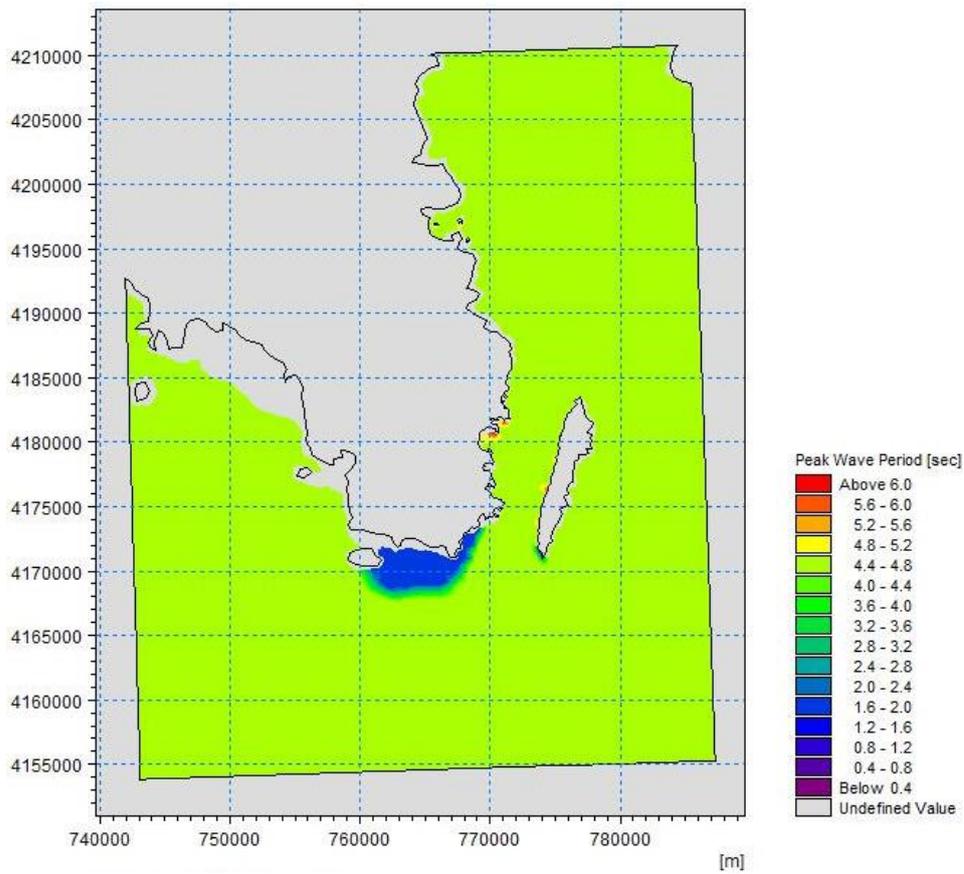


Figure 24 1st stationary run result for  $T_p$

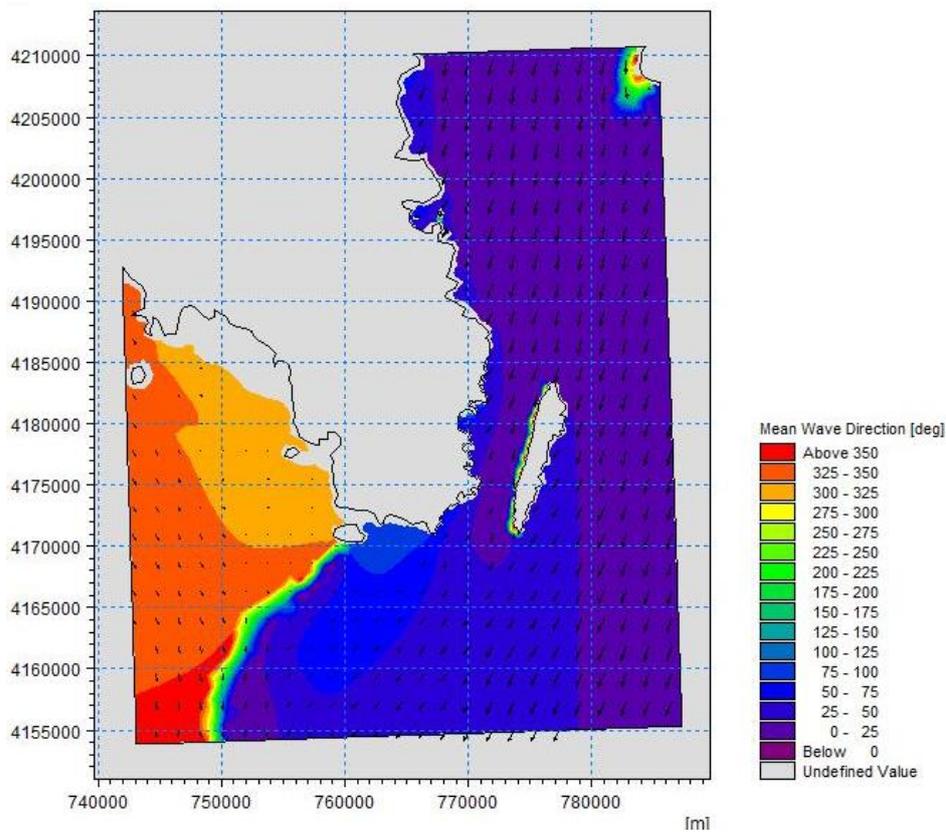


Figure 25 1st stationary run result for  $\theta$

On the second run, the only parameters that changed were the mean wave direction  $\vartheta$  and wind direction  $\vartheta_w$  from  $15^\circ$  N to  $230^\circ$  S-SW. As expected, the significant wave height was highest on the west south and southeast boundaries ( $> 0.6$  m) and lower on the north-northeast boundaries (Fig. 26). Peak period resulted the same as the input period in the whole domain (Fig. 27) and mean wave direction coincided with the wind direction for the most part of the area (Fig. 28).

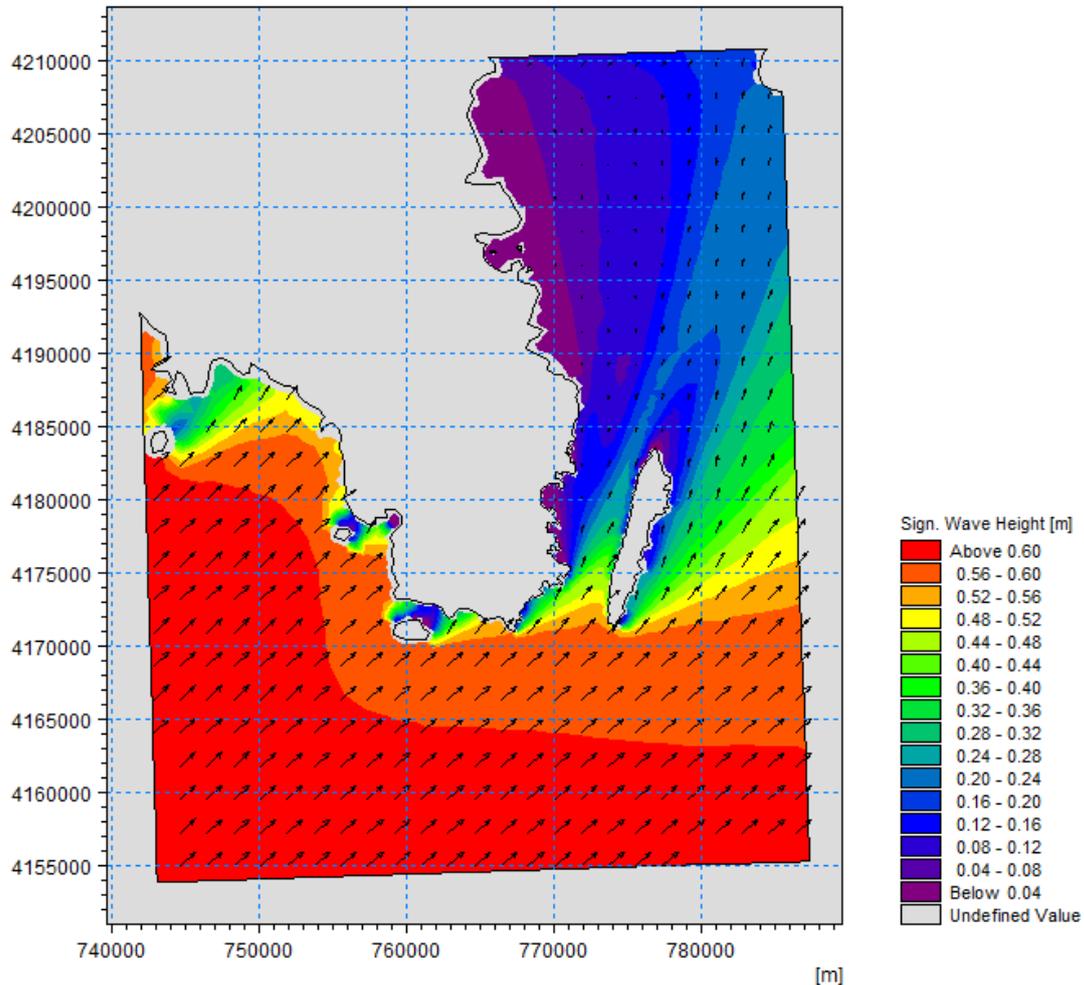


Figure 26 2nd stationary run result for Hs

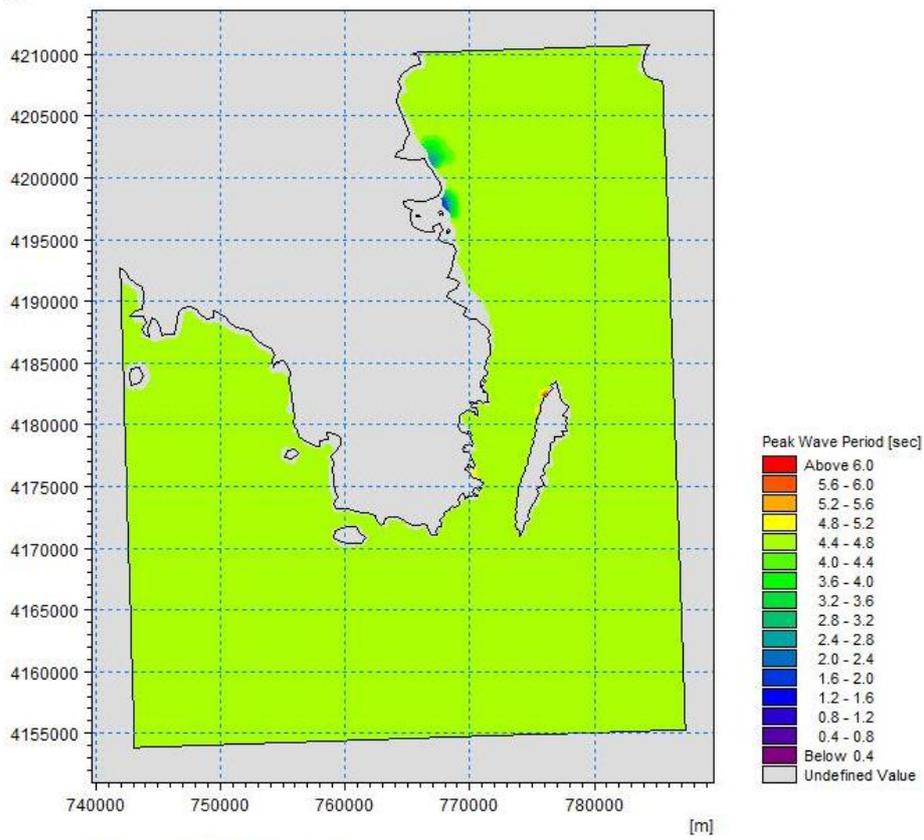


Figure 27 2nd stationary run result for  $T_p$

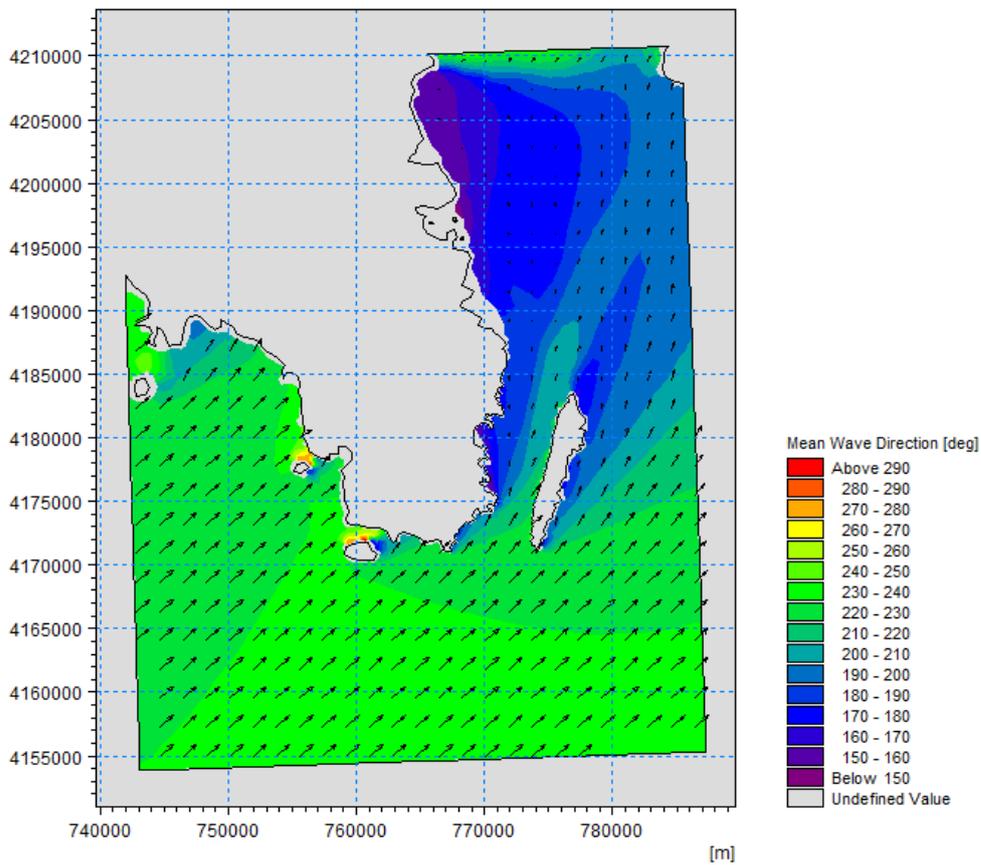


Figure 28 2nd stationary run result for  $\theta$

The third stationary run differed on the magnitude of the wave and wind parameters (they were slightly lower as shown in Table 1), and as a consequence the results for  $H_s$ ,  $T_p$  and  $\vartheta$  (Fig. 29-31) illustrate the same characteristics as the first run but with lower parameter values.

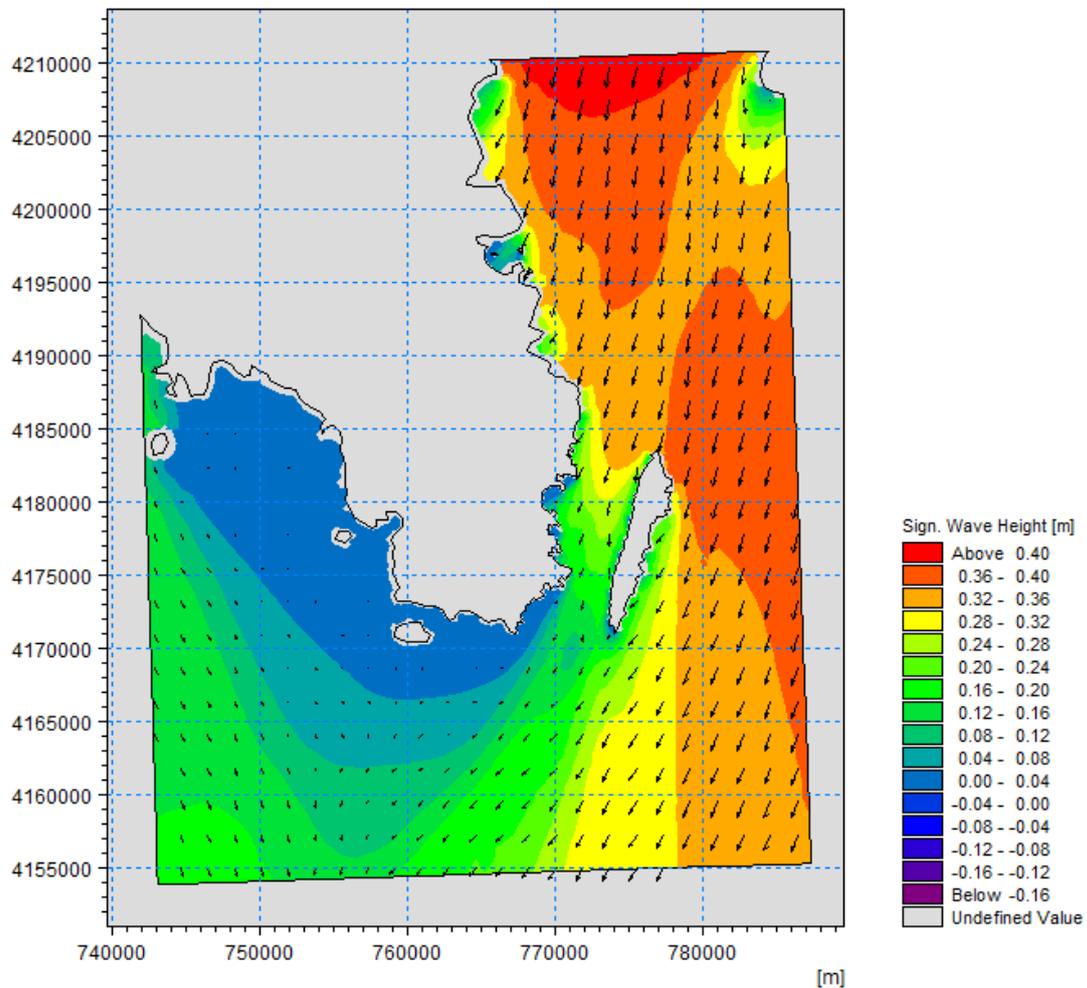


Figure 29 3rd stationary run result for  $H_s$

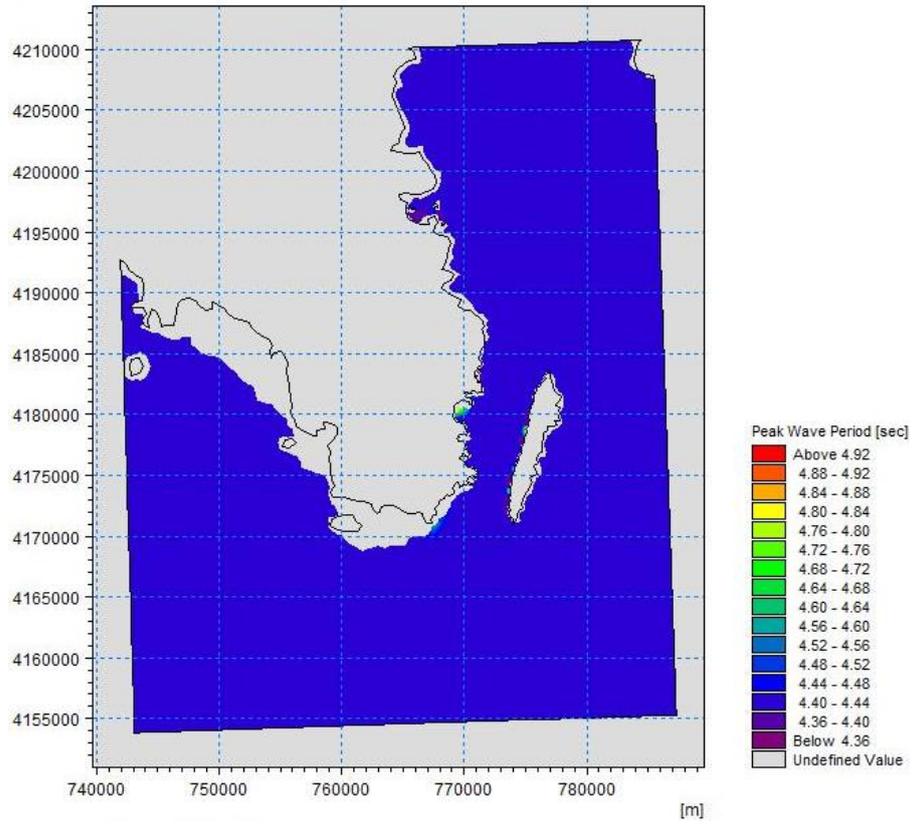


Figure 30 3rd stationary run for  $T_p$

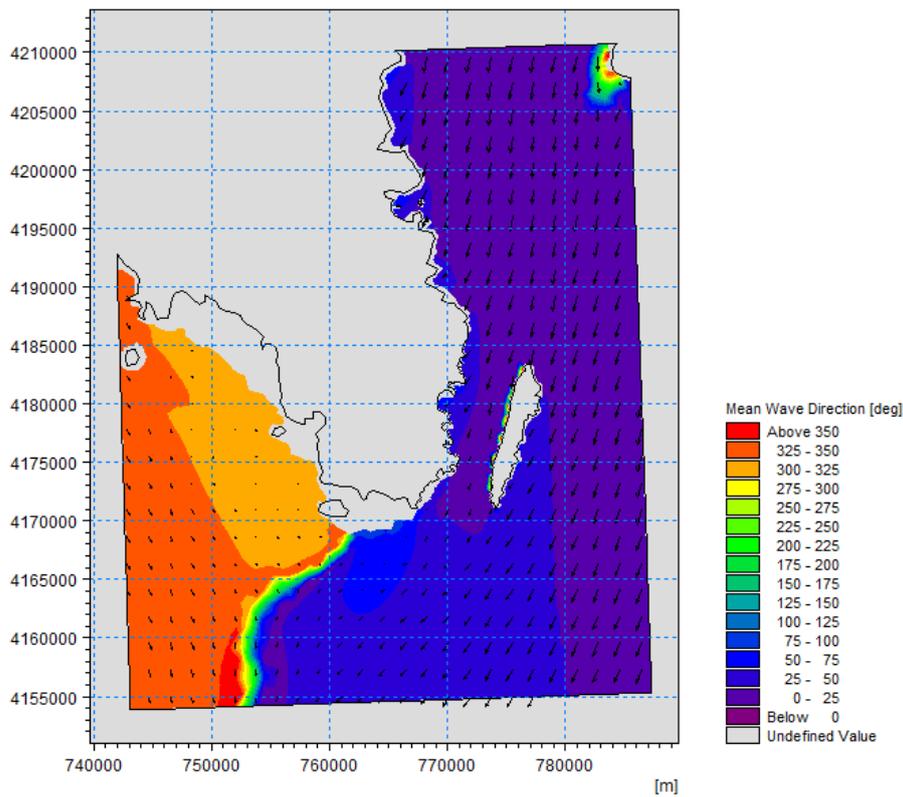


Figure 31 3rd stationary run for  $\theta$

To continue, the fourth run exhibited the same pattern as the second run, with the highest significant wave height being on the west south and southeast boundaries (Fig. 32) and a very clear effect of shading behind three islets near Vouliagmeni, Anavissos and Sounio, as well as the eastern part of Attica and Makronisos. Peak period results the same as the input period (Fig. 33) and mean wave direction follows the direction of the wind (Fig. 34).

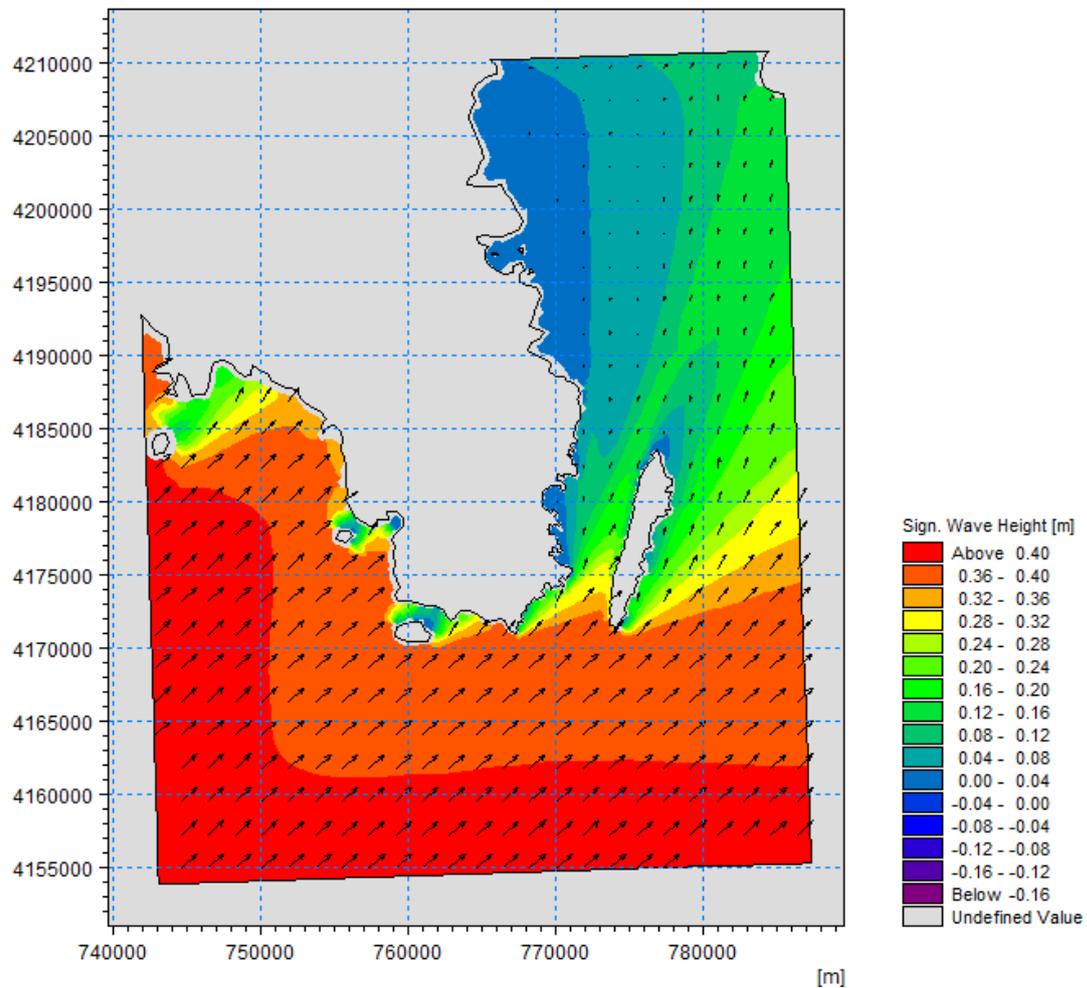


Figure 32 4th stationary run result for Hs

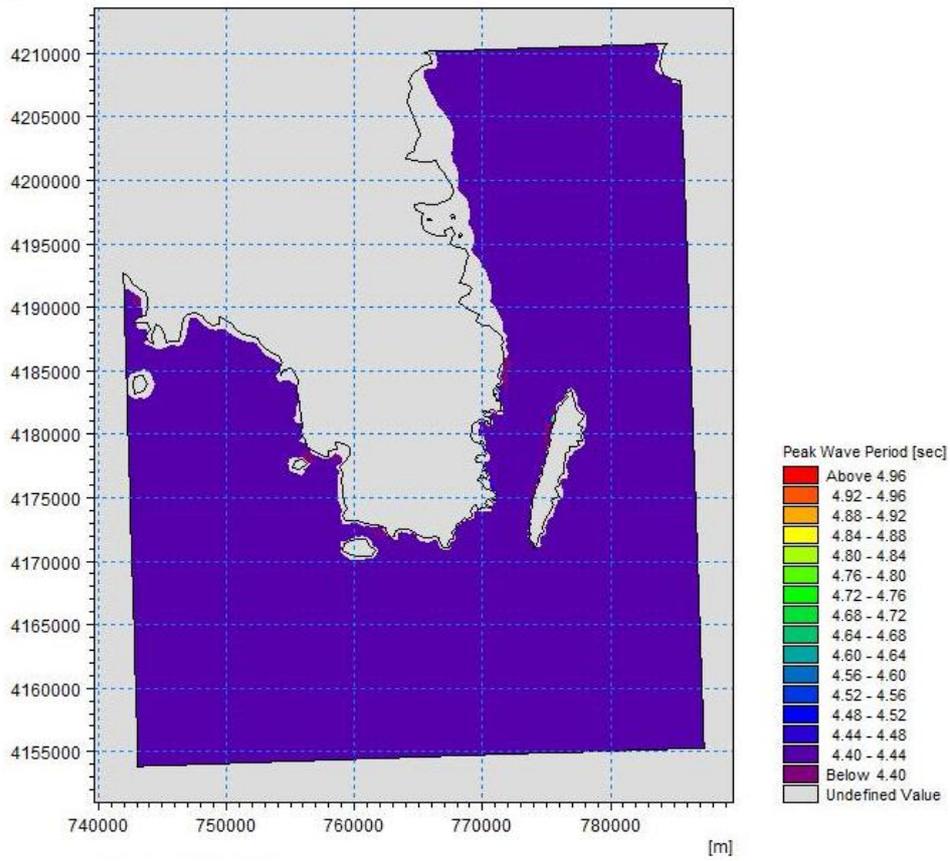


Figure 33 4th stationary run for  $T_p$

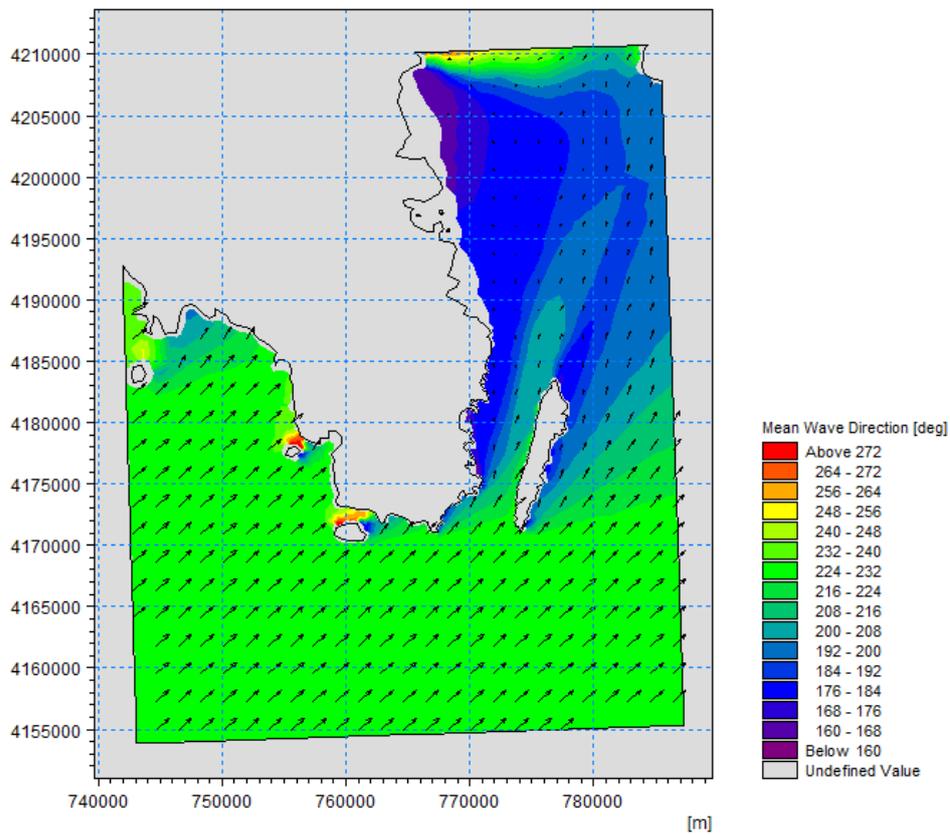


Figure 34 4th stationary run for  $\theta$

As for the fifth stationary run, the main difference from the previous runs was the higher peak period input. As the figures illustrate (Fig. 35-37), the results show great similarity with the first and third stationary runs concerning all the wave parameters. Although the peak period in this case (Fig. 36) has the same magnitude as the input period for most of the domain, it shows values  $> 10$  s in nearshore areas, *e.g.* from Vouliagmeni to Saronida and south of Sounio. Furthermore, it has a lower value on the southern entrance of Makronisos island ( $\sim 9.7$  s). Mean wave direction (Fig. 37) also follows the same pattern as in first and third runs, however in the west-southwest boundary there is a gradient of change which could be due to the gradient of the bathymetry in that particular area.

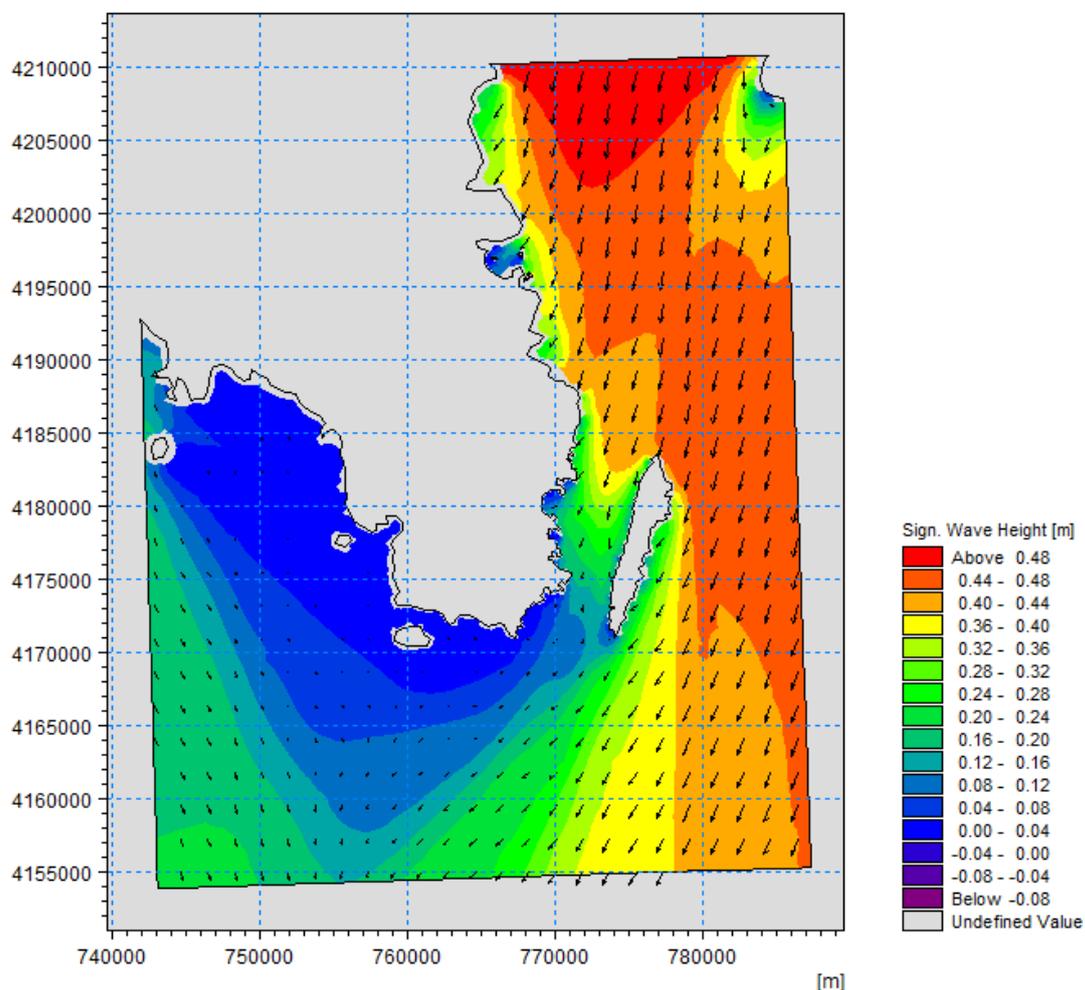


Figure 35 5th stationary run result for Hs

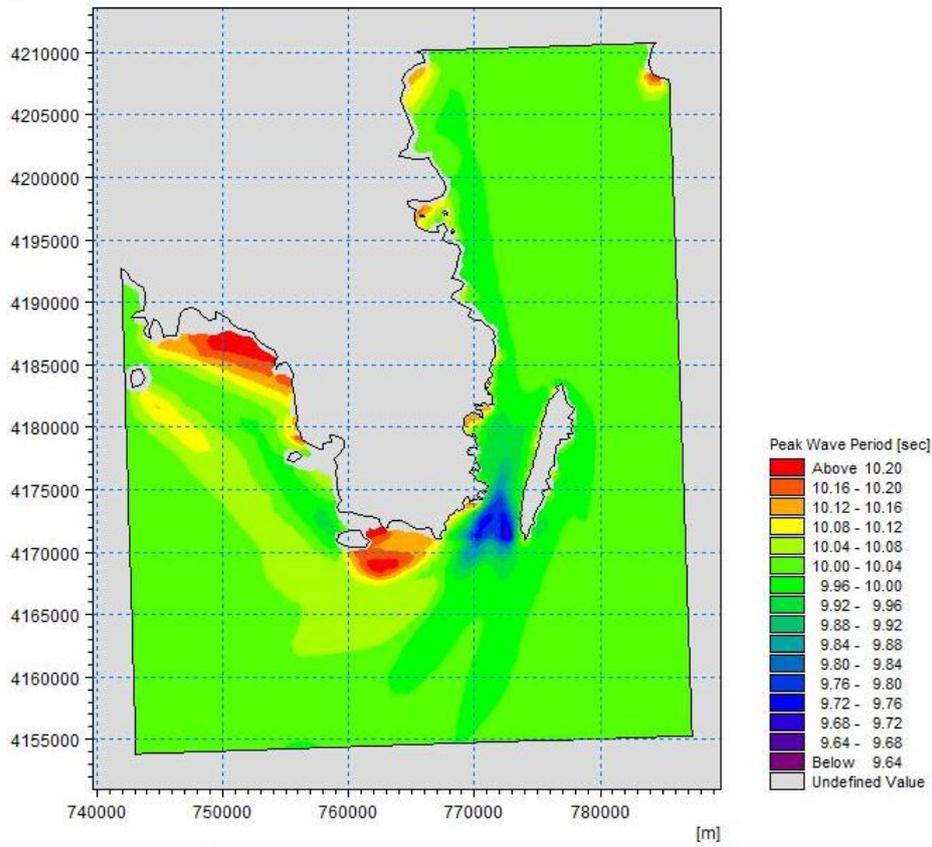


Figure 36 5th stationary run result for  $T_p$

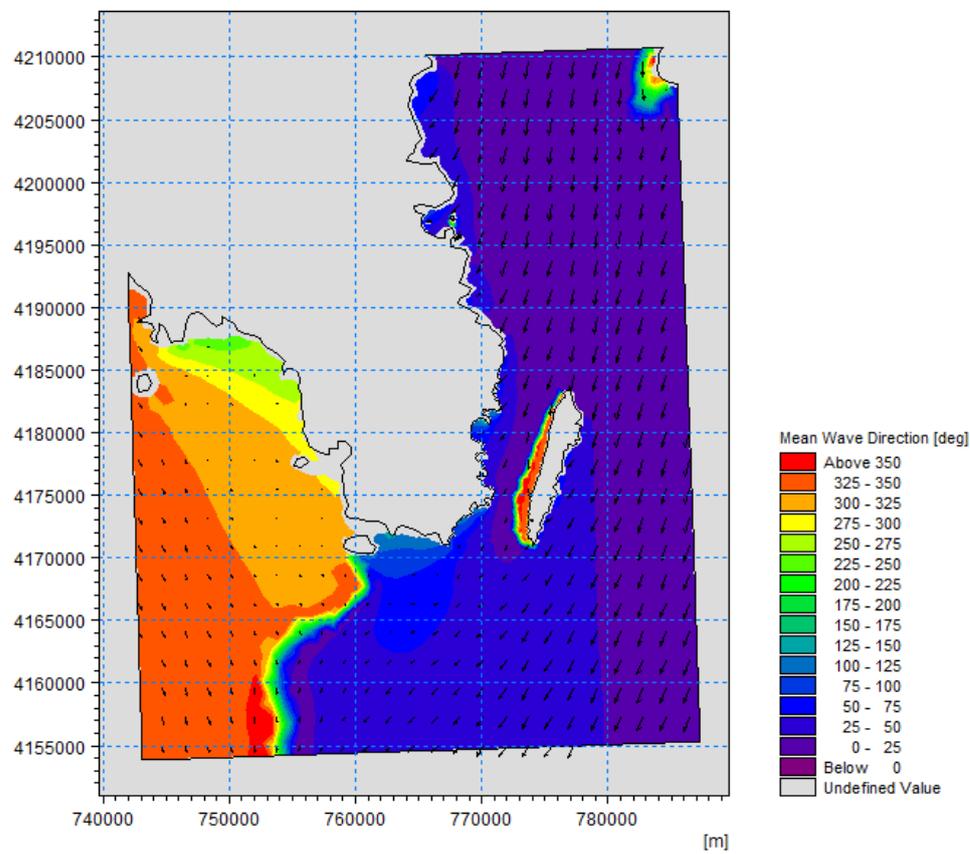


Figure 37 5th stationary run result for  $\theta$

The sixth stationary run had the same magnitude of  $H_s$ ,  $T_p$  and  $u_w$  but the wind direction  $\vartheta_w$  and the mean wave direction  $\vartheta$  were from  $230^\circ$ . Here the results show great similarity with all the previous results with the same wind direction regarding  $H_s$ , which is the development of higher waves on the west-south boundaries (Fig. 38). Peak period is uniform on the domain except on the upper entrance of Makronisos channel where its magnitude drops by approximately 0.5 s (Fig. 39). Mean wave direction is the same as the wind direction but changes on the north-northeast boundary (Fig. 40).

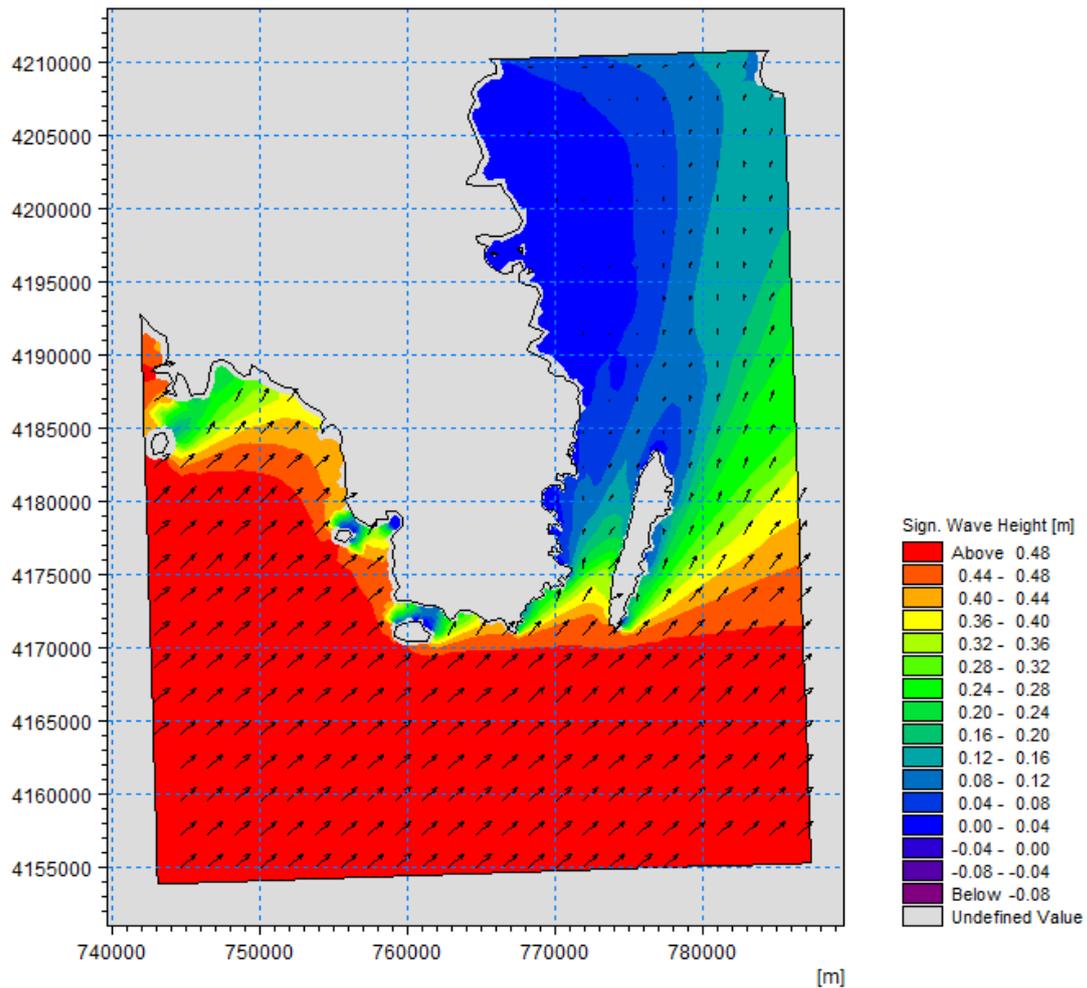


Figure 38 6th stationary run result for  $H_s$

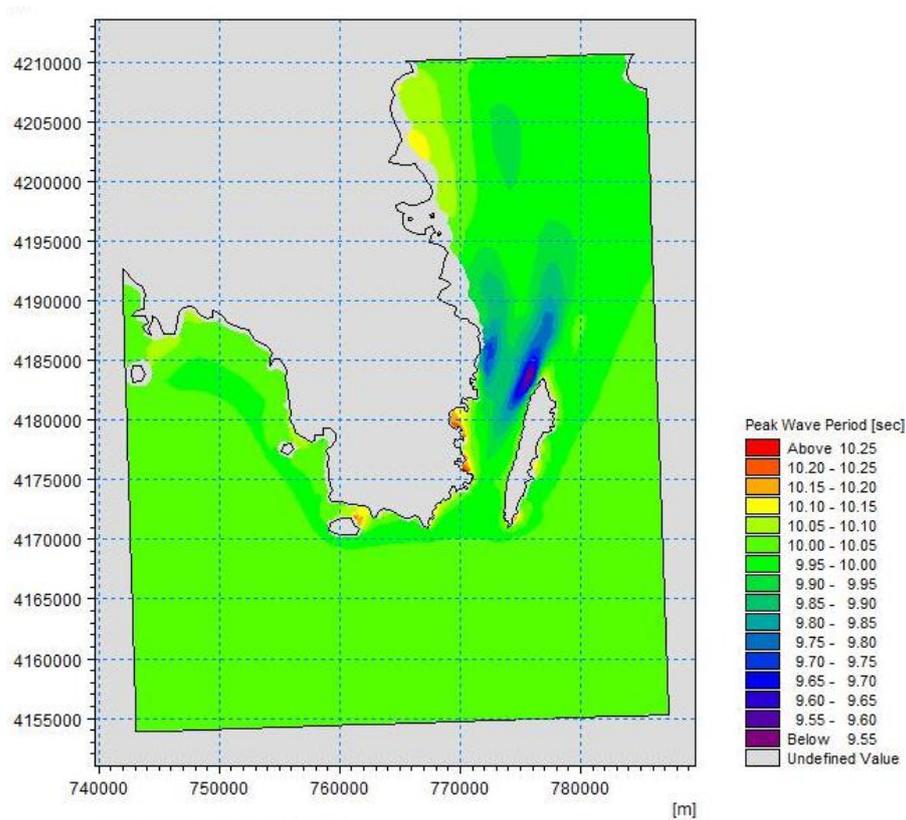


Figure 39 6th stationary run result for  $T_p$

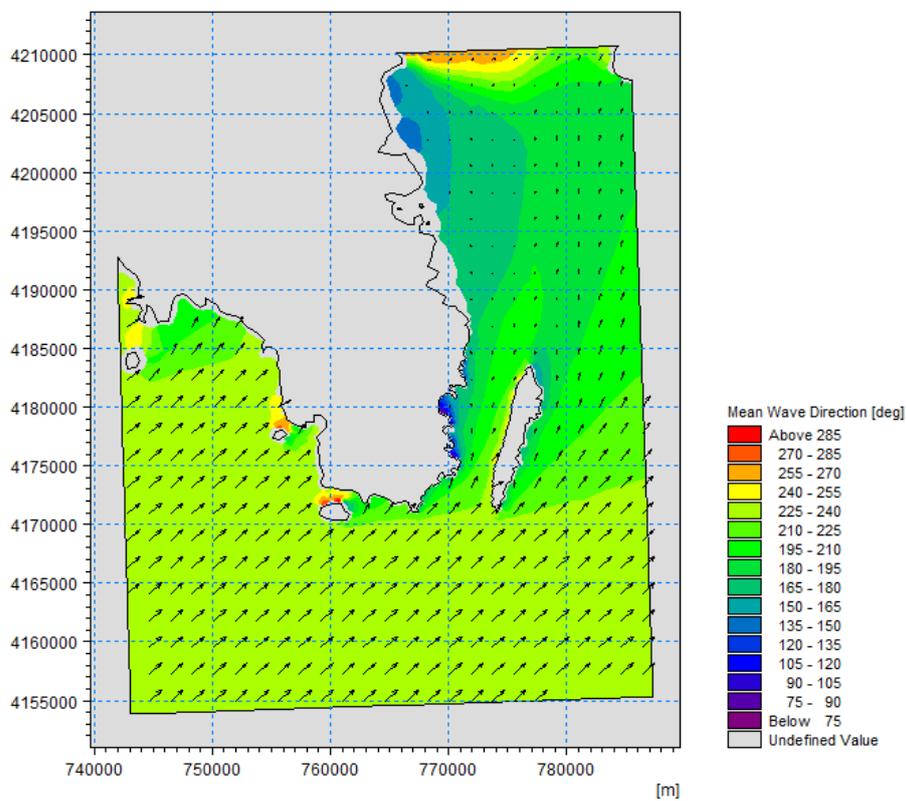


Figure 40 6th stationary run result for  $\theta$

Of all the stationary runs, the seventh as well as the eighth runs involved the highest significant wave height ( $H_s = 2$  m) and wind speed input values ( $u_w = 12.3$  m/s). The results of the seventh run concerning the significant wave height demonstrate the same pattern as all the previous runs which included wind coming from  $15^\circ$  N, meaning that the highest waves are found on the north-east boundaries ( $> 2.10$  m) while on the west nearshore part of the area as well as some protected embayments at eastern Attica (e.g. Porto Rafti, Thoriko, Lavrio) there is the appearance of the shading effect (Fig. 41). The result of the peak period suggest that it is the same as the input period except at the southwest of the area where it drops from  $\sim 7$ s to 2.5 s (Fig. 42). Mean wave direction appears the same as the wind direction input except on the eastern nearshore areas of Attica Makronisos and Megalonisos, possibly due to the bathymetry of those areas (Fig. 43).

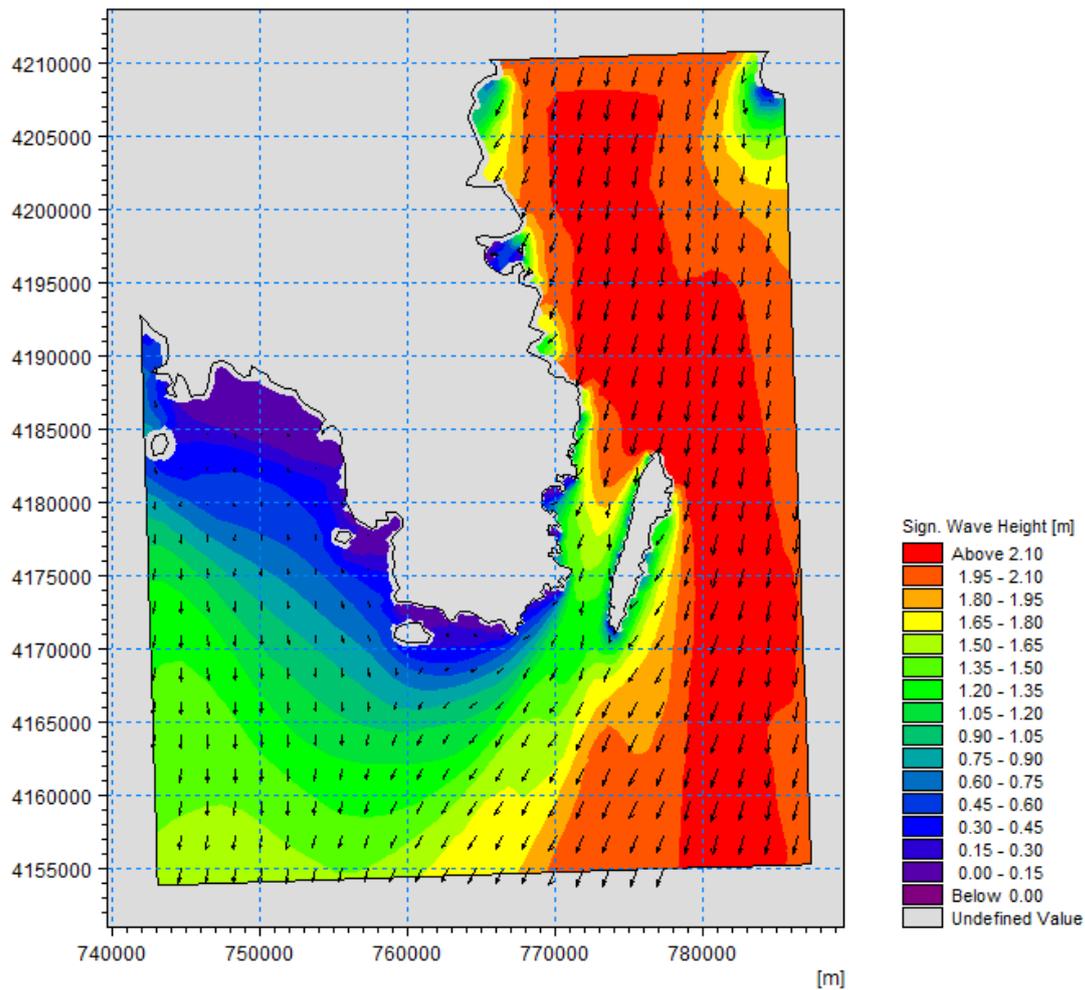


Figure 41 7th stationary run result for  $H_s$

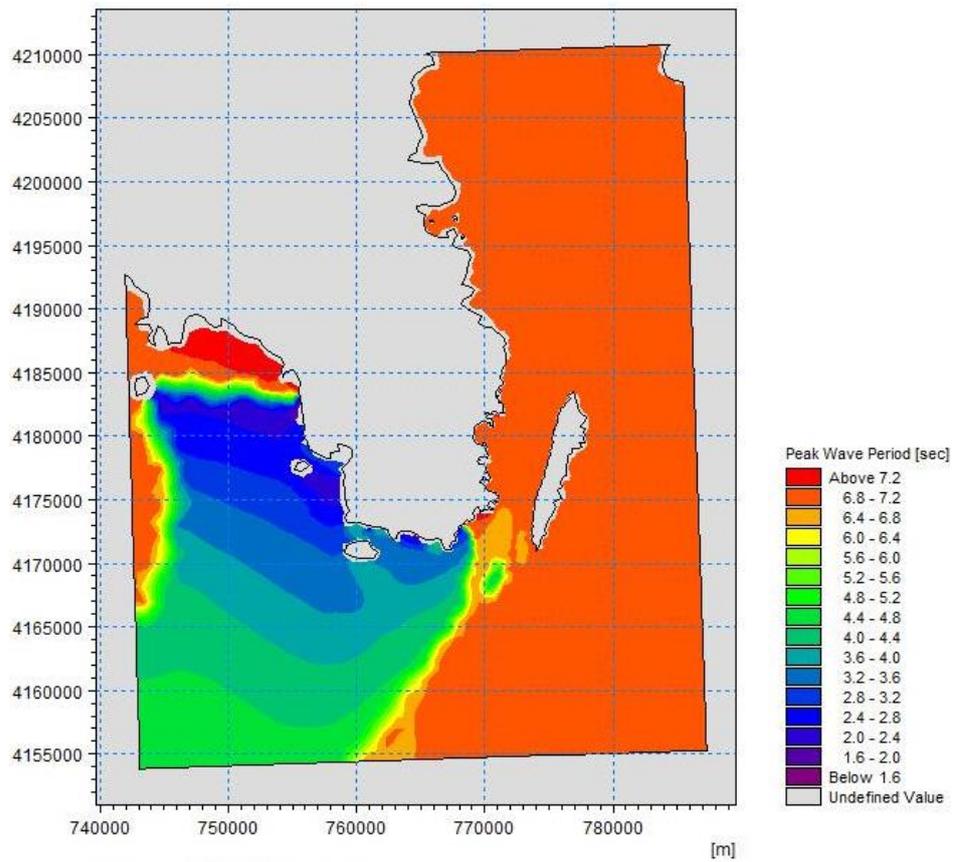


Figure 42 7th stationary run result for  $T_p$

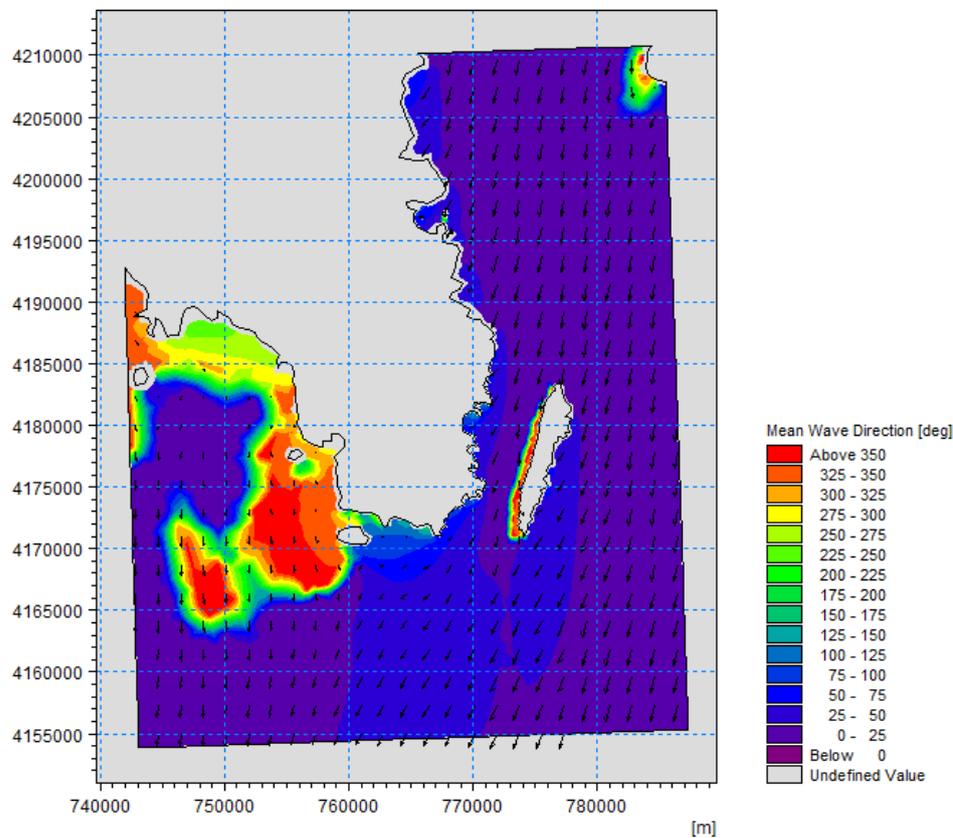


Figure 43 7th stationary run result for  $\theta$

Finally, the eighth stationary run had the same input parameters as the previous run except  $u_w$  and  $\vartheta$  which were set to  $230^\circ$  S-SW. The results of  $H_s$ ,  $T_p$  and  $\vartheta$  coincide with the previous run with the same direction of wind (Fig. 44-46). After reviewing all the results from the stationary runs, an important observation is that with the given inputs, Makronisos channel appears to be affected by presumably more intense waves on the north and south entrances. On the one hand, when the wind blows from  $15^\circ$  N Sounio nearshore area and the western part of Attica are protected from high waves and Makronisos channel is affected mostly at its north entrance but on the other hand when wind direction is  $230^\circ$  S-SW Sounio is exposed to the incoming waves but Makronisos channel is not affected as much except its southern entrance.

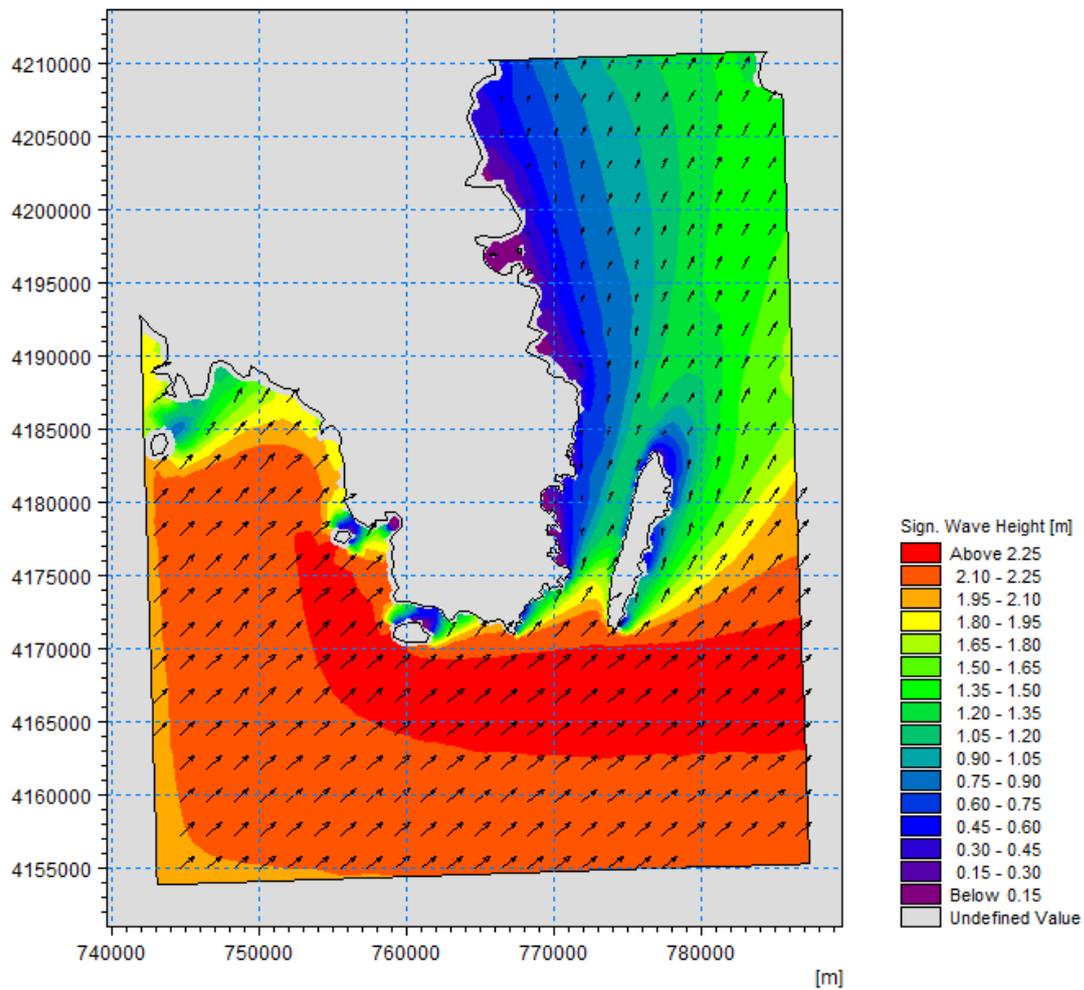


Figure 44 8th stationary run result for  $H_s$

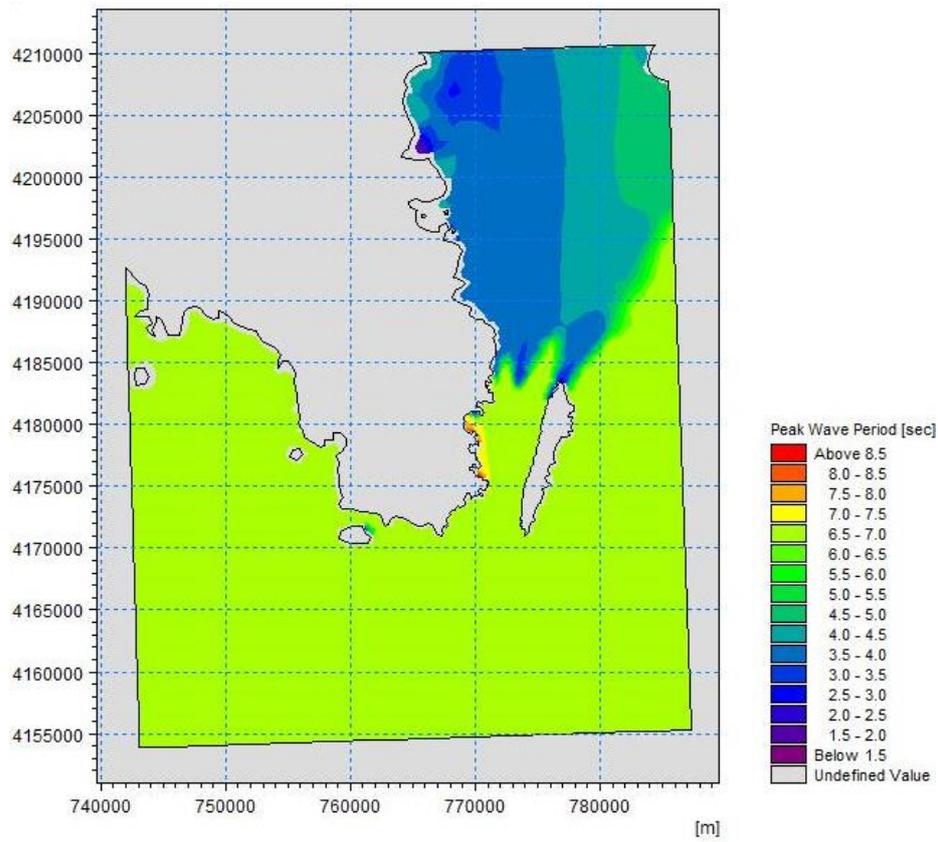


Figure 45 8th stationary run result for  $T_p$

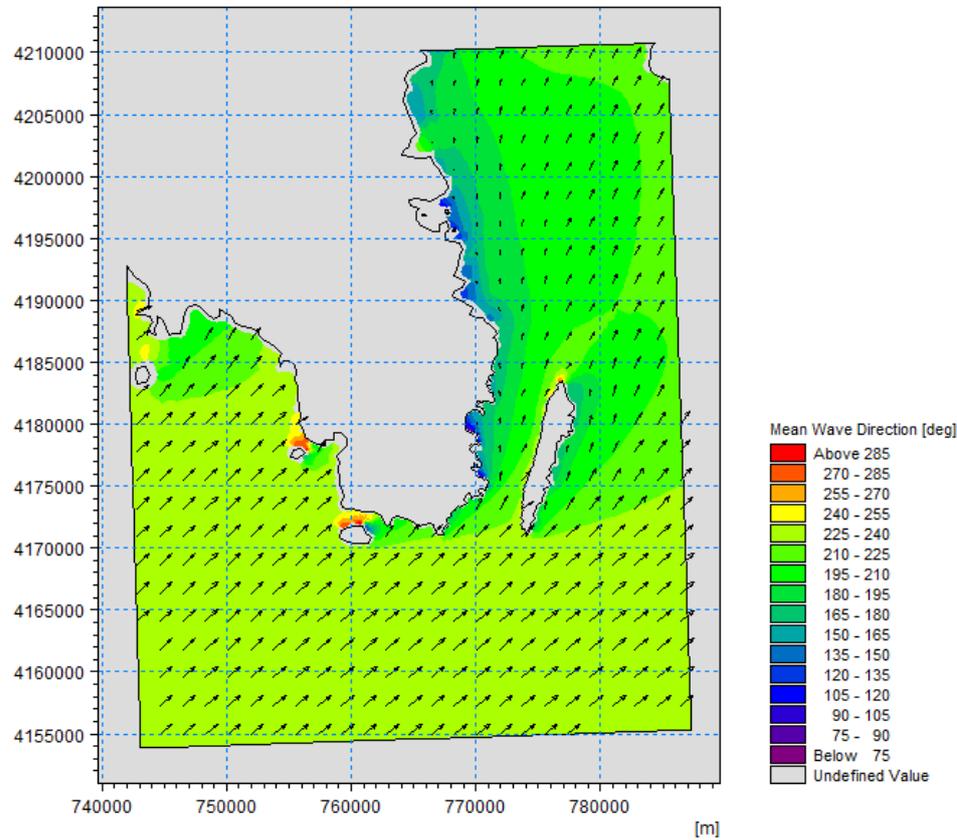


Figure 46 8th stationary run result for  $\theta$

### 4.3.2. Three year hindcast (1980-1982)

In this part, MIKE 21 SW was employed to simulate three years of wave conditions in the study region. In contrast with the stationary runs, this effort was directed towards point series results, and in this case five (5) points were selected inside Makronisos channel (as seen in Fig. 47). The coordinates of the points are seen in Table 2 in UTM34 system, and were transformed to WGS84 with Datum Convert tool from Mike Toolbox (Table 3). The model's time step (dt) was set to 4 hours. Both wind forcing and boundary conditions are products of the European Center for Medium-Range Weather Forecasts (ECMWF) ERA5 data set (<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>), which is the fifth generation of atmospheric reanalysis in production and replaced ERA-Interim reanalysis (Hersbach et al., 2020).

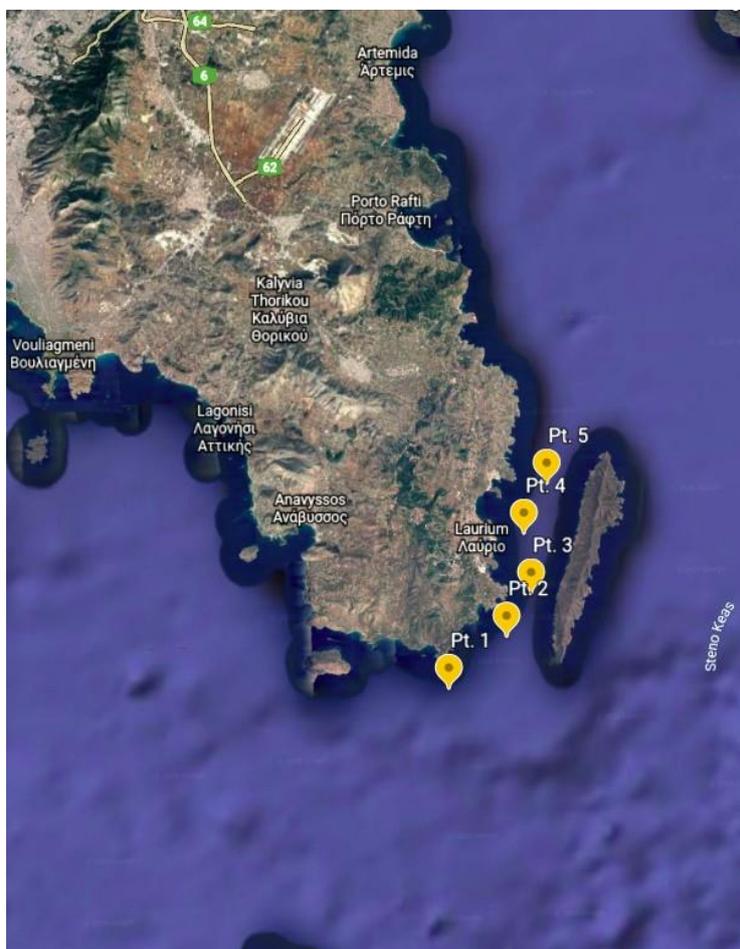


Figure 47 5 selected points for the wave simulation

	<b>Longitude</b>	<b>Latitude</b>
<b>1</b>	767437.623112626	4169292.46101007
<b>2</b>	770880.978509862	4172342.29007619
<b>3</b>	772356.702251534	4174900.21122842
<b>4</b>	771963.175920421	4178441.94820844
<b>5</b>	773340.518079316	4181393.39569178

Table 2. Simulation point coordinates in UTM 34 system

	<b>Longitude</b>	<b>Latitude</b>
<b>1</b>	24.03°	37.63°
<b>2</b>	24.07°	37.65°
<b>3</b>	24.08°	37.68°
<b>4</b>	24.08°	37.71°
<b>5</b>	24.10°	37.73°

Table 3. Simulation point coordinates in WGS84

In the case of wind forcing, wind speed components at 10 meters height ( $u_{10}$  and  $v_{10}$  in m/s) were downloaded from Copernicus Climate Data Store (CDS), <https://cds.climate.copernicus.eu/#!/home>, as hourly fields with approximately 27.75 x 27.75 km (0.25° x 0.25°) resolution. MIKE 21 SW's wind input consists of two options: (a) either time series for wind speed and wind direction which will be constant in the whole field or (b) a 2-D input file containing  $u$  and  $v$  wind components. The first option was chosen and as a consequence the magnitude of the wind speed's vector was calculated considering wind speed as a complex number  $w = u_{10} + iv_{10}$ . As a result,  $w$ 's magnitude is given by  $|w| = |u_{10} + iv_{10}|$ . Given the assumption that wave direction cannot differ significantly from wind direction, the values of the mean wave direction at the offshore point [37.5°N, 24°E] were used in place of wind direction.

With regards to boundary conditions, the wave field used in this study was downloaded from CDS along with wind data, as hourly fields with a coarser resolution of 55.5 x 55.5 km. And for this particular reason, it was decided to assign to the open boundaries constant values along their lines provided that there was only one offshore point with ERA5 wave data in accordance with the size of the study area. Specifically, the distance between the southern and northern point of the area is 1/2 of a degree, ~ 55.5 km, which coincides with the distance between west and east corner (1/2 degree ~ 55.5 km). This consequently translates into only one available point of wave data which is the offshore point at [37.5°N, 24°E]. The next offshore points with available wave data were at [37.5°N, 24.5°E] and [38°N, 24.5°E]. Thus, the wave data from [37.5°N, 24°E] were used to cover the west and

south boundaries of the domain, and the other two offshore points, [37.5°N, 24°E] and [38°N, 24.5°E], were carried over to the east and north boundaries respectively.

After having completed the assignment of each offshore point to the corresponding boundary, wind and wave data were sub-sampled in order to transform them to 4-hour step time series. It is notable that the input of the wave field for the boundary conditions requires values for significant wave height  $H_s$ , peak wave period  $T_p$ , mean wave direction  $\vartheta$  and directional spreading index. The first three parameters were given from ERA5 dataset. The value assigned to the directional spreading index was 5.

## 5. Wave climatology inside Makronisos channel and extreme value analysis

The simulation of the wave characteristics inside Makronisos channel and Sounio nearshore region was not completed for the whole period that it was intended to (meaning the 40 year period of 1980-2019) owing to insufficient computational resources. Instead, three year (1980-1982) time series of selected wave parameters ( $H_s$ ,  $T_p$  and  $\vartheta$ ) at specific locations were effectively produced in order to analyze and extract a conclusion about the wave conditions at the specific site as well as to assess any possible extreme events which could take place.

### 5.1. Univariate distributions of nearshore data

The statistical analysis for all five target points (whose coordinates are presented in Table 2 and 3) was conducted in the same manner as for the offshore data, and as mentioned before in this document, it is explained more in depth in Appendix A and for further reading it is suggested to see the work of Athanassoulis & Belibassakis (2002). At this point it is worth mentioning that the model performed well in general terms, although when the predicted significant wave height value was zero, the results for peak period and mean wave direction were also zero. For this reason, and because a period of 0s does not have any physical explanation, peak period values for all target points were filtered, resulting in the extraction of  $T_p$  values which were less than 1 s although this was only applied for the univariate distributions. The first approach is with the univariate distributions. Starting from the southernmost point of Makronisos channel, Target Point 1 (TP1), according to the statistics the most probable significant wave height is between the interval 0-0.5 m (82%) and the probability that it will exceed 1.5 m is 2% (Fig. 48).

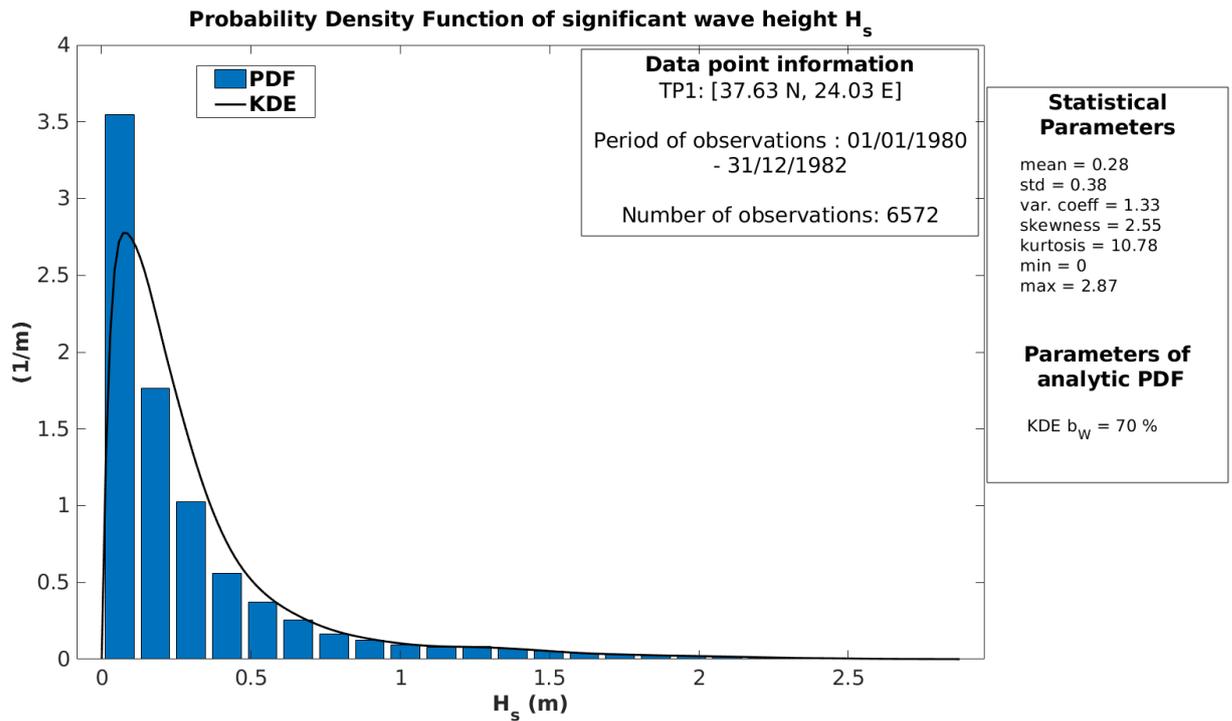


Figure 48 Probability density function of  $H_s$  at TP1

As for the peak period (Fig. 49), the highest probability lies between 3-4 s (31 %) followed by 4-5 sec (22 %). With regards to mean wave direction (Fig. 50), the most probable directions are between  $0^\circ$ - $15^\circ$  (12 %),  $30^\circ$ - $60^\circ$  (35 %) and  $180^\circ$  -  $240^\circ$  (30 %).

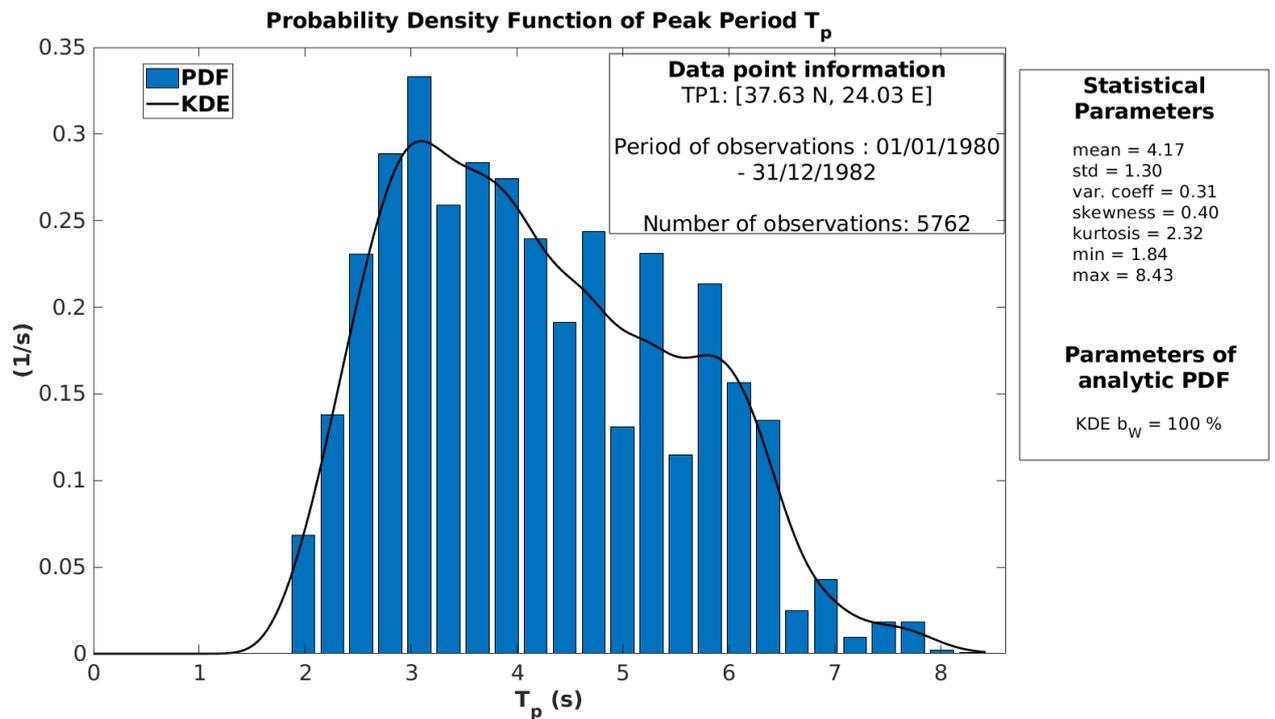


Figure 49 Probability density function of  $T_p$  at TP1

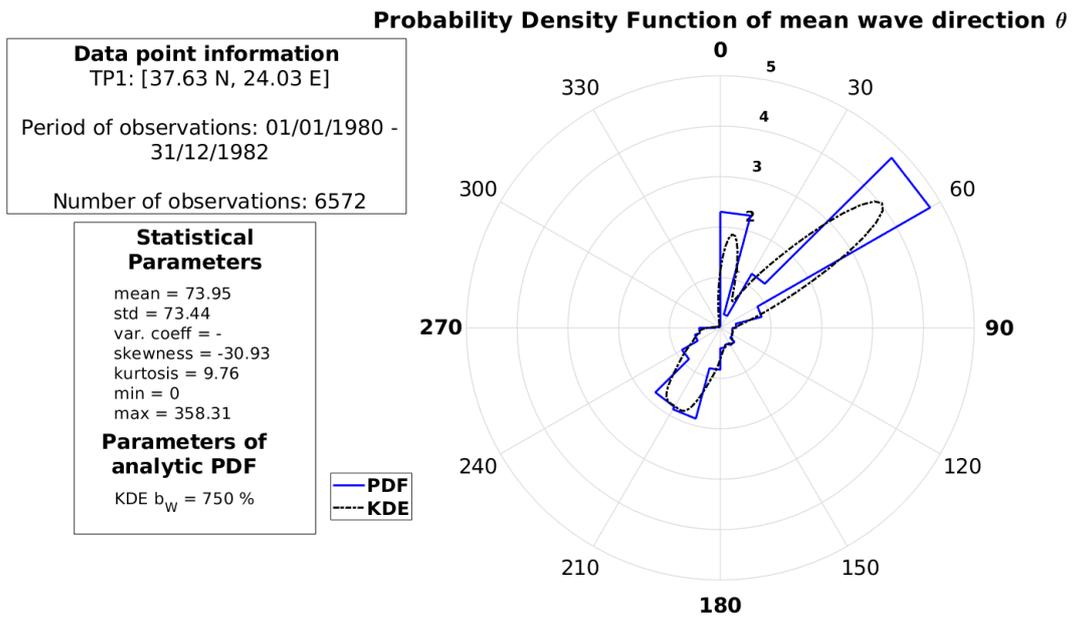


Figure 50 Probability density function of  $\theta$  at TP1

At TP2, the 79% of the predicted significant wave height was between 0-0.5 m with a 2% probability to exceed 1.5 m (Fig. 51), while the two most probable peak period intervals were between 5-6 s (25 %) and 4-5 s (22 %) (Fig. 52). In this case the most probable mean wave directions appears to be from 15°-30° (47 %) and 200°-220° (15 %) (Fig. 53).

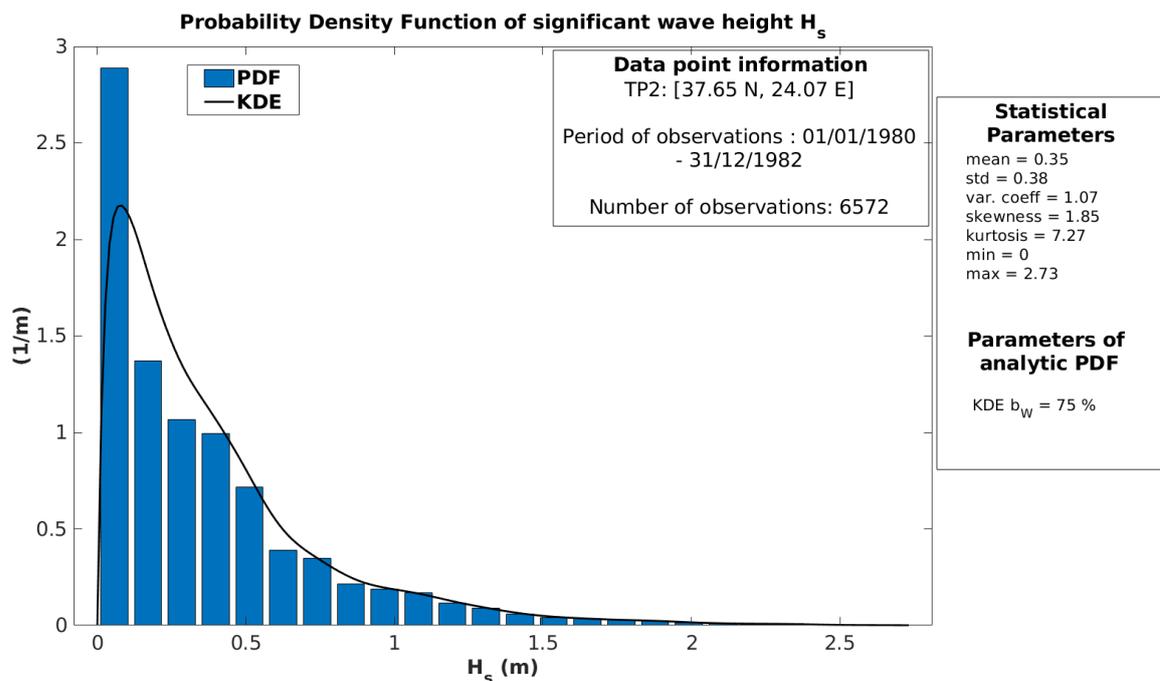


Figure 51 Probability density function of  $H_s$  at TP2

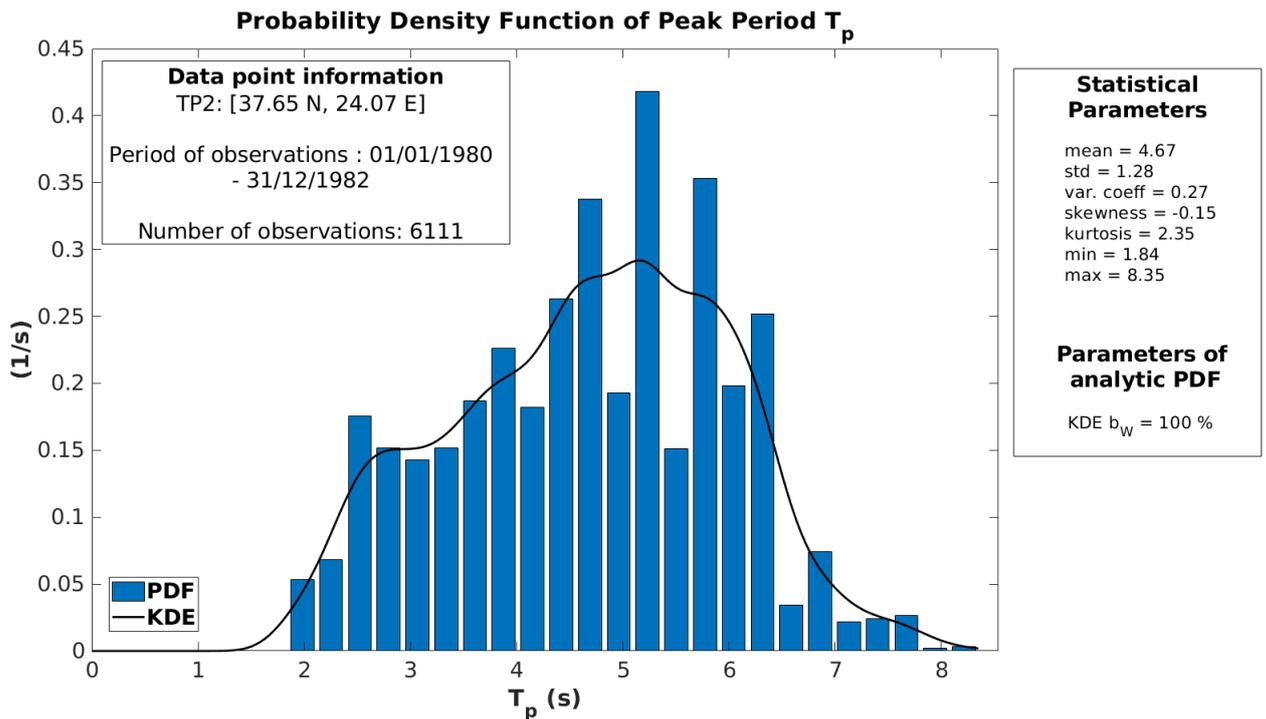


Figure 52 Probability density function of  $T_p$  at TP2

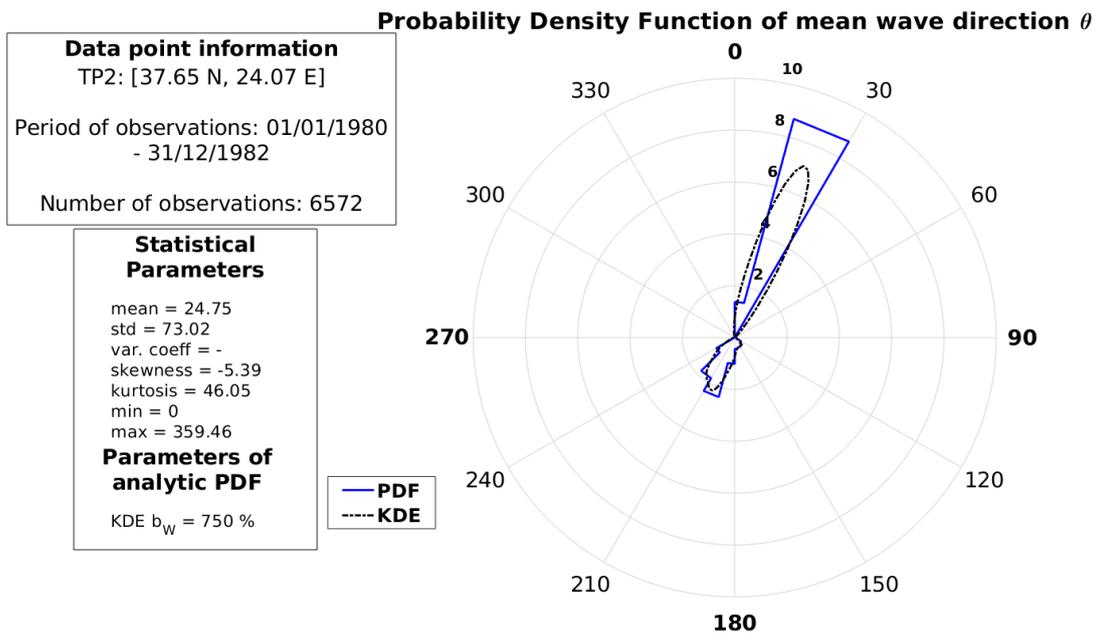


Figure 53 Probability density function of  $\theta$  at TP2

To continue, at the next target point TP3, the highest probability with regards to the significant wave height is between 0-0.5 m (69 %) with only 1% exceeding 1.5 m (Fig. 54). The most probable peak period (Fig. 55) lies between 4-6 s (50%) and mean wave direction has a 56 % probability at  $0^\circ$ - $15^\circ$  and 27 % at  $190^\circ$ - $210^\circ$  (Fig. 56).

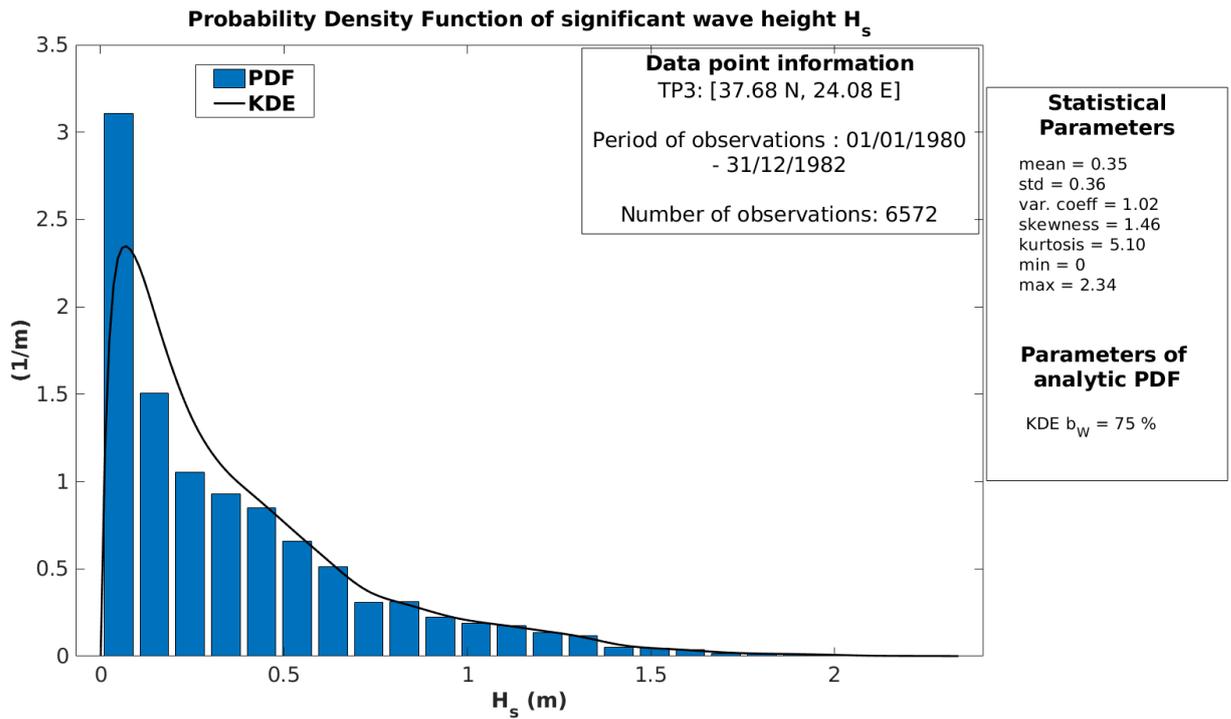


Figure 54 Probability density function of  $H_s$  at TP3

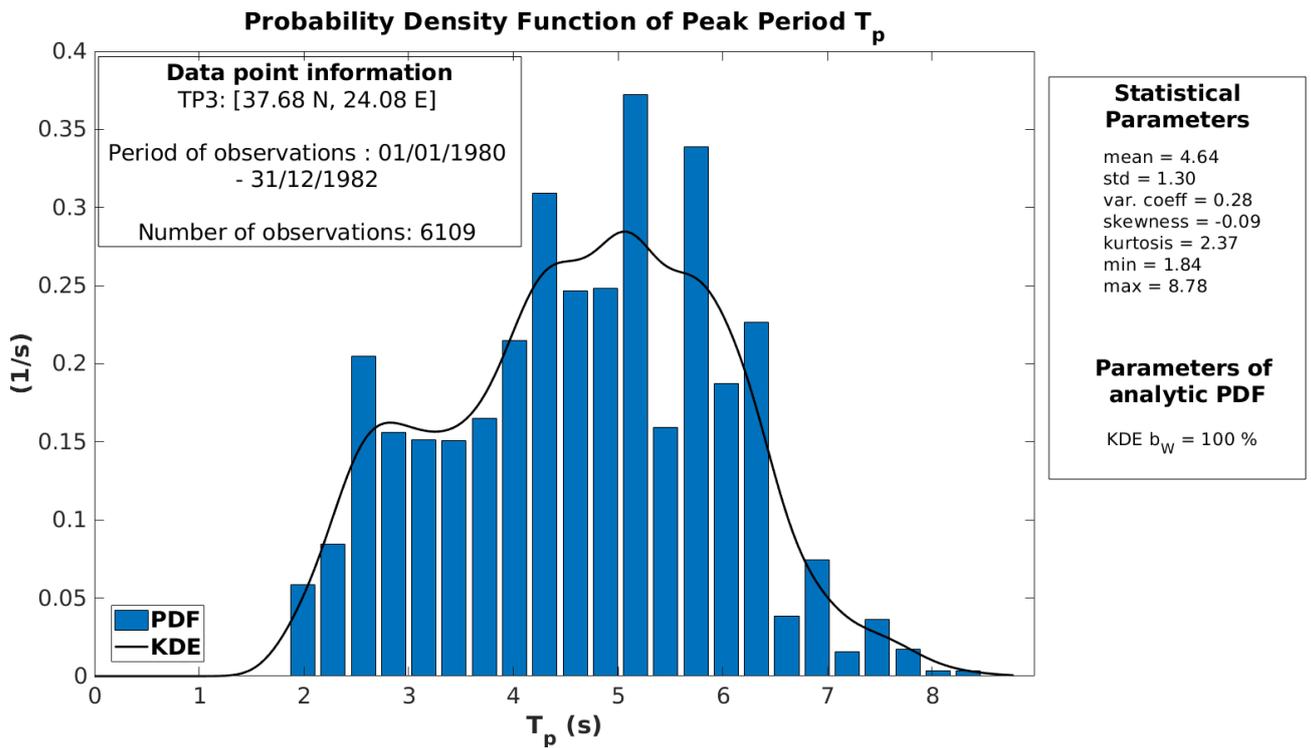


Figure 55 Probability density function of  $T_p$  at TP3

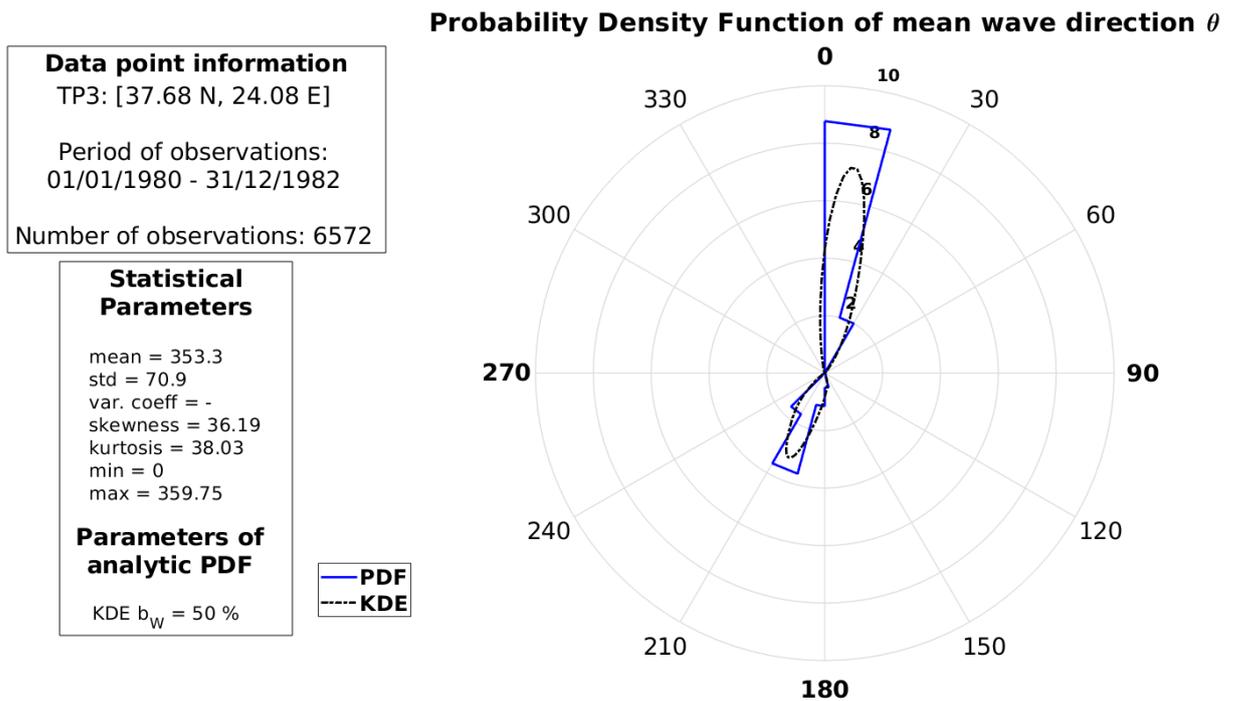


Figure 56 Probability density function of  $\theta$  at TP3

In TP4 there is a 77 % probability that the significant wave height will be between 0-0.5 m (Fig. 57) and less than 1 % that it will exceed 1.5 m. The highest in terms of probability interval in which peak period appears is between 4-6 s (45 %) (Fig. 58) and the two most probable mean wave directions are from 15°-30° (46 %) and 180°-195° (31 %) (Fig. 59).

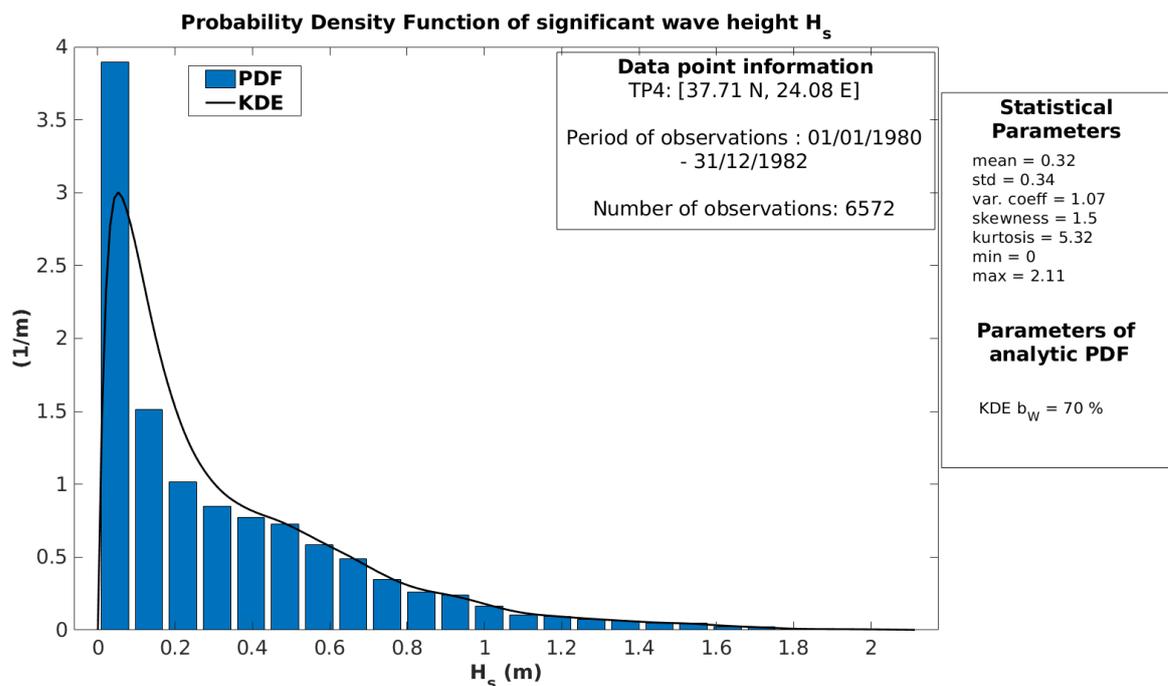


Figure 57 Probability density function of  $H_s$  at TP4

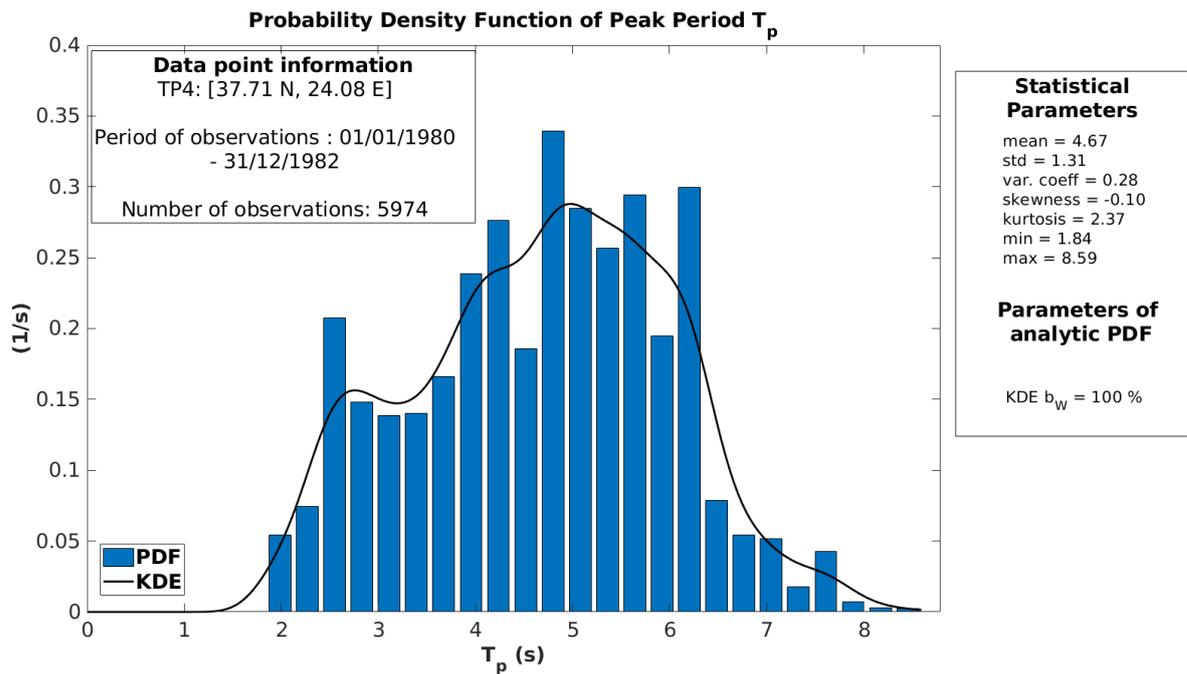


Figure 58 Probability density function of  $T_p$  at TP4

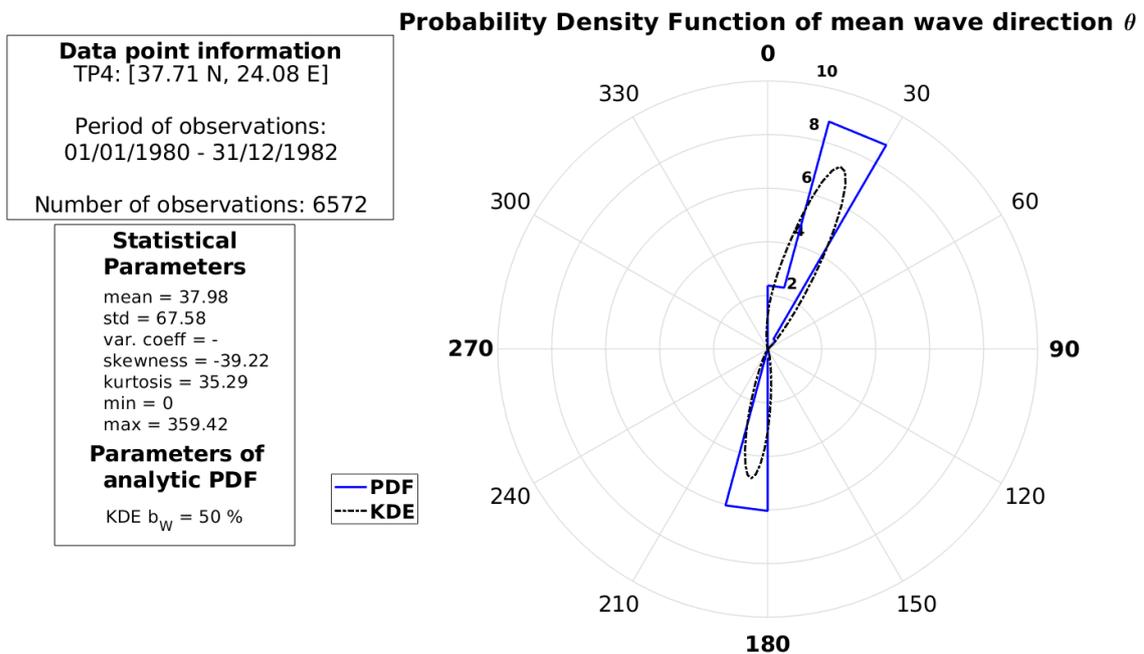


Figure 59 Probability density function of  $\theta$  at TP4

Finally, in the northernmost target point (TP5) significant wave height has a 63 % probability to be expected between 0-0.5 m and 3 % to exceed 1.5 m (Fig. 60). Peak period is most probable to be between 4-6 s (46 %) (Fig. 61) and the two main mean wave directions are 0°-30° (60 %) and 180°-200° (32 %) (Fig. 62).

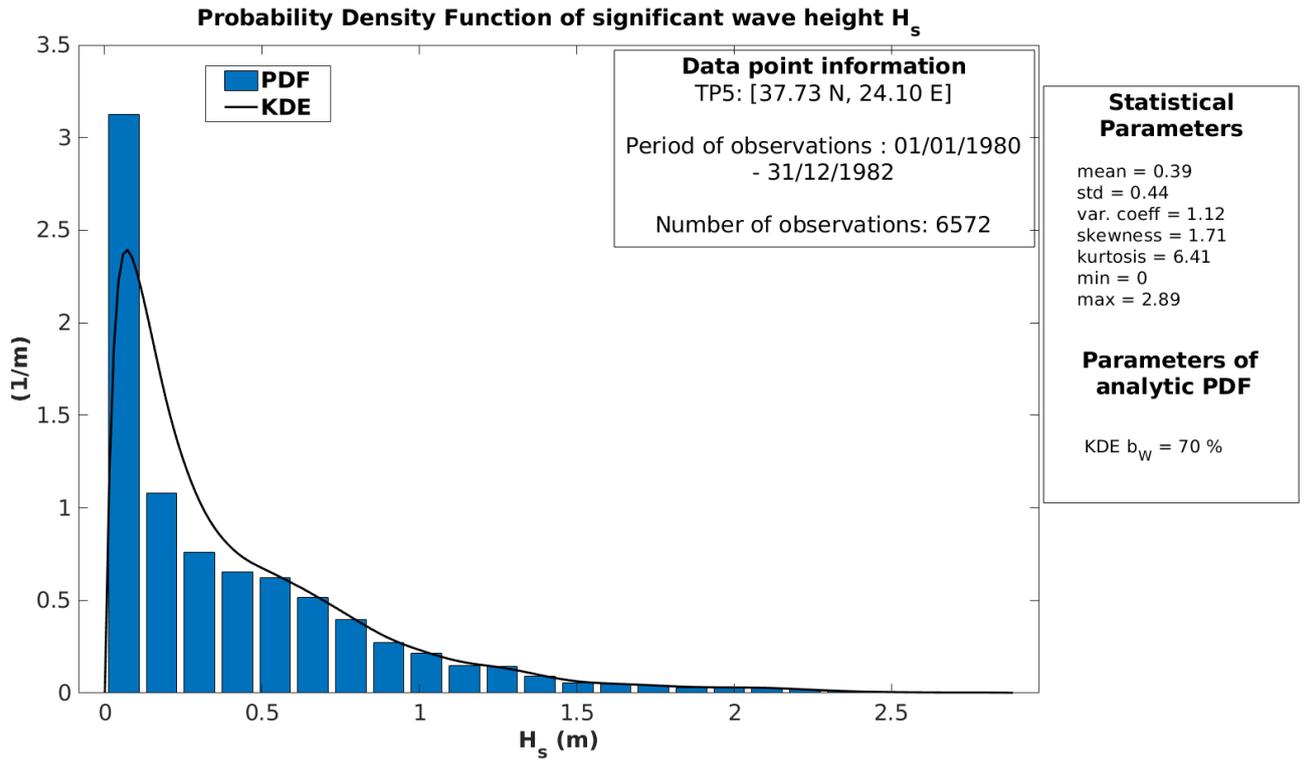


Figure 60 Probability density function of  $H_s$  at TP5

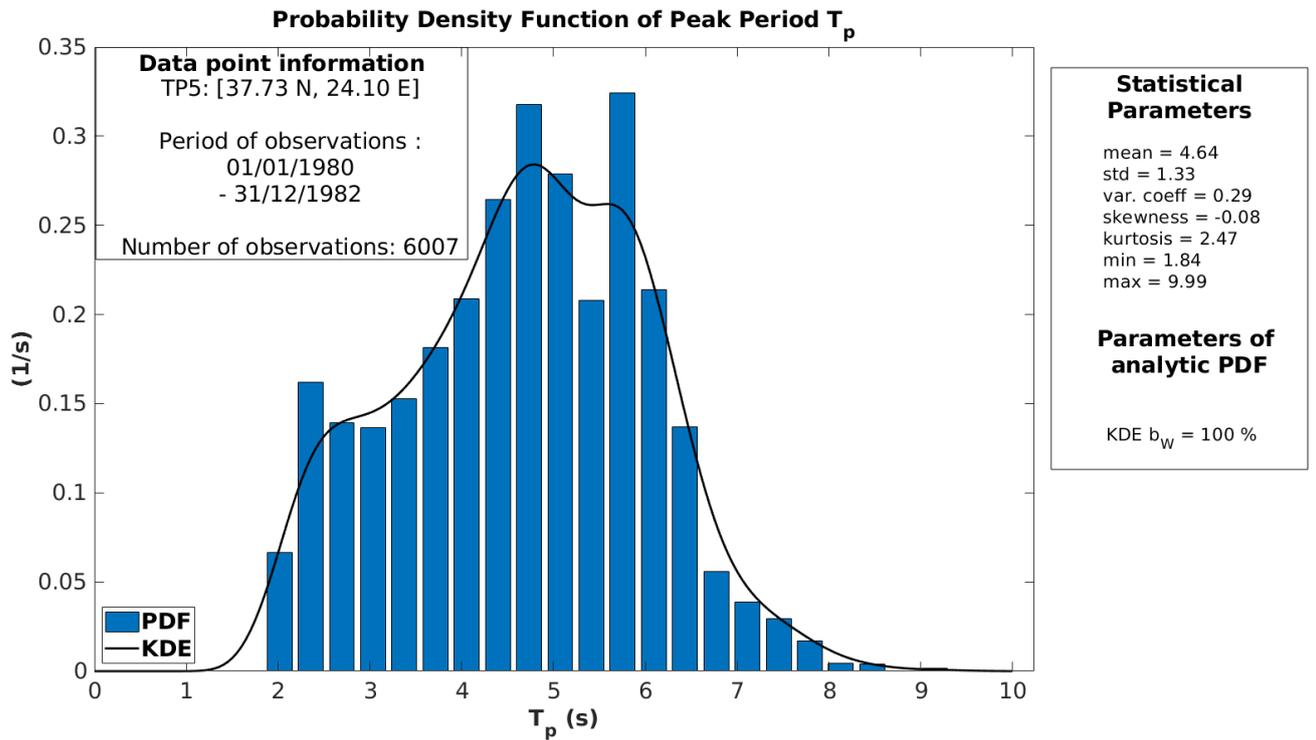


Figure 61 Probability density function of  $T_p$  at TP5

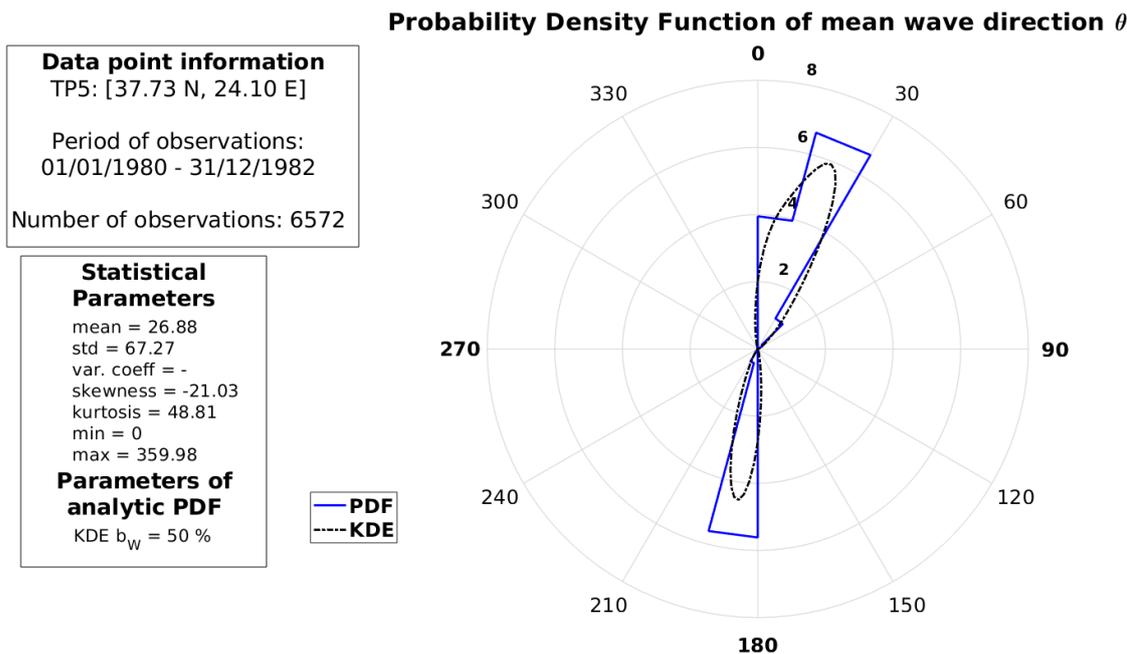


Figure 62 Probability density function of  $\theta$  at TP5

## 5.2. Bivariate distributions of nearshore data

The bivariate distributions  $(H_s, T_p)$  of all five target points are illustrated in Fig. 63-67. They present quasi the same characteristics with very subtle differences. At TP1 the most likely combination is of  $\{0 < H_s < 0.4 \text{ m}\}$  and  $\{2 \leq T_p < 3.5 \text{ s}\}$ . At TP2 the highest probability is detected at  $\{0 < H_s < 0.25 \text{ m}\}$  and  $\{2 \leq T_p < 2.5 \text{ s}\}$  followed by  $\{H_s \approx 0.5 \text{ m}, 5 < T_p < 6 \text{ s}\}$ . Next, at TP3 the most probable combination is of  $\{0 < H_s < 0.25 \text{ m}\}$  and  $\{2 \leq T_p < 3 \text{ s}\}$  as well as  $\{0.25 < H_s < 0.5 \text{ m}\}$  and  $\{4.5 < T_p < 5.2 \text{ s}\}$ . As for TP4,  $\{0 < H_s < 0.25 \text{ m}\}$  and  $\{2 \leq T_p < 3 \text{ s}\}$  is also the most probable combination along with  $\{0.3 < H_s < 0.6 \text{ m}\}$  and  $\{4 \leq T_p < 5.5 \text{ s}\}$ . Finally, at TP5 the combinations of  $H_s$  and  $T_p$  which appear most are first of  $\{0 < H_s < 0.2 \text{ m}\}$  and  $\{2 \leq T_p < 3 \text{ s}\}$  and second of  $\{0.4 < H_s < 0.6 \text{ m}\}$  and  $\{4 < T_p < 5 \text{ s}\}$ .

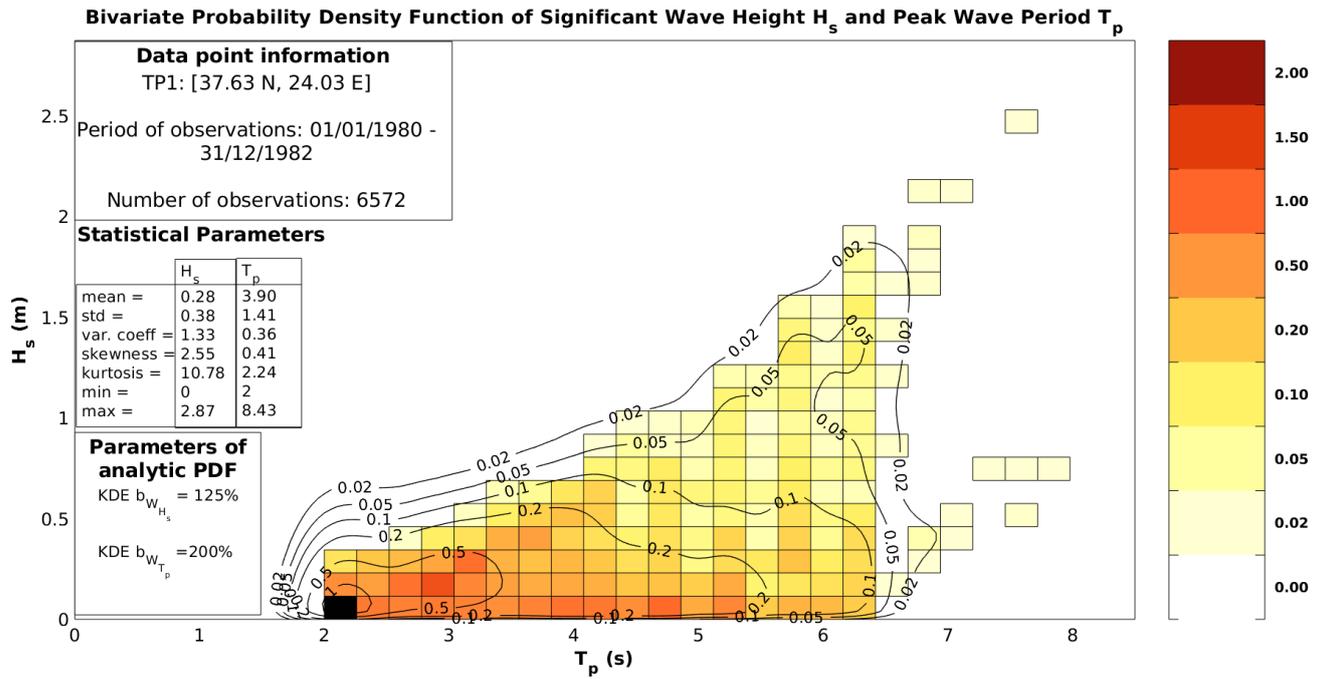


Figure 63 Bivariate Distribution of  $H_s$  &  $T_p$  at TP1

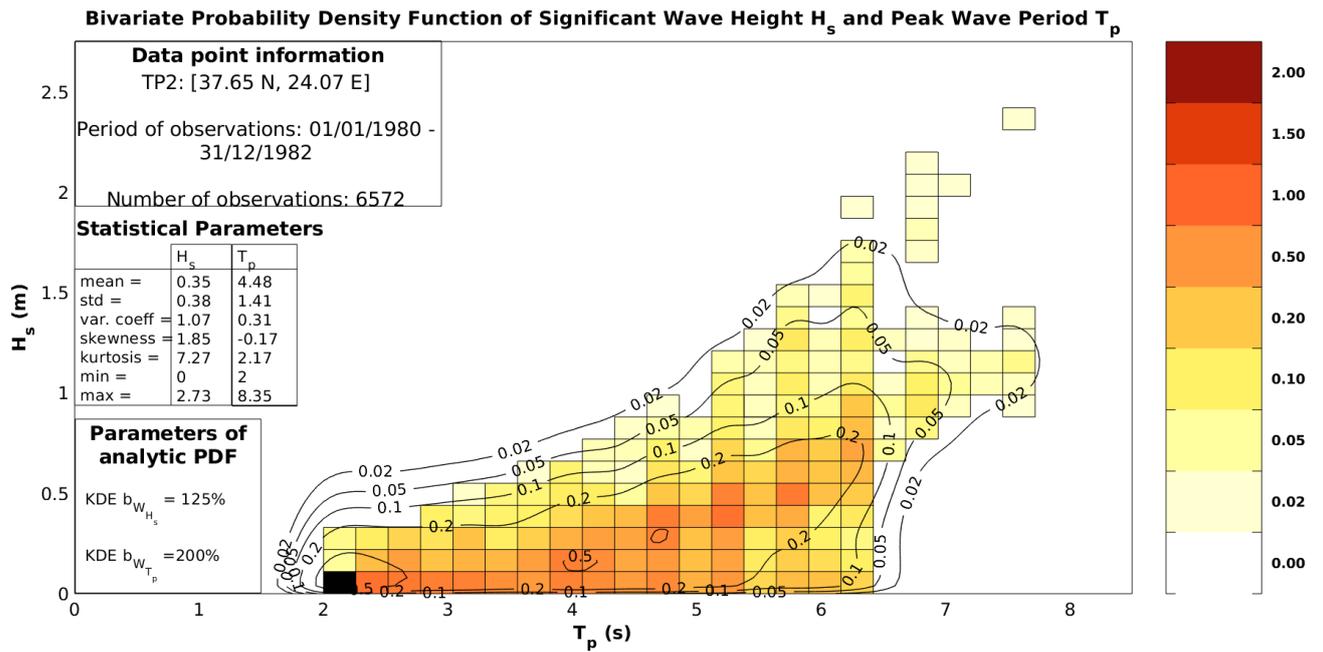


Figure 64 Bivariate Distribution of  $H_s$  &  $T_p$  at TP2

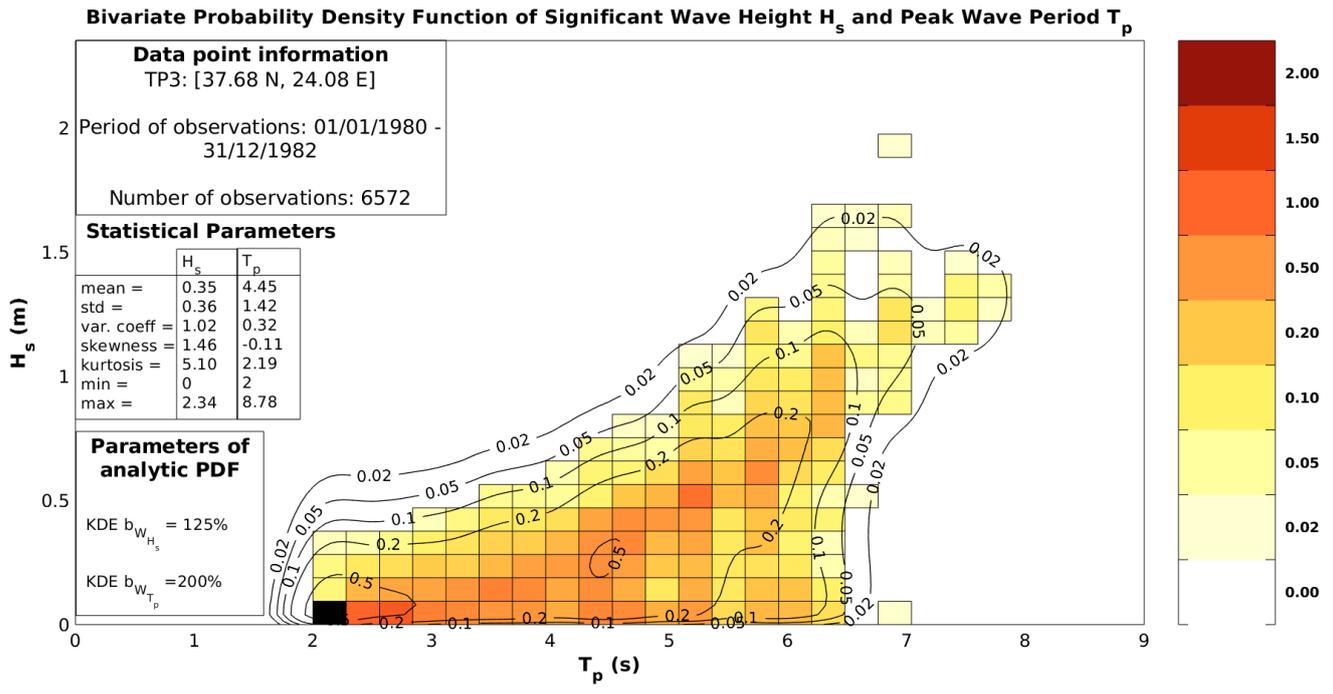


Figure 65 Bivariate Distribution of  $H_s$  &  $T_p$  at TP3

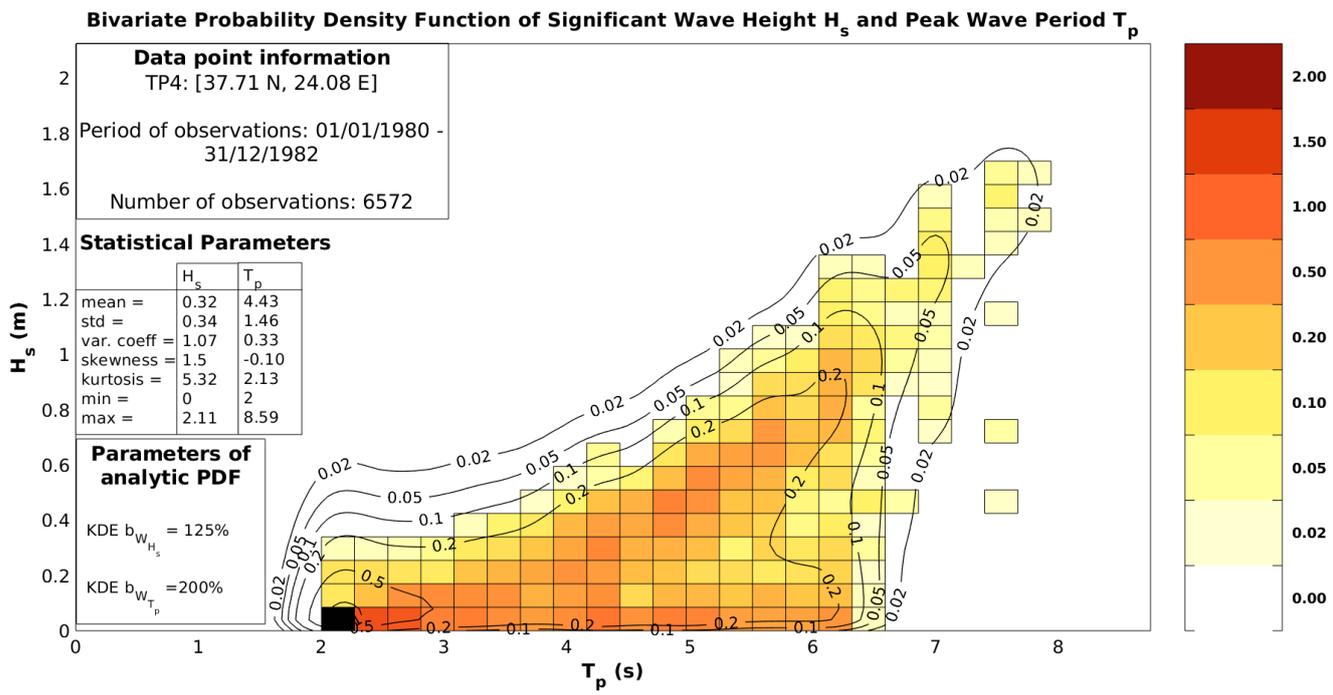
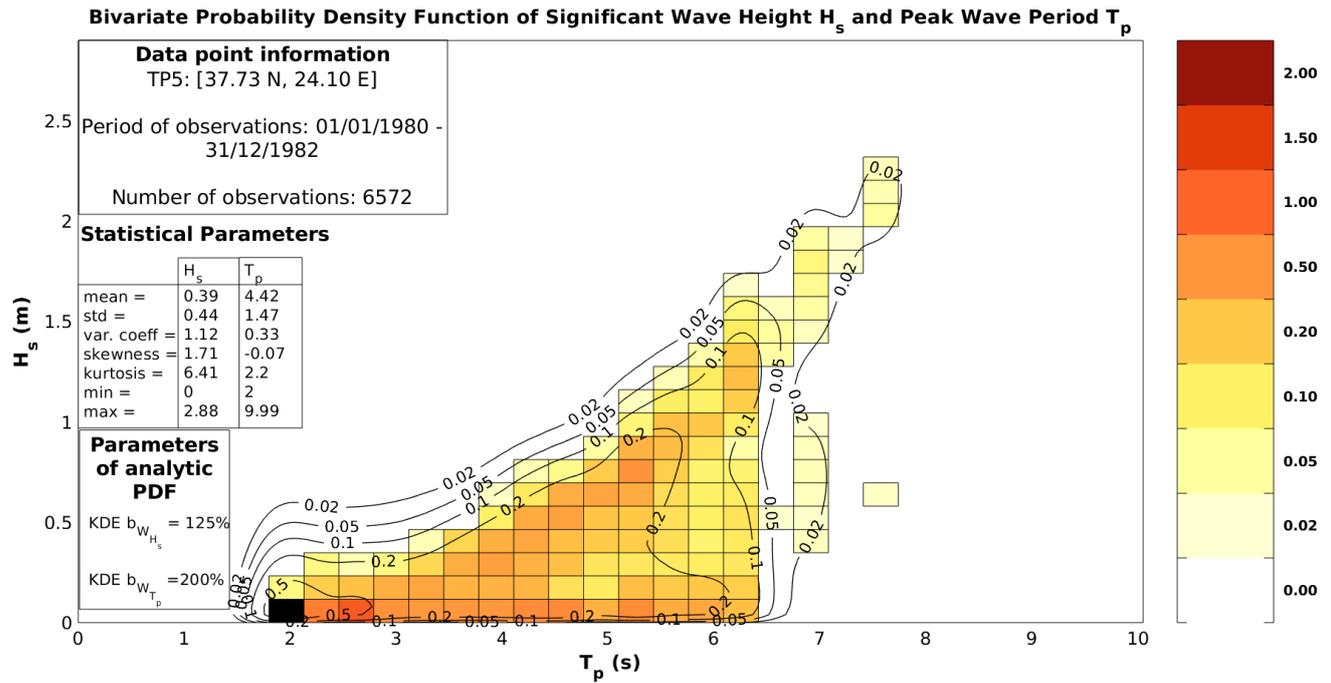


Figure 66 Bivariate Distribution of  $H_s$  &  $T_p$  at TP4



### 5.3. Wave climatology at the target points in Makronisos channel

At this point, after having discussed the probability of the wave parameters for the offshore and nearshore points, it is worth making a comparison of the actual timeseries of  $H_s$ ,  $T_p$  and  $\vartheta$  (both offshore and nearshore) regarding the same time period 1980-1983. First, the timeseries for the offshore points O1-O3 are illustrated in Fig. 68-70 and for the nearshore points in Fig. 71-75. Between the three offshore data points, O2 (the easternmost point) appears to have more intense waves in terms of  $H_s$ . As for  $T_p$  and  $\vartheta$ , in general they exhibit the same characteristics for all three offshore data points. Nearshore points on the other hand, TP1-TP5, show slightly different characteristics in comparison with the offshore points, which is expected taking into account the location of the points. Consequently,  $H_s$  and  $T_p$  are slightly smaller than at O1-O3. Mean wave direction  $\vartheta$  on the nearshore points appears to have a smaller range of angles when compared to the offshore points. As discussed earlier in chapter 3, offshore  $\vartheta$  ranges approximately between  $345^\circ - 30^\circ$  and  $150^\circ - 240^\circ$ . On the contrary, nearshore point data have a smaller interval concerning  $\vartheta$ , which is evident on Fig. 71-75. As stated earlier in this chapter, MIKE 21 SW results for  $T_p$  where close to zero (or even zero) for very small significant wave heights. The same behavior occurred for  $\vartheta$  results leading to an accumulation of mean wave direction values to zero. As for the seasonality of wave climate/intensity/parameters, it is evident that from a first look  $H_s$  is more intense during late fall-winter compared to summer months.  $T_p$  does not always seem to follow the same pattern as  $H_s$  thus it would be imperative to analyze in depth the

timeseries so as to come to a conclusion. The same applies for mean wave direction as there is no evident pattern. It should be noted that there seems to be a correlation between peaks of  $H_s$  and  $T_p$  as well as a wave direction  $\vartheta$  which comes from the north. As for the comparison between offshore and nearshore wave data, there is a difference of magnitude regarding significant wave height, meaning that offshore waves are more intense than at nearshore locations. This can also be perceived by viewing Tables 4 and 5, which present the mean values for all wave parameters at offshore and nearshore points respectively. Among the offshore wave data, O2 presents the highest mean  $H_s$  value while O1 has the lowest mean value. Correspondingly, TP5 has the highest mean  $H_s$  value while TP4 has the lowest. There is a notable difference between mean  $H_s$  values of offshore and nearshore wave data, for example  $\overline{H_s}$  (O1) = 0.44 m and  $\overline{H_s}$  (TP1) = 0.28 m.

**O1 [37.5 N, 24 E]**

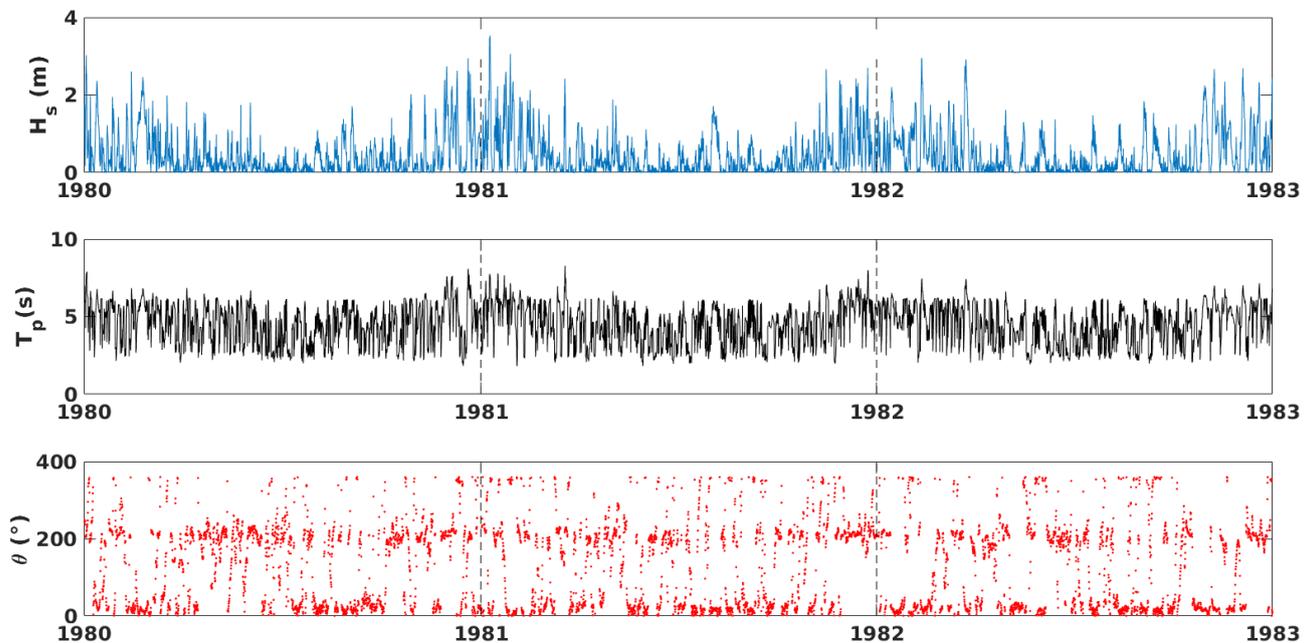


Figure 68 O1 timeseries for the period 1980-1983

O2 [37.5 N, 24.25 E]

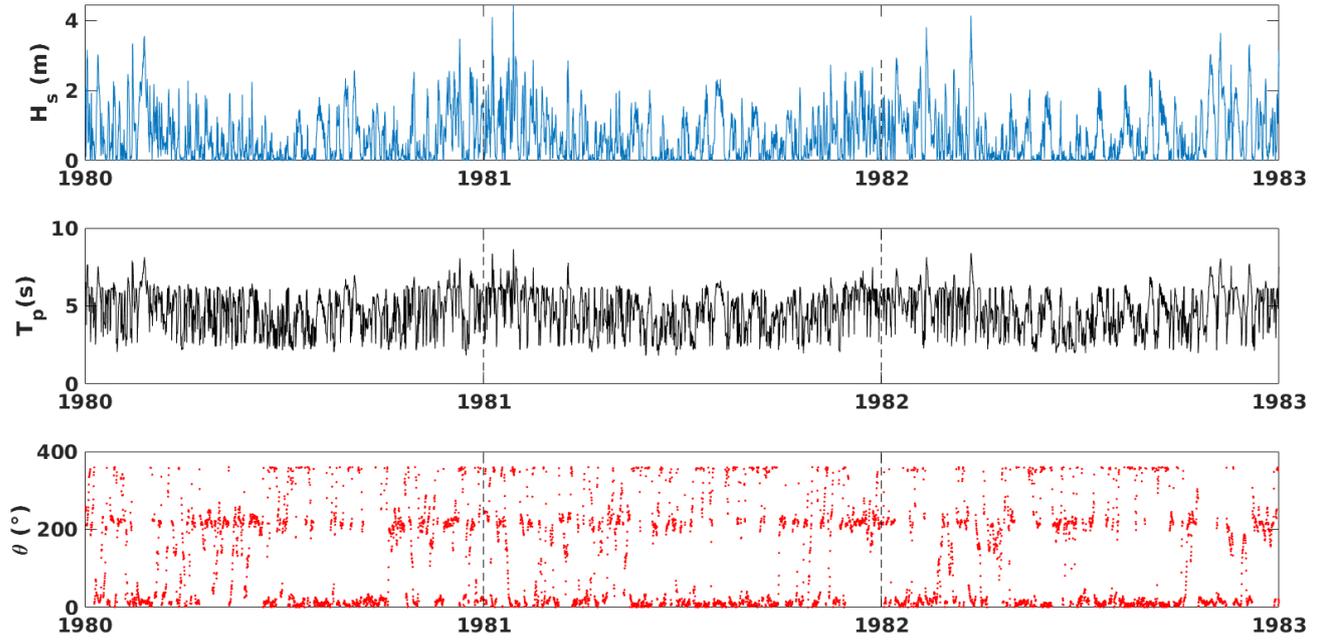


Figure 69 O2 timeseries for the period 1980-1983

O3 [38 N, 24.25 E]

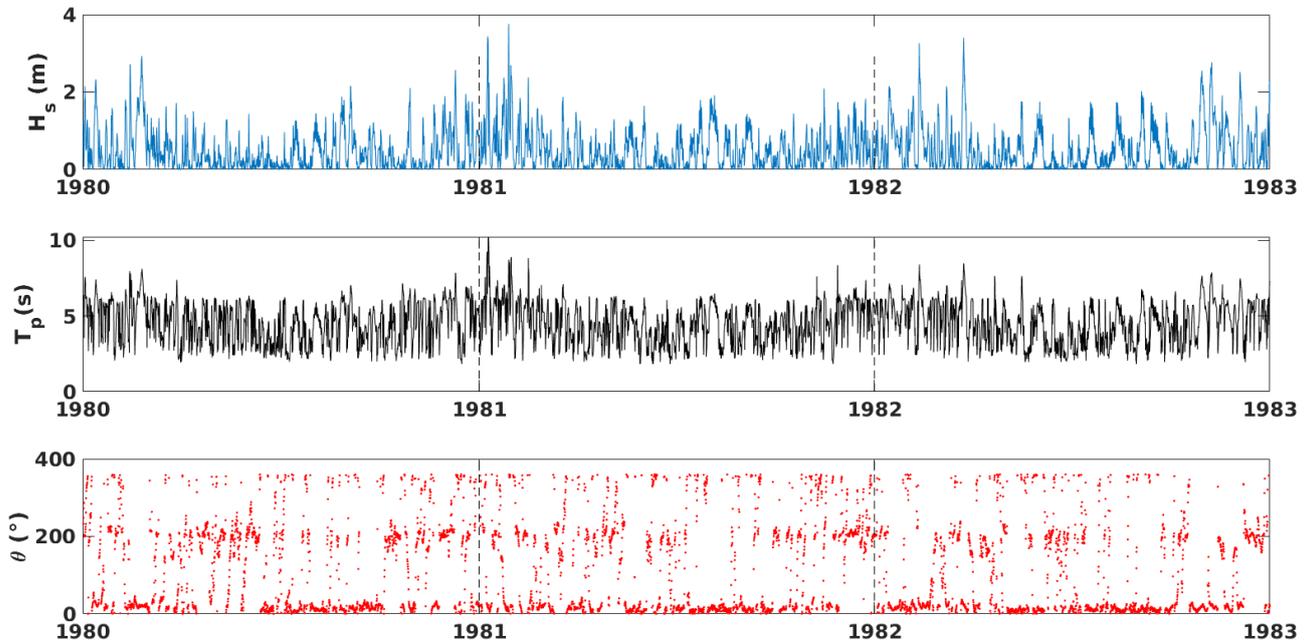


Figure 70 O3 timeseries for the period 1980-1983

TP1 [37.63 N, 24.03 E]

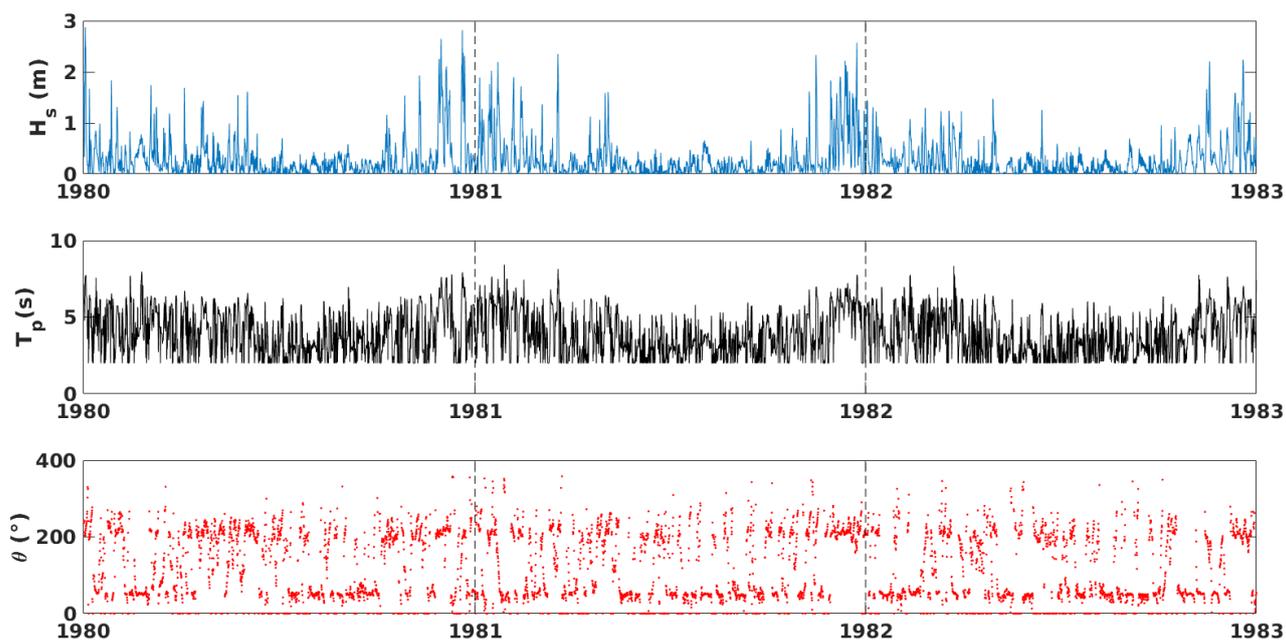


Figure 71 TP1 timeseries for the period 1980-1983

TP2 [37.65 N, 24.07 E]

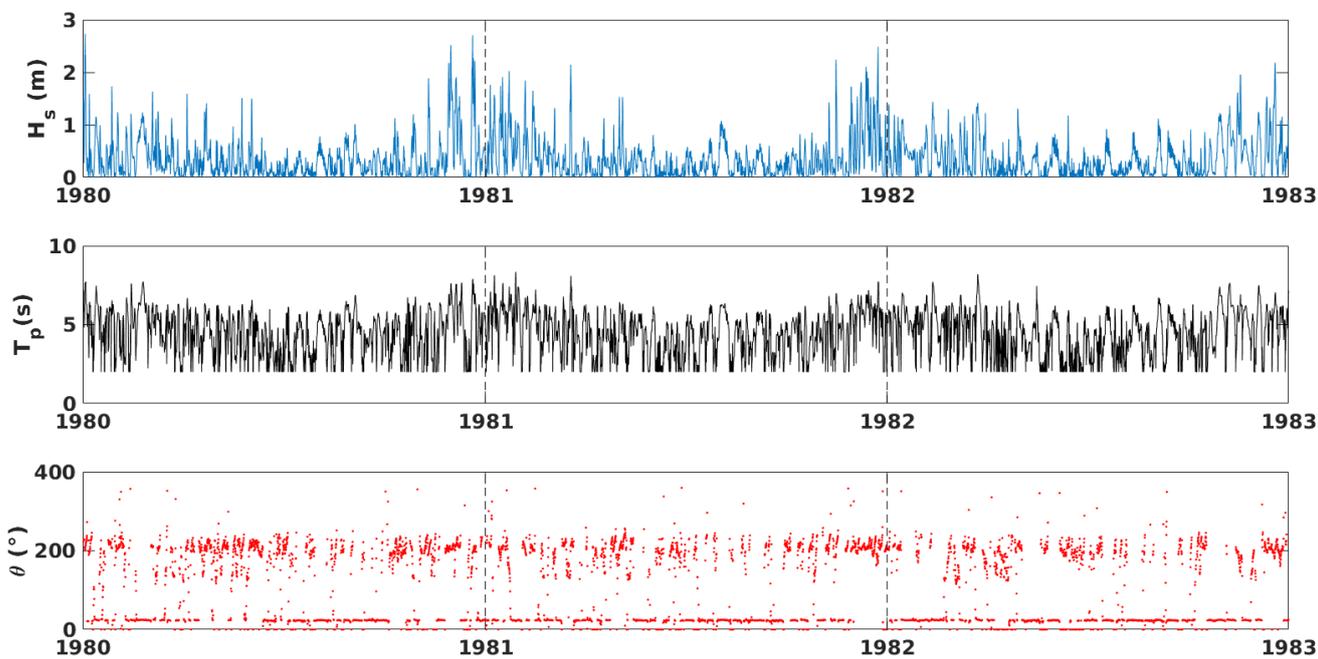


Figure 72 TP2 timeseries for the period 1980-1983

**TP3 [37.68 N, 24.08 E]**

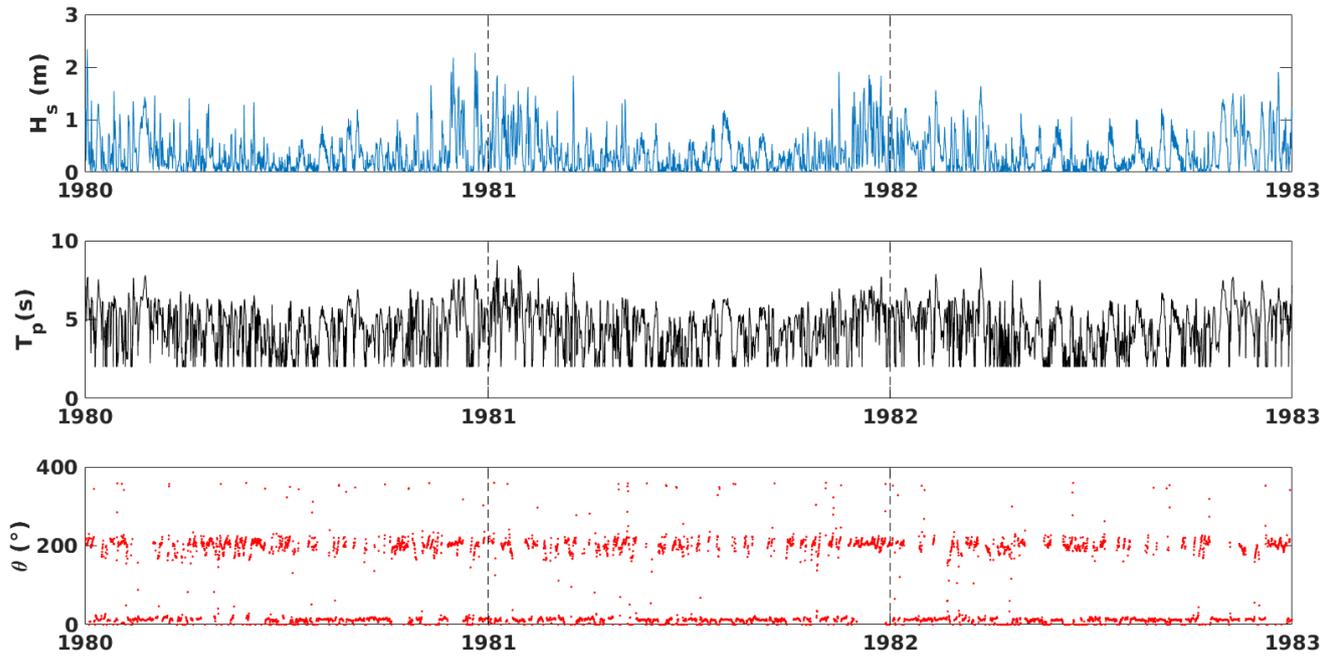


Figure 73 TP3 timeseries for the period 1980-1983

**TP4 [37.71 N, 24.08 E]**

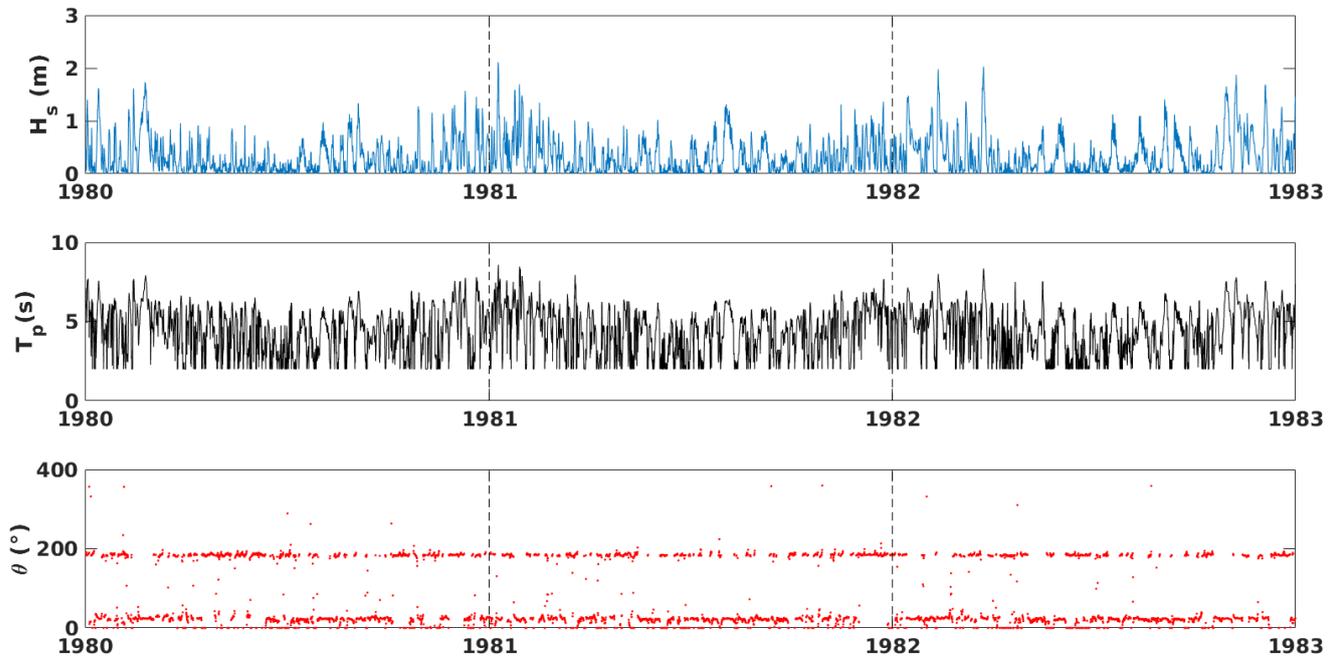


Figure 74 TP4 timeseries for the period 1980-1983

TP5 [37.73 N, 24.10 E]

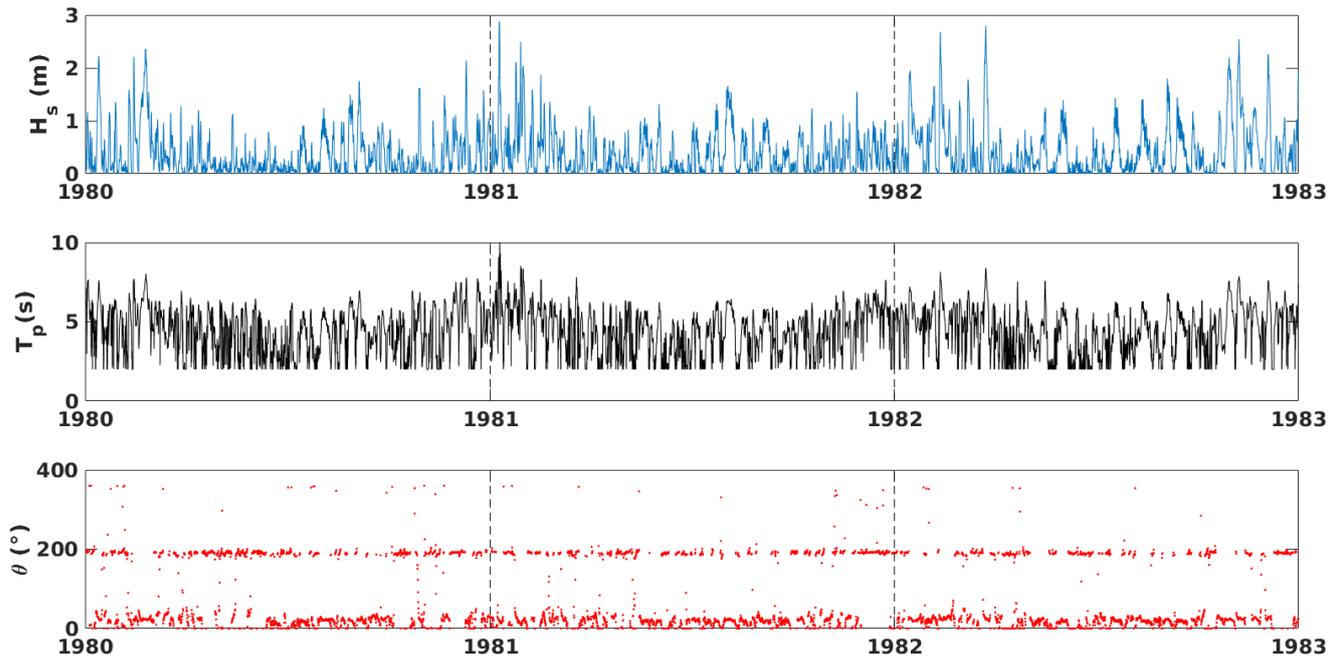


Figure 75 TP5 timeseries for the period 1980-1983

Mean	$H_s$ (m)	$T_p$ (s)	$\theta$ (°)
O1	0.44	4.23	5.39
O2	0.69	4.38	346.91
O3	0.5	4.24	10.97

Table 4 Mean values of wave parameters ( $H_s$ ,  $T_p$  and  $\theta$ ) at offshore points O1-O3

Mean	$H_s$ (m)	$T_p$ (s)	$\theta$ (°)
TP1	0.28	4.17	73.95
TP2	0.35	4.67	24.75
TP3	0.35	4.64	353.3
TP4	0.32	4.67	37.98
TP5	0.39	4.64	26.88

Table 5 Mean values of wave parameters ( $H_s$ ,  $T_p$  and  $\theta$ ) at nearshore points TP1-TP5

#### 5.4. Multivariate regression for significant wave height $H_s$

Moving forward to the multivariate regression, the produced three-year time series of significant wave height  $H_s$  from MIKE 21 SW simulation were extended applying a multivariate regression model. Three different models were tested to evaluate the errors between the observed and the predicted values, (i) a first order (ii) a second order and (iii) a third order multivariate regression model.

The first order model includes 4 terms, the second 10 terms while the third 20 terms, which are both shown in eq. (5), (6) and (7) respectively :

$$p_{H_s} = \beta_1 + \beta_2x + \beta_3y + \beta_4z \quad (5)$$

$$\{p_{H_s}\}_2 = \beta_1 + \beta_2x + \beta_3y + \beta_4z + \beta_5x^2 + \beta_6y^2 + \beta_7z^2 + \beta_8xy + \beta_9xz + \beta_{10}yz \quad (6)$$

$$\{p_{H_s}\}_3 = \beta_1 + \beta_2x + \beta_3y + \beta_4z + \beta_5x^2 + \beta_6y^2 + \beta_7z^2 + \beta_8xy + \beta_9xz + \beta_{10}yz + \beta_{11}x^3 + \beta_{12}y^3 + \beta_{13}z^3 + \beta_{14}x^2y + \beta_{15}x^2z + \beta_{16}xy^2 + \beta_{17}xz^2 + \beta_{18}y^2z + \beta_{19}yz^2 + \beta_{20}xyz \quad (7)$$

where  $p_{H_s}$  is the predicted value for the significant wave height, which is dependent on the observed values (x, y, z) of significant wave height for the whole forty year period, namely the  $H_s$  timeseries of the three boundaries from ERA5 dataset, and  $\beta_n$  are the coefficients of each model. The prediction of  $H_s$  for the whole period was calculated using two out of three years from MIKE results, keeping the third year as a means of estimating the errors between the observed and the estimated data. Among these three different order models, it was observed that the second order model could perform better in terms of having the least deviation from the observed time series. The first order model performed well but not enough to be selected. As for the third order model, while it was expected to perform better than the second order model due to the inclusion of twice as many coefficients, the results showed that there were higher discrepancies between the observed and the estimated data, meaning it was overestimating  $H_s$ . This could be possibly explained due to over fitting.

To continue, the predicted (alongside with the observed) three-year (1980-1982)  $H_s$  timeseries of the five distinct nearshore target points are illustrated in Fig. 76-80

including the first and second order regression models as a means of comparison between them. The results suggest that in almost all target points the first order model can describe the wave regime by means of  $H_s$  adequately although it slightly underestimates it, but the second order model lessens the differences between observed and predicted values. Interestingly, in TP4 (Fig. 79) it is evident that first and second order model results do not differ substantially. It is worth noting that TP4 has the lowest significant wave height values of all target points.

In addition, the root mean square error (RMSE) was calculated for the first and second order regression model results for each target point and they are illustrated in the table below (Table 6). It is evident that the improvement in RMSE after employing the 2nd order regression model is not high, but rather low, even on the edge of not being statistically significant.

	TP1	TP2	TP3	TP4	TP5
1st order	0.21	0.16	0.13	0.10	0.14
2nd order	0.19	0.15	0.12	0.09	0.13

Table 6 RMSE of first and second order regression for all target points.

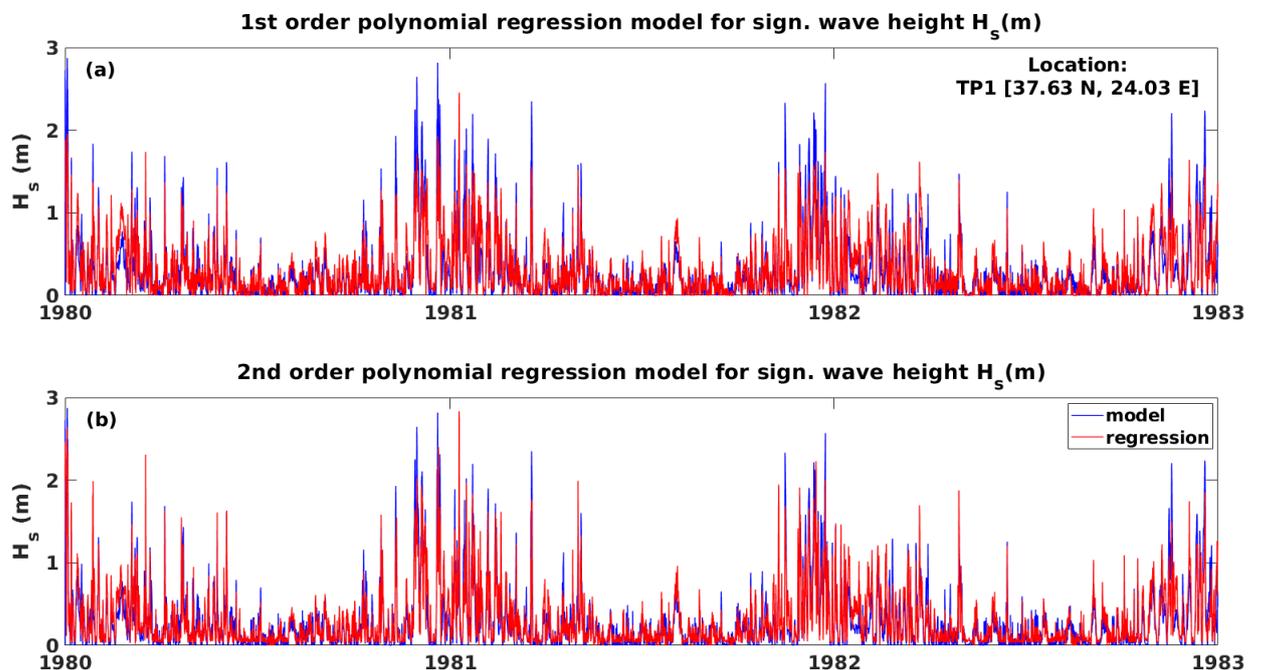


Figure 76 Multivariate regression of  $H_s$  for TP1: (a) 1st order, (b) 2nd order. Blue line is MIKE results, red line is the product of regression.

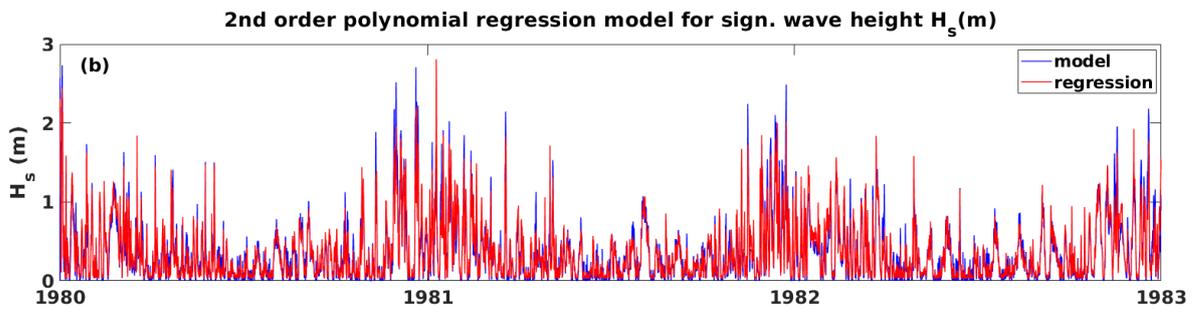
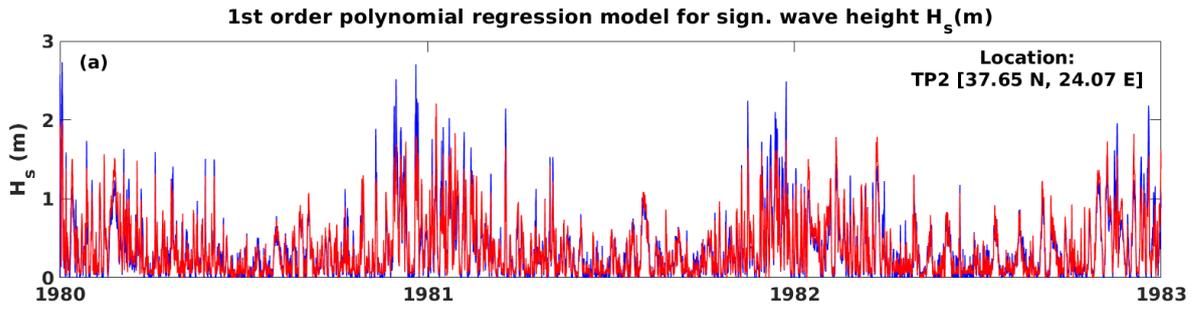


Figure 77 Multivariate regression of  $H_s$  for TP2: (a) 1st order, (b) 2nd order . Blue line is MIKE results, red line is the product of regression.

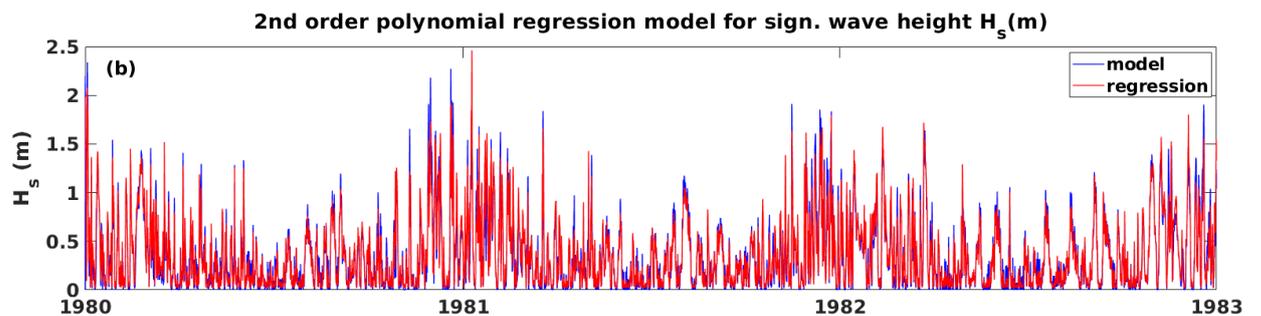
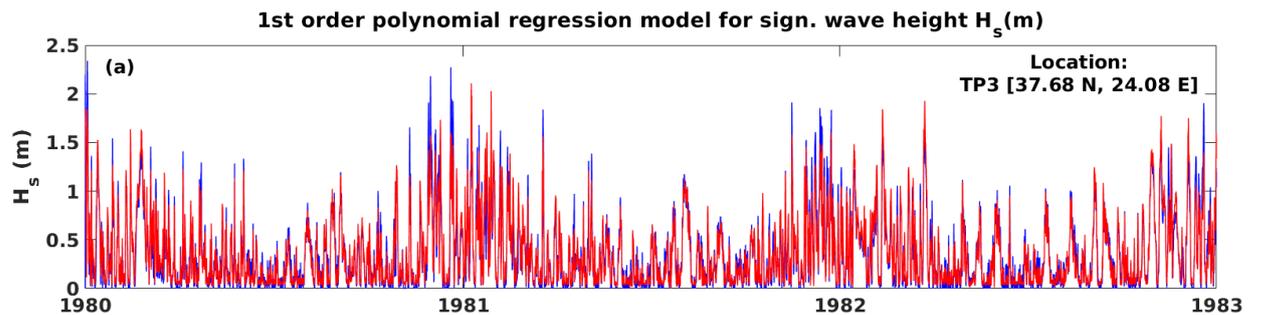


Figure 78 Multivariate regression of  $H_s$  for TP3: (a) 1st order, (b) 2nd order. Blue line is MIKE results, red line is the product of regression.

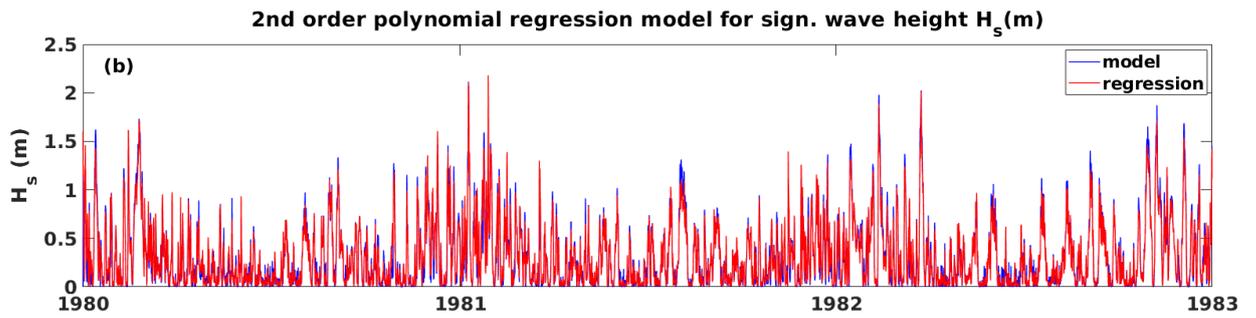
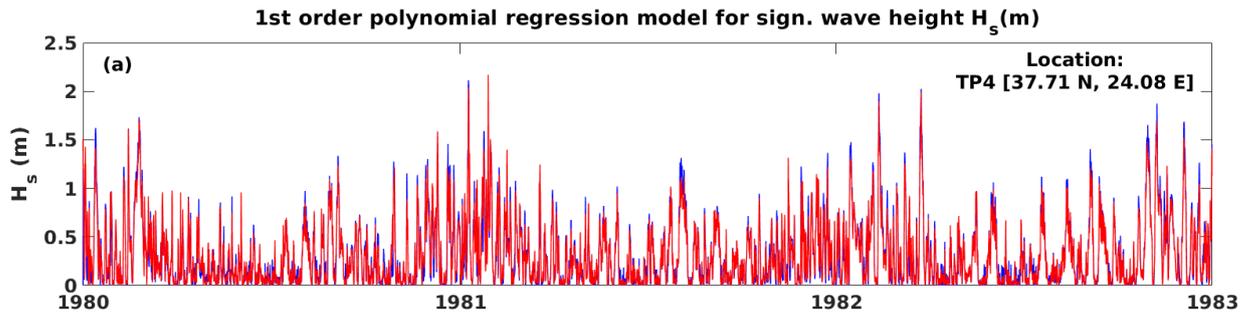


Figure 79 Multivariate regression of  $H_s$  for TP4: (a) 1st order, (b) 2nd order. Blue line is MIKE results, red line is the product of regression.

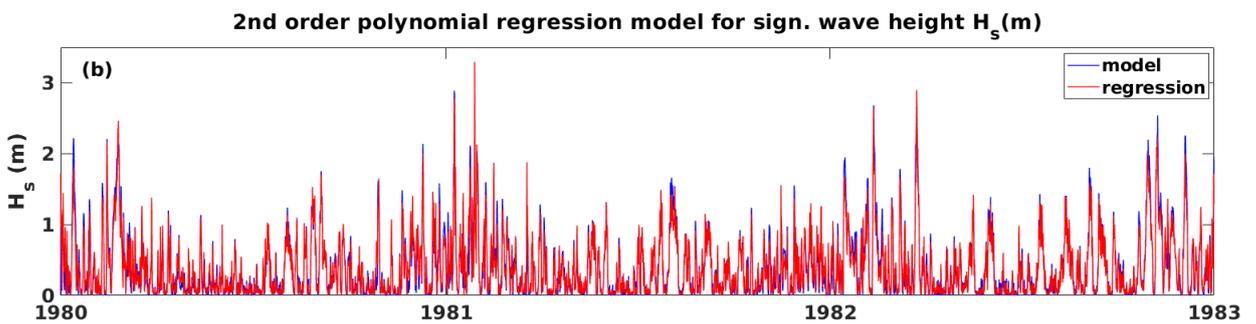
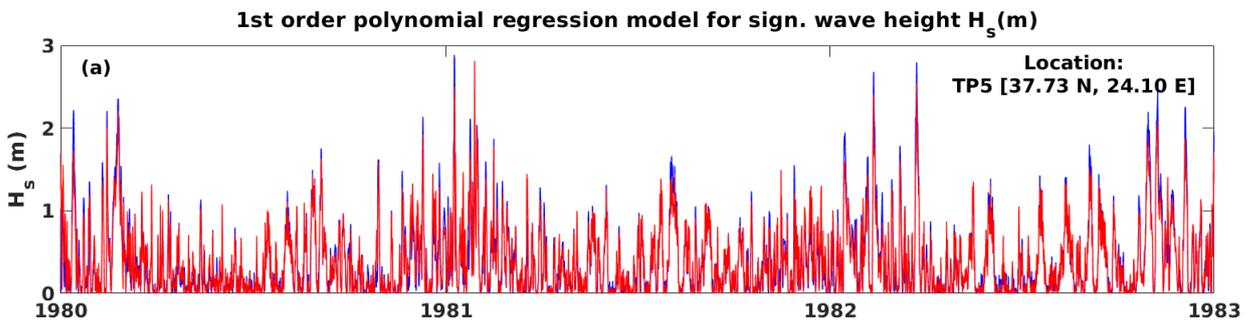


Figure 80 Multivariate regression of  $H_s$  for TP5: (a) 1st order, (b) 2nd order. Blue line is MIKE results, red line is the product of regression.

## 5.5. Extreme value analysis using the POT method

Finally, extreme value analysis was applied at the extended timeseries of the nearshore target points as well as the offshore boundary points using the POT method which is described in detail in Appendix C. The threshold for the exceedances was selected in accordance with the maximum values of  $H_s$  for each location in a manner that facilitates the estimation of the Generalized Pareto (GP) cumulative distribution function (cdf) parameters  $\xi$  and  $\sigma_u$ . Furthermore, the parameters of the GP cdf were estimated using the Maximum Likelihood method (MLM).

The results of the analysis are illustrated in Fig. 81-96, including the GP cdf plot along with the distribution parameters and design values for each data point. The design values include the return level  $x_T$  and return period  $T$  of an event. The return level is a value that is expected to become equivalent or exceeded once every  $T$  years with a probability  $\frac{1}{T}$ . The return level  $x_T$  can be estimated via a distribution function  $F$  (in this case the GP cdf) where  $x_T$  is the  $1 - \frac{1}{T}$  quantile of  $F$  (Kapelonis et al., 2015).

To continue, the chosen threshold and GP parameters for all wave data points are shown in Table 7 and the significant wave heights which correspond to return periods of 10, 20 and 30 years are presented in Table 8.

	O1	O2	O3	TP1	TP2	TP3	TP4	TP5
<b>u (m)</b>	2.2	2.2	2.2	2.2	2.2	2.2	2.2	2.2
<b><math>\xi</math></b>	-0.18	-0.11	-0.18	0.29	0.02	-0.66	-0.22	-0.22
<b><math>\sigma</math></b>	0.39	0.54	0.5	0.21	0.22	0.32	0.09	0.43

Table 7 POT threshold and GP parameters

$H_s^{(T)}$ (m)	O1	O2	O3	TP1	TP2	TP3	TP4	TP5
<b>10</b>	0.74	1.1	0.94	0.68	0.53	0.38	0.17	0.78
<b>20</b>	0.9	1.38	1.16	0.99	0.69	0.42	0.20	0.95
<b>30</b>	0.99	1.53	1.27	1.2	0.79	0.44	0.22	1.04

Table 8 Significant wave height (m) for return periods of 10, 20 and 30 years

According to the results, on the one side, the highest significant wave height for each return period belongs to O2 with  $H_s^{(10)} = 1.1$  m,  $H_s^{(20)} = 1.38$  m and  $H_s^{(30)}$  approximately 1.53 m. On the other side the lowest significant wave height for all selected return periods is observed at TP4 with  $H_s^{(10)} = 0.17$  m,  $H_s^{(20)} = 0.20$  m and  $H_s^{(30)} = 0.22$  m. Among the nearshore points TP1 exhibits the highest  $H_s^{(T)}$  values with  $H_s^{(10)} = 0.68$  m,  $H_s^{(20)} = 0.99$  m and  $H_s^{(30)} = 1.2$  m along with TP5 which correspondingly has  $H_s^{(10)} = 0.78$  m,  $H_s^{(20)} = 0.95$  m and  $H_s^{(30)} = 1.04$  m.

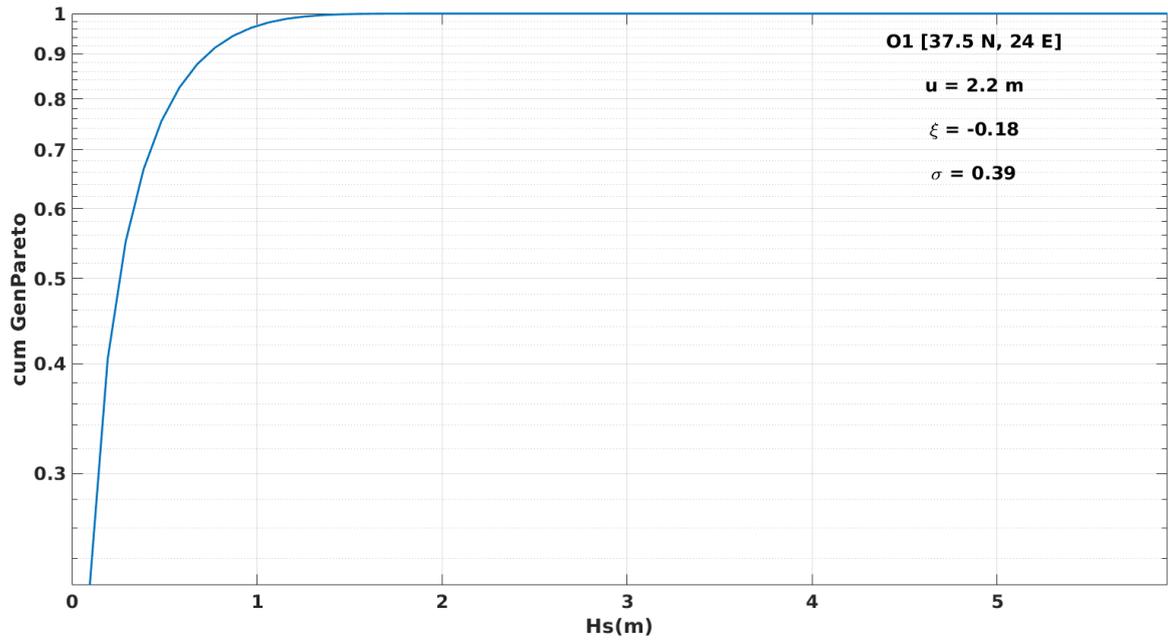


Figure 81 Cumulative Generalized Pareto distribution for Hs at O1

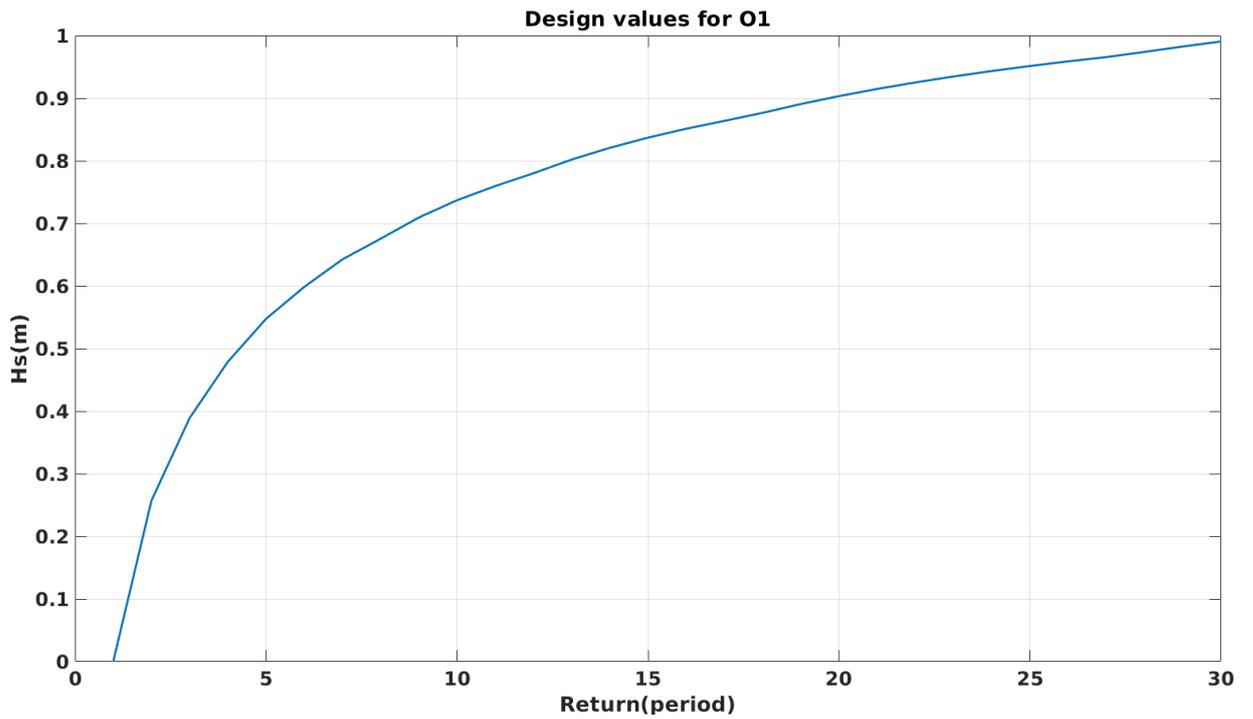


Figure 82 Design values (return period) for O1

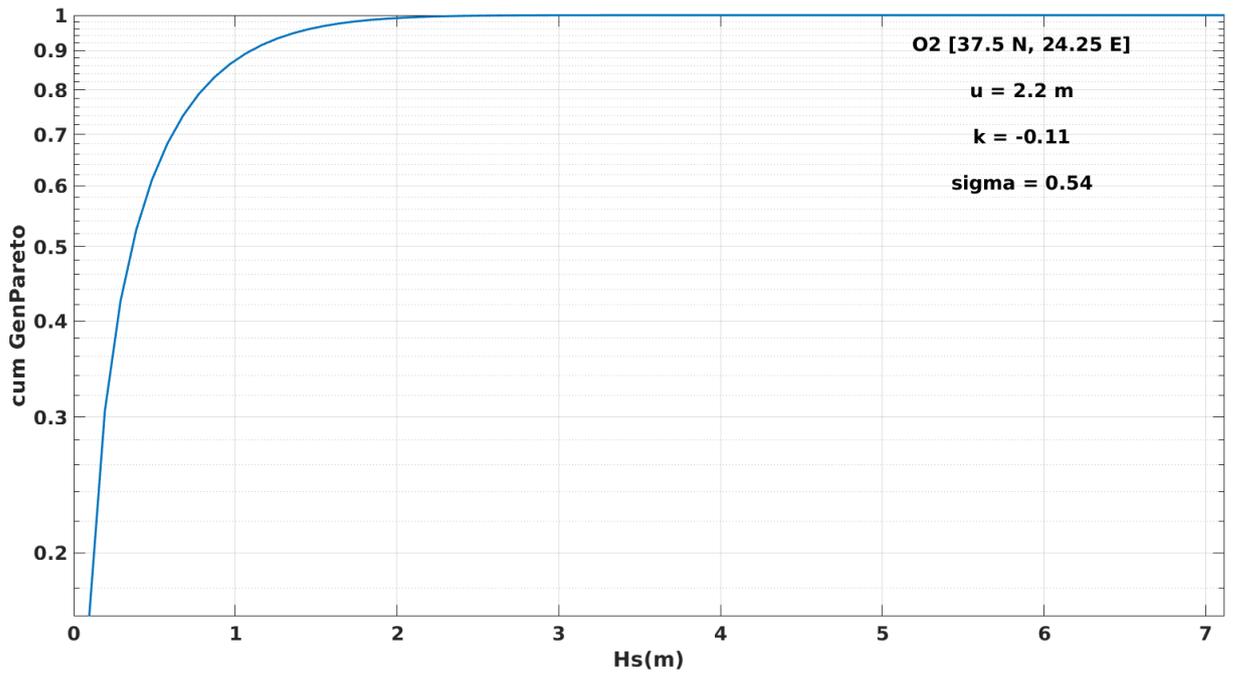


Figure 83 Cumulative Generalized Pareto distribution for Hs at O2

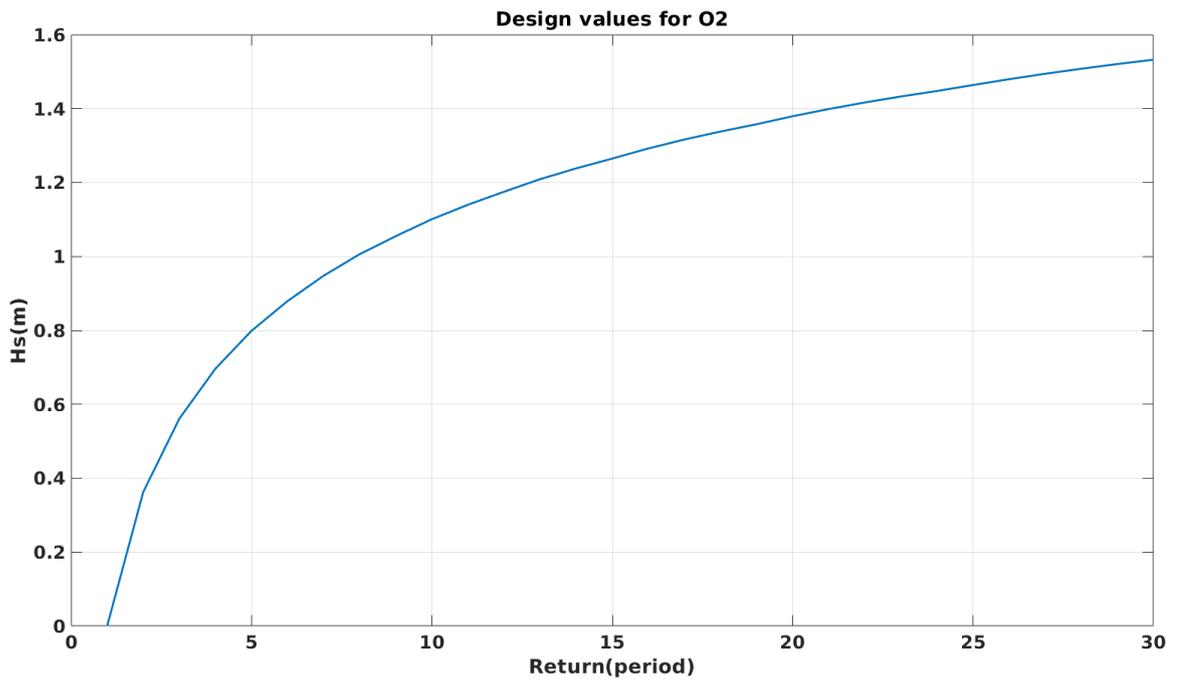


Figure 84 Design values (return period) for O2

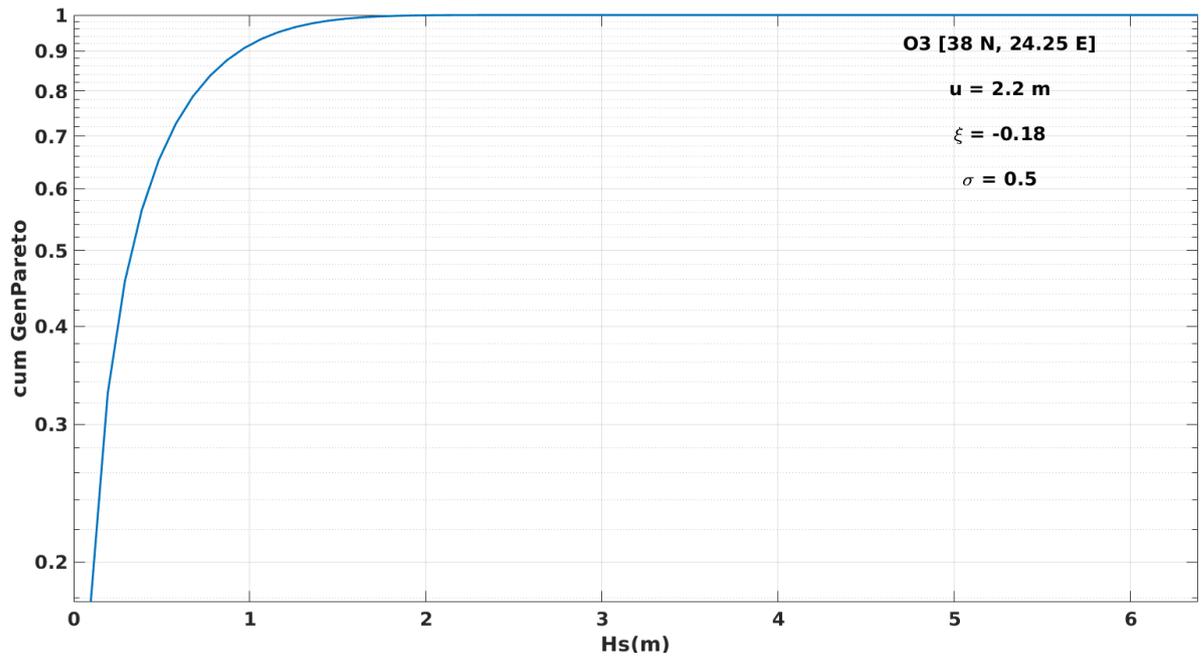


Figure 85 Cumulative Generalized Pareto distribution for Hs at O3

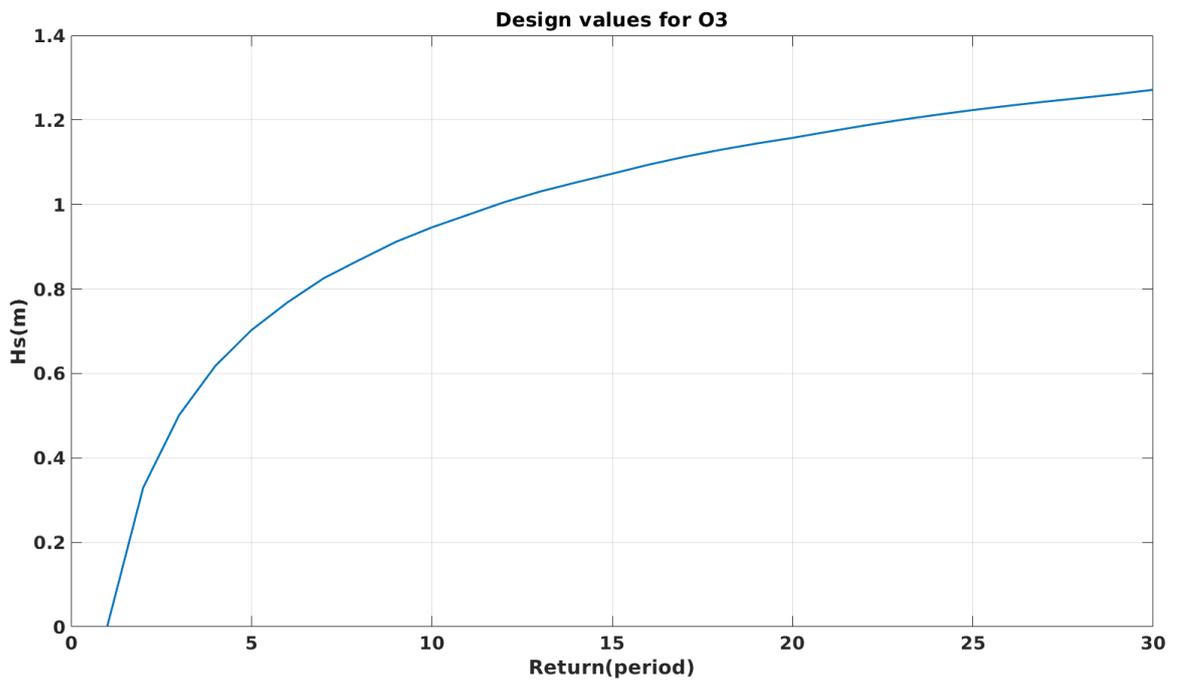


Figure 86 Design values (return period) for O3

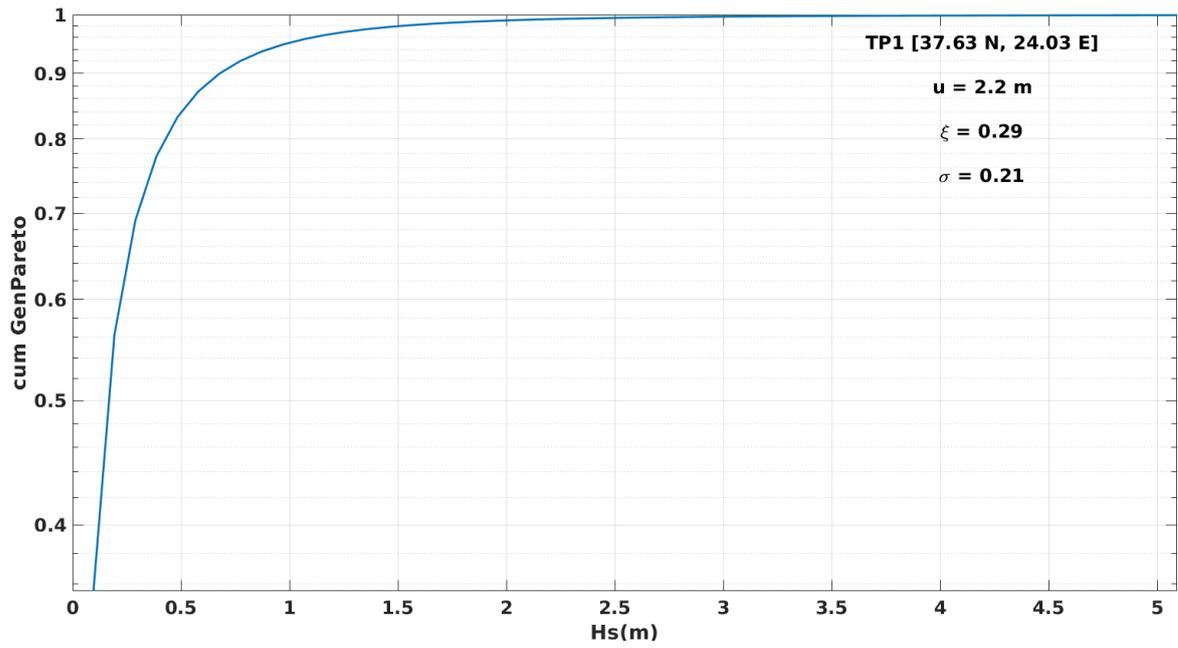


Figure 87 Cumulative Generalized Pareto distribution for Hs at TP1

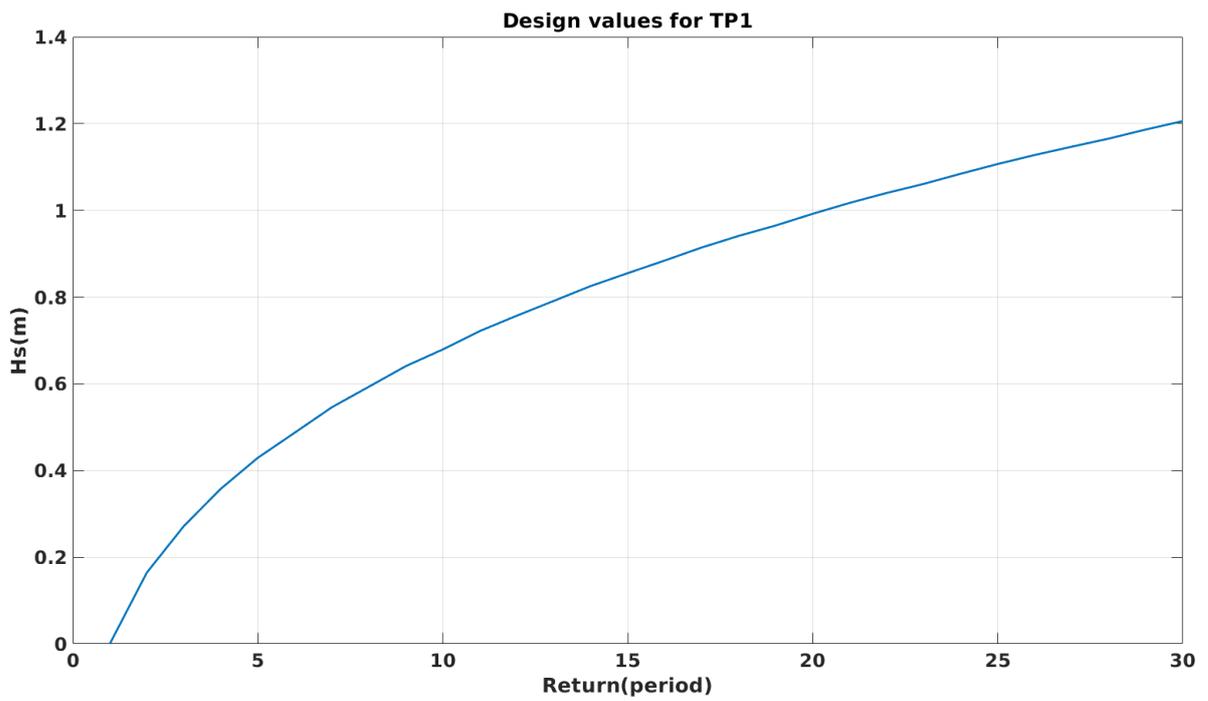


Figure 88 Design values (return period) for TP1

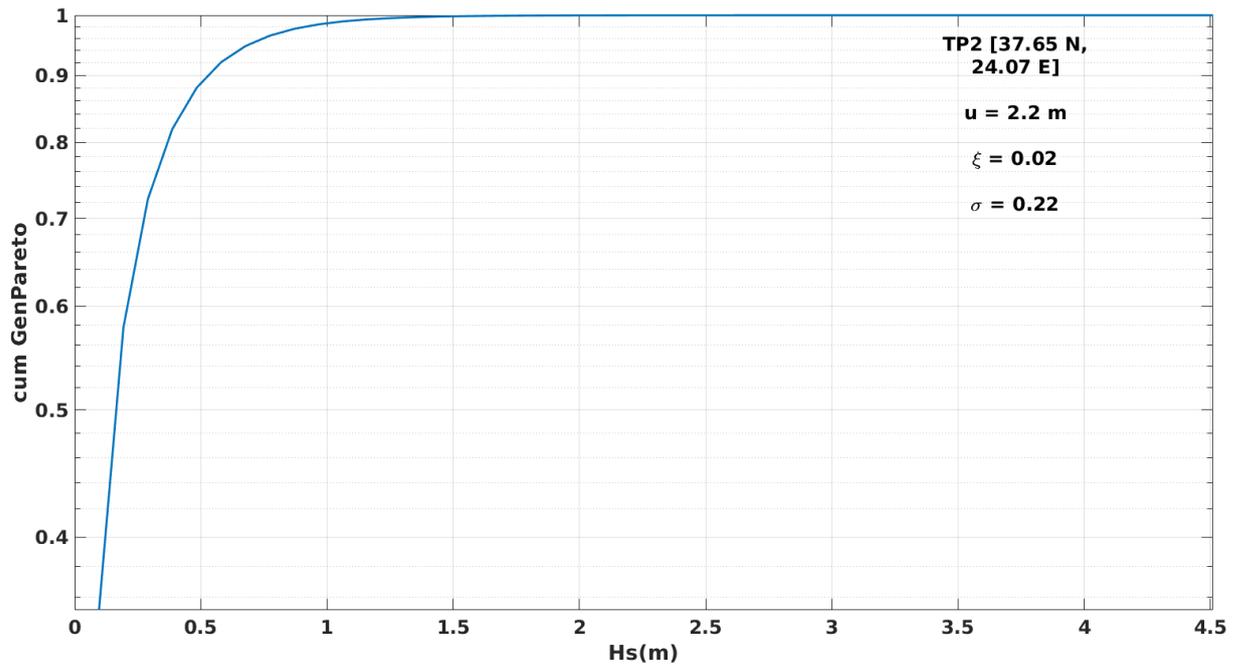


Figure 89 Cumulative Generalized Pareto distribution for Hs at TP2

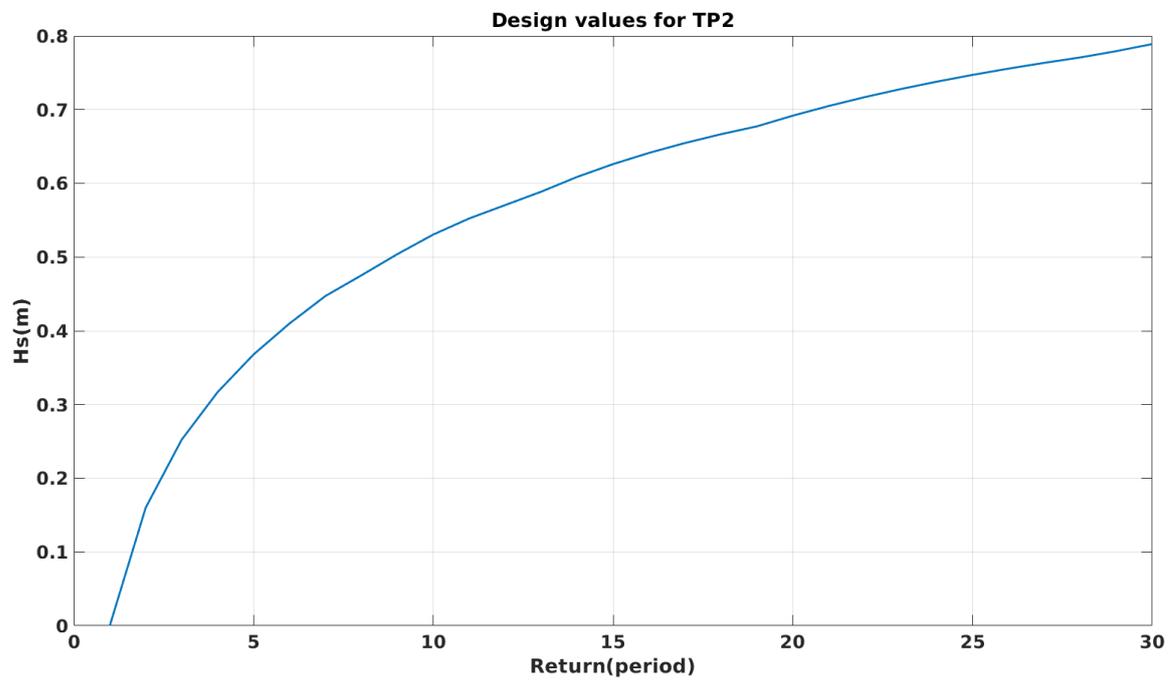


Figure 90 Design values (return period) for TP2

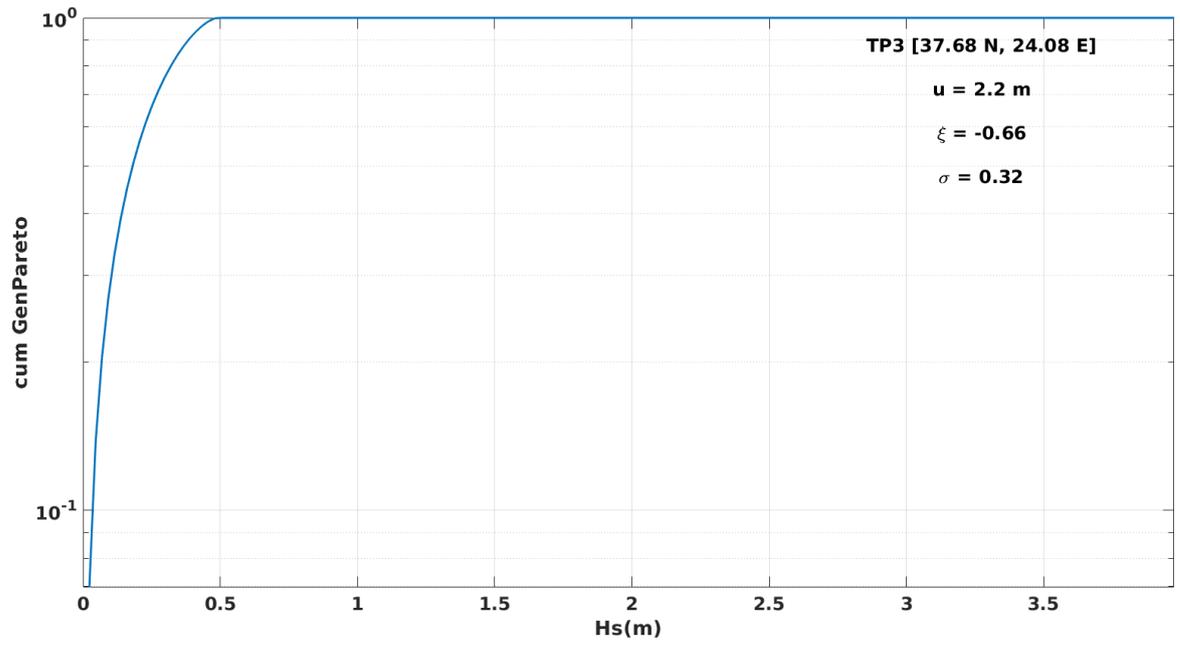


Figure 91 Cumulative Generalized Pareto distribution for Hs at TP3

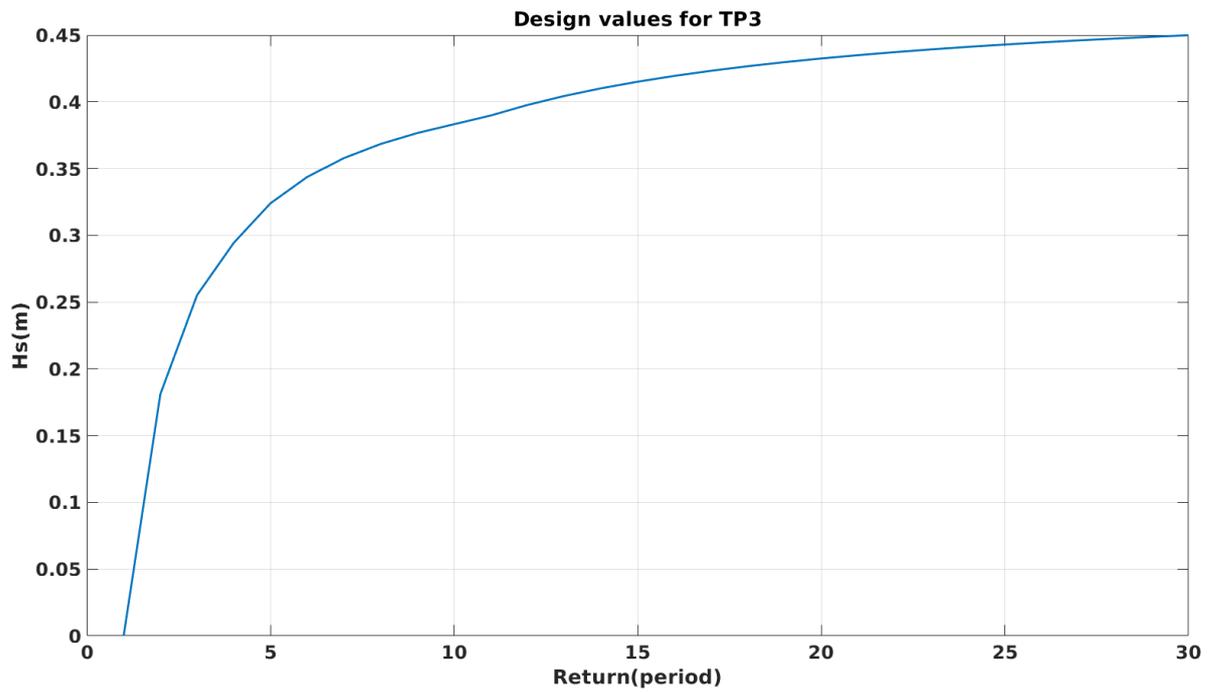


Figure 92 Design values (return period) for TP3

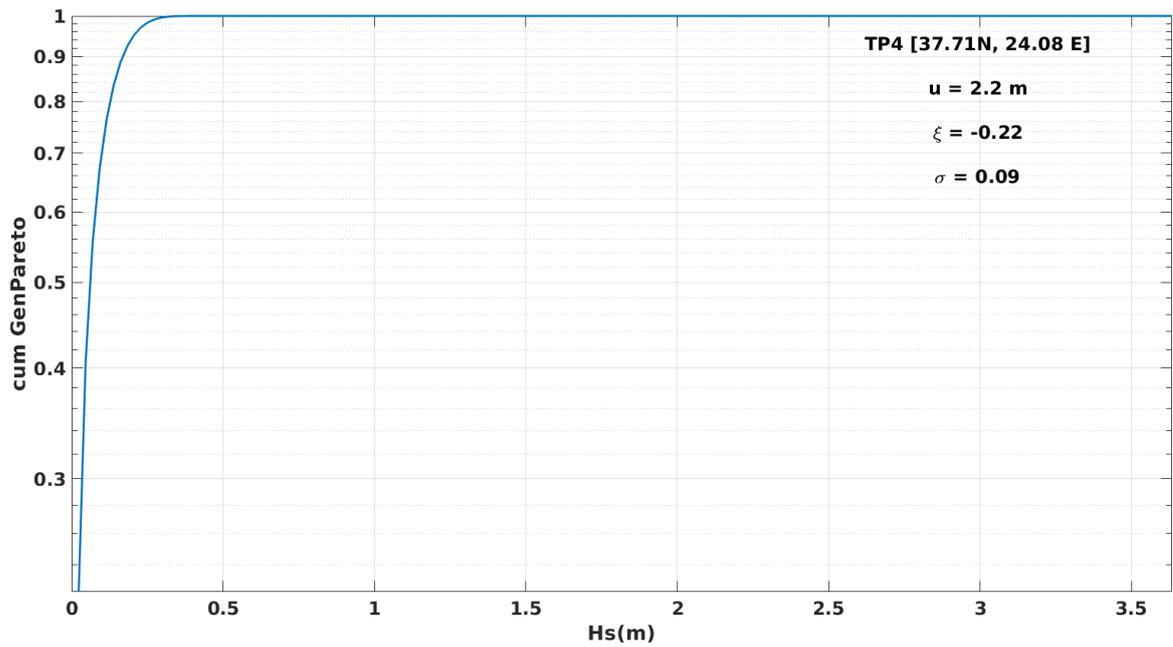


Figure 93 Cumulative Generalized Pareto distribution for Hs at TP4

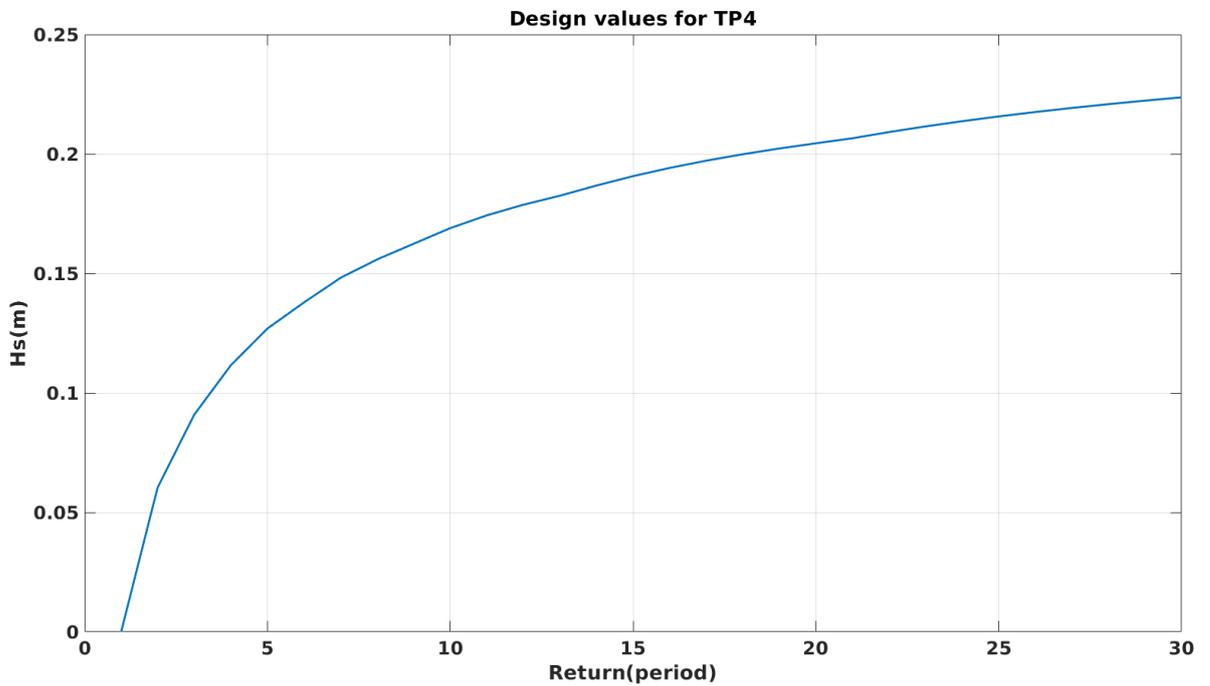


Figure 94 Design values (return period) for TP4

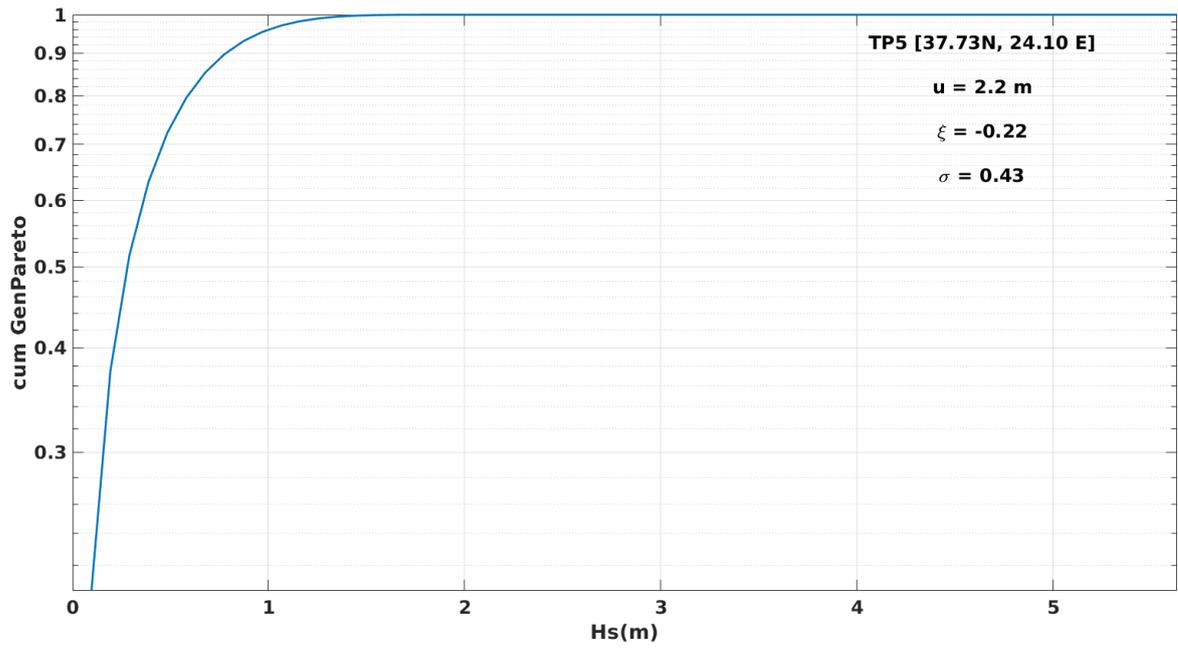


Figure 95 Cumulative Generalized Pareto distribution for Hs at TP5

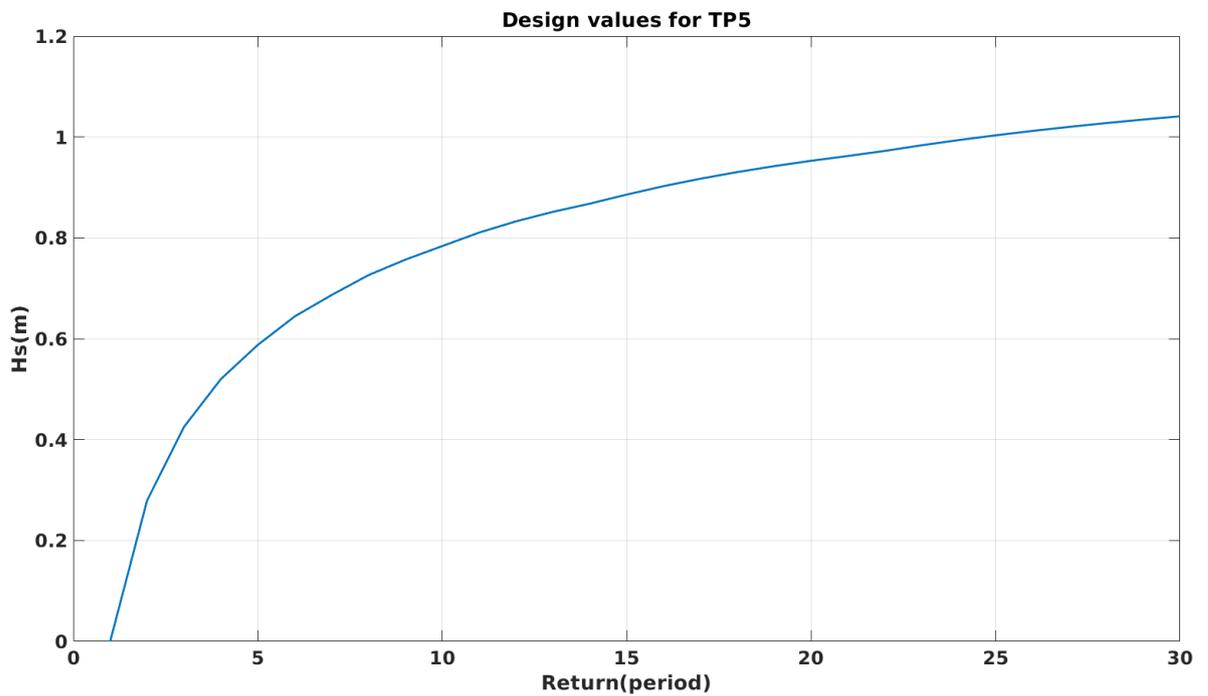


Figure 96 Design values (return period) for TP5

## 6. Discussion and conclusions

To summarize, the main purpose of this thesis is the simulation of the wave regime at Makronisos channel for the period of 1980-2019 in order to assess it with statistical methods and conduct an extreme value analysis. The modeling tool for this study is MIKE 21 Spectral Waves Model, transforming the offshore wave conditions to selected nearshore locations.

The setup of the model contained wind and wave data from ERA5 reanalysis database which served as atmospheric and lateral boundary conditions. The computational mesh, which was varying in resolution, was based on EMODnet bathymetric data and a digitized coastline created from a naval map of HNHS. In order to ensure that the generated mesh functions properly, eight stationary runs were performed using boundary conditions from Medatlas Climatology (Wind and Wave Atlas of the Mediterranean Sea, 2004). As for the hindcast runs, the overall computational cost was high, approximately 104 hours per year's run, given that it was conducted serially in a personal computer. Consequently, the simulation time for the three years produced was 312 hours (13 days). This implies that it would be impossible to complete in time the desired forty year period by using the same resources. Although MIKE 21 offers the ability for parallel runs, this was not possible in this case. Given the above information, three years (1980-1982) out of the expected 40 year period were effectively produced and as a consequence the usage of a multivariate regression model was required so as to extend the results of the significant wave height. Three different models were examined, first, second and third order, from which the latter was excluded due to the overestimation of  $H_s$ . RMSE for the first and second order model showed that there is not quite a significant difference between them, nevertheless the second order model was found to be more fit for the intended purpose.

Furthermore, the assessment of the wave climate both for the offshore and nearshore data was accomplished with the probabilistic description of wave parameters ( $H_s, T_p, \vartheta$ ), implementing Kernel Density Models (univariate and bivariate) thus making an estimation of the parameter's probability distributions. The statistic results for the offshore point data suggest that the wave conditions at the boundaries of the study area are quite mild with more than 50% probability of the significant wave height being between the interval of [0, 0.5] m. The same applies for the nearshore points as they also exhibited quite small wave heights with more than 50% probability of  $H_s$  between [0, 0.5] m. Concerning peak wave period the interval of [4, 6] sec was observed to be the most probable to occur for offshore and nearshore points (except TP1 where [3, 5] sec was the most probable  $T_p$  interval). The governing mean wave direction is 0°-30° N-NE for the offshore points and 15°-30° N-NE for the nearshore points. Additionally at O1, TP1, TP4 and TP5 a secondary

governing mean wave direction ranges from 180°-230° S-SE. The most energetic offshore point in terms of the maximum value of significant wave height observed is O2 (max  $H_s = 0.69$  m), and the most energetic nearshore point is TP5 (max  $H_s = 0.39$  m). A pattern that is observed from the results timeseries regarding the seasonality of significant wave height is that it is higher during winter and early spring and lower during the summer months. Additionally, after reviewing the timeseries of offshore and nearshore points, a correlation between peaks of  $H_s$  with peaks of  $T_p$  and a wave direction coming from the north appears, which would deserve a more in depth analysis so as to be verified.

To continue, the POT method was used in order to assess the extreme events of the significant wave height at offshore and nearshore points, estimating the parameters of the cumulative Generalized Pareto distribution and the significant wave heights corresponding to the return levels of 10, 20 and 30 years. At this point it should be noted that although it is suggested to apply some declustering methodology at the timeseries under examination before conducting the POT method, in this study it was omitted given that the regression already introduced a level of uncertainty to the results. As observed from the results, O2 which is the southeastern corner of the study area, exhibits the highest significant wave heights for all return periods ( $H_s^{(10)} = 1.1$  m,  $H_s^{(20)} = 1.38$  m and  $H_s^{(30)} = 1.53$  m). On the contrary TP4 which is a bit north from Lavrio port seems to be the most protected in terms of high waves since it presented the lowest significant wave heights for all return periods ( $H_s^{(10)} = 0.17$  m,  $H_s^{(20)} = 0.2$  m and  $H_s^{(30)} = 0.22$  m). Between the nearshore points the highest return values are presented at TP1 ( $H_s^{(30)} = 1.2$  m) and TP5 ( $H_s^{(30)} = 1.04$  m), which are situated at the southern and northern entrance of Makronisos channel respectively.

As a general remark, the absence of in situ data did not allow the calibration of the model and the validation of the produced data for the purpose of reducing the level of uncertainty. Additionally, the usage of ERA5 wave model data for boundary conditions magnified the uncertainty because its resolution is 0.5° (approximately 55.5 km) which is the same as the length of both the southern and eastern boundaries of the computational grid. This is a potential underspecification problem and a possible solution would be to make a downscaling effort by producing wave data results for a wider area and then applying the results as boundary conditions for a new simulation at the desired area. As for the method of estimation of the GP parameters, it would be advisable to apply the PWM method since according to Caires (2016) for a data sample less than 50 years PWM is superior to ML method in terms of robustness and error characteristics that relate to non-existence of solutions. By extension, it would be worth to further assess the threshold selection depending on the peculiarities of each data point as it directly impacts the return values. As a last proposal towards future research, it would be of great interest to

simulate and examine wave-current interaction given the unique characteristic of Makronisos channel, provided that there are available current data in or near the channel. Furthermore, in the case where wave predictions existed, it would be interesting to compare the estimated return values with the predictions and lastly, the MED Waves (Ravdas et al., 2018) data from Copernicus Marine Environment Monitoring Service (CMEMS) could be used to compare with MIKE 21 SW results, considering the availability of information close to one nearshore point (*e.g* TP1).

## Appendix A - Probability distributions of linear and circular variables

One part of this thesis is focused on the statistical analysis of wave climatology at the study area during a specific period. The wave parameters included in this particular approach are the significant wave height  $H_s$ , peak period  $T_p$ , which are both linear variables, and mean wave direction  $\vartheta$  which is a circular variable. The analysis was carried out with a nonparametric method, estimating the kernel density function (kdf) for both univariate and bivariate distributions. The kernel distribution then is the representation of the kdf. The advantage of nonparametric methods is that they do not depend on parametric assumptions since the probability density function (pdf) is estimated in accordance with the information contained in the data sample (Karathanasi, 2020). The most frequent scenarios on the basis of which the nonparametric methods are preferred are:

(a) a small sample size, (b) when there is no knowledge concerning the pdf of the data or when it is known that it follows a normal distribution, and (c) the presence of outliers in the data.

The kdf which were introduced by Fix and Hodges (1951) have been widely used in data analysis (Athanasoulis & Belibassakis, 2002 ; Karathanasi, 2020).

Let  $(Y_1, Y_2, \dots, Y_N)$  be independent and identically distributed sample (univariate) from a population, then the unknown pdf  $f(x)$  can be estimated from (Athanasoulis & Belibassakis, 2002):

$$f(x) = \frac{1}{N} \sum_{n=1}^N K(x; Y_n, h) \quad (A.1)$$

where  $K(x; Y_n, h)$ ,  $n=1, 2, \dots, N$  are the kdfs which are smooth and non-negative functions and  $h$  is the bandwidth parameter which controls the smoothness of the density curve. Furthermore, a generalized univariate model that can estimate and serve as the analytic representation of the pdf of any meteorological or oceanographic parameter, as described by Athanasoulis & Belibassakis (2002) is

$$f(x) = \sum_{i=1}^I v_i K(x; y_i, h_i) \quad (A.2)$$

where  $v_i$  are non-negative constants and  $h_i$  the local bandwidth parameters.

As mentioned in Karathanasi (2020), the conditions that need to be satisfied in order for a function to be utilized as a kernel are:

- a.  $K(x; y, h) \geq 0 \forall x$ , is a non-negative function
- b.  $\int K(x; y, h)dx = 1$
- c.  $\int xK(x; y, h)dx = 0$  because of symmetry
- d.  $0 < \int x^2 K(x; y, h)dx < \infty$ , is of second order

There are several types of kernel functions commonly used e.g. Uniform, Triangle, Epanechnikov, Biweight, Gaussian, Gamma etc. In this study and for the univariate distribution of linear variables the Gamma type was employed, which is described in the following paragraphs.

As for circular variables, like mean wave direction or wind direction, it is essential to use a 360°periodic kdf (Athanasoulis & Belibassakis, 2002). An example of the latter is the Wrapped-Gaussian distribution (Mardia, 1972):

$$K_{WG}(x; y, h) = \frac{1}{\sqrt{2\pi}h} \sum_{k=-\infty}^{\infty} \exp \left[ -\frac{1}{2} \left( \frac{x - y - 360k}{h} \right)^2 \right] \quad (A.3)$$

where  $y$  is the positioning parameter,  $h$  is the bandwidth and  $k$  is the central term. According to Athanasoulis & Belibassakis (2002), the Wrapped-Gaussian distribution can be accurately estimated from the three central terms of the sum ( $k=-1, 0, 1$ ) when  $h < 90^\circ$ .

Consequently, the density for the Wrapped-Gaussian distribution is expressed as follows:

$$f_{KWG}(x; h) = \frac{1}{nh} \sum_{i=1}^n K_{WN}(x; y, h) \quad (A.4)$$

In addition to the univariate approach, the bivariate probability distributions appear as a more appropriate method in the field of ocean engineering due to the fact that the main parameters such as waves and winds are overall non-independent variables (Karathanasi, 2020). In such a case, the corresponding kdf is the extension of eq. (A.2):

$$f(x) = \sum_{i=1}^I \sum_{j=1}^J v_{ij} K(x; y_{ij}, h_{ij}) \quad (A.5)$$

where  $x = (x_1, x_2), y_{ij} = (y_{ij_1}, y_{ij_2})$   $i = 1, 2, \dots, I, j = 1, 2, \dots, J$ , is a set of values spreading on the support of  $f(x)$  and  $h_{ij}$  are local bandwidth matrices (Athanasoulis & Belibassakis, 2002). The skewed kernel that was used in the bivariate distribution is the Gamma kernel:

$$K_{\Gamma}(x; y, h) = \frac{x^{\lambda-1} e^{\left(\frac{-x}{\mu}\right)}}{\mu^{\lambda} \Gamma(\lambda)} \quad (A.6)$$

where  $\lambda(y, h) > 1$  and  $\mu(y, h) > 0$  are the shape and scale parameters (Athanasoulis & Belibassakis, 2002), and  $\Gamma$  is the Gamma function which can be expressed as  $\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} dx$ . Both aforementioned parameters  $\lambda$  and  $\mu$  can be linked with the location parameter  $y$  and bandwidth  $h$  via the following equations:

$$\lambda = \frac{1}{2} \left[ 2 + \left(\frac{y}{h}\right)^2 + \sqrt{\left(2 + \left(\frac{y}{h}\right)^2\right)^2 - 4} \right] \quad (A.7)$$

and

$$\mu = \frac{y}{\lambda - 1}. \quad (A.8)$$

## Appendix B - Regression analysis

The forecast and/or simulation of a physical phenomenon like the characteristics of wave climate is a crucial part of coastal structure studies and by extension of the offshore oil/gas industry or even wind farm installations given that it provides key information about possible environmental loads, as well as it is crucial to maritime transport in terms of voyage safety. However, such endeavor requires significant computational resources for its completion. In the case of insufficient resources, regression modeling methods can provide an efficient solution to the problem. In the present thesis, due to computational power limitations, multivariate regression was used to model the evolution of significant wave height  $H_s$  for the 37 year period that remained using the 3 year results obtained from MIKE 21 SW simulations.

The evolution of wave characteristics ( $H_s$ ,  $T_p$  and  $\vartheta$ ) in time and space is a phenomenon governed by specific equations and dependent on various conditions. With the help of regression analysis, and specifically multivariate regression, it is possible to discover the relationship between the wave parameter outcome (whether  $H_s$ ,  $T_p$  or  $\vartheta$ ) and its predictor variables, so as to predict the development of the desired parameter.

The procedure that constructs the regression model consists of (a) assuming models on observed data thought to affect the phenomenon, (b) the estimation of the model parameters and (c) the evaluation of the estimated models in order to select the optimal model (Konishi, 2014).

### B.1. Multivariate polynomial regression

Multivariate polynomial regression is a special case of multiple linear regression. Essentially it refers to the development of a non linear relationship between a result (dependent variable  $y$ ) and the parameters that affect it (independent variables  $x_i$ ). However linearity exists in the sense that the regression function  $E(y|x)$  is linear in the unknown coefficients that are estimated from the data.

The equation of multivariate polynomial regression of m-th order is

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_m x_i^m + \varepsilon_i (i = 1, 2, \dots, n) \quad (B.1)$$

which can be also written in matrix form corresponding to a response vector  $\mathbf{y}$ , a coefficient vector  $\boldsymbol{\beta}$ , a design matrix  $X$ , and a vector of random errors  $\boldsymbol{\varepsilon}$ .

More specifically, it can be viewed as a system of linear equations:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^m \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

and as a matrix notation

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (\text{B. 2})$$

Two methods that are usually employed to estimate regression models are least squares and maximum likelihood. In the first method the best regression model is that which minimizes the sum of squared errors between the observed data and the model. The advantage of least squares method is that the error terms  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  are uncorrelated  $E[\varepsilon_i \varepsilon_j] = 0$  ( $i \neq j$ ), the mean  $E[\varepsilon_i \varepsilon_j] = 0$  and variance  $E[\varepsilon_i^2] = \sigma^2$ . The vector of least squares regression coefficient  $\boldsymbol{\beta}$  can be estimated from (Konishi, 2014).

$$S(\boldsymbol{\beta}) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{y} - X\boldsymbol{\beta})^T (\mathbf{y} - X\boldsymbol{\beta}) \quad (\text{B. 3})$$

$$S(\boldsymbol{\beta}) = \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T X\boldsymbol{\beta} + \boldsymbol{\beta}^T X^T X\boldsymbol{\beta} \quad (\text{B. 4})$$

In order to proceed, we differentiate both sides of eq. (B.4) with respect to  $\boldsymbol{\beta}$  which yields

$$\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} (\mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T X\boldsymbol{\beta} + \boldsymbol{\beta}^T X^T X\boldsymbol{\beta}) = -2X^T \mathbf{y} + 2X^T X\boldsymbol{\beta} = 0 \quad (\text{B. 5})$$

The least squares coefficient  $\boldsymbol{\beta}$  can be therefore estimated solving eq. (B.5) given that the inverse of  $X^T X$  matrix exists, leading to the following expression:

$$\boldsymbol{\beta} = (X^T X)^{-1} X^T \mathbf{y} \quad (\text{B. 6})$$

which is the ordinary least squares estimator function that minimizes the sum of squares of residuals and according to Gauss-Markov theorem is the best linear unbiased estimator (BLUE).

In the second method, maximum likelihood assumes that the data for the response variable have a normal distribution, and that the best fitting model is one that maximizes the likelihood function. In this case, besides from  $E[\varepsilon_i \varepsilon_j] = 0$  ( $i \neq j$ ) and  $E[\varepsilon_i^2] = \sigma^2$ , the components of the error vector  $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$  are mutually independent and distributed according to a normal distribution  $N(0, \sigma^2)$  (Konishi, 2014). In this way, the probability density function of the error  $\boldsymbol{\varepsilon}$  is

$$f(\boldsymbol{\varepsilon}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}}{2\sigma^2}\right) \quad (B.7)$$

The likelihood function is

$$L(\boldsymbol{\beta}, \sigma^2) = f(\mathbf{y}|X; \boldsymbol{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2} (\mathbf{y} - X\boldsymbol{\beta})^T (\mathbf{y} - X\boldsymbol{\beta})\right] \quad (B.8)$$

and consequently the log-likelihood function is

$$\ell(\boldsymbol{\beta}, \sigma^2) = \log L(\boldsymbol{\beta}, \sigma^2) = -\frac{n}{2} \log(2\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{y} - X\boldsymbol{\beta})^T (\mathbf{y} - X\boldsymbol{\beta}) \quad (B.9)$$

The solution to the likelihood equations which maximizes (B.9) is

$$\frac{\partial \ell(\boldsymbol{\beta}, \sigma^2)}{\partial \boldsymbol{\beta}} = \frac{1}{\sigma^2} (X^T \mathbf{y} - X^T X \boldsymbol{\beta}) \quad (B.10)$$

$$\frac{\partial \ell(\boldsymbol{\beta}, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} (\mathbf{y} - X\boldsymbol{\beta})^T (\mathbf{y} - X\boldsymbol{\beta}) = 0 \quad (B.11)$$

Finally, the maximum likelihood estimates for  $\boldsymbol{\beta}$  and  $\sigma^2$  are given from the following expressions:

$$\boldsymbol{\beta} = (X^T X)^{-1} X^T \mathbf{y}, \quad \widehat{\sigma^2} = \frac{1}{n} (\mathbf{y} - X\boldsymbol{\beta})^T (\mathbf{y} - X\boldsymbol{\beta}) \quad (B.12)$$

## Appendix C - Extreme value analysis methods

Extreme value analysis is a tool of great importance for ocean-coastal engineering applications considering the impact of extreme metocean conditions on coastal and/or offshore structures. The most common approach to describe the severity of sea state is the examination of significant wave height  $H_s$ , which is defined as the average wave height of the highest one-third of the waves. Additionally, a reliable extreme value analysis can be performed providing that quality long time series (at least 30 years) of  $H_s$  are available.

According to Ochi (1998), order statistics, which are a very important tool in non-parametric statistics and inference, are the basis for estimating extreme values. Let a sample of significant wave height observations  $(x_1, x_2, \dots, x_n)$ , where each element  $x_i$  is assumed to be statistically independent with the same probability density function  $f(x)$ . If we classify the sample in ascending order of magnitude where the largest significant wave height is  $y_n$  and the smallest  $y_1$ , the sequence  $(y_1, y_2, \dots, y_n)$  would be the ordered sample, where each  $y_i$  would be a statistically independent random variable with a distinct probability density function. Then, the cumulative distribution function and probability density function of  $y_n$  would be:

$$G(y_n) = [F(x)]^n_{x=y_n} \quad (C.1)$$

and

$$g(y_n) = n[f(x)\{F(x)\}^{n-1}]_{x=y_n} \quad (C.2)$$

respectively.

If we assign to  $M_n = \max\{x_1, x_2, \dots, x_n\}$ ,  $n \in \mathbb{N}$  the maximum values from the original sample  $(x_1, x_2, \dots, x_n)$  then the cumulative distribution function of  $M_n$  can be estimated from the following expression (Karathanasi, 2020):

$$F_{M_n} = P(M_n \leq x) = \prod_{i=1}^n P(x_i \leq x) = [F(x)]^n, x \in \mathbb{R}, \quad n \in \mathbb{N}. \quad (C.3)$$

In most cases the cumulative distribution function  $F$  is unknown, thus the most common approach is to assess the asymptotic behavior of  $M_n$  at its right tail. This approach was first developed by Fréchet (1927), as well as the same distribution, and later Fisher and Tippett (1928) suggested two more distributions. Lets suppose that there are two quantities  $a_n$  and  $b_n$  ( $a_n > 0$ )  $\forall n$  and a distribution  $G(x)$  such that (Pickands, 1975):

$$\lim_{n \rightarrow \infty} P \left\{ \frac{M_n - b_n}{a_n} \leq x \right\} = \lim_{n \rightarrow \infty} F_{M_n}(a_n x + b_n) = G(x) \quad (C.4)$$

$\forall x$  that  $G(x)$  is continuous. Then  $G(x)$  is an extremal distribution function and  $F(x)$  is in its domain of attraction (Pickands, 1975). According to Gnedenko (1943), only three distributions satisfy the condition required for the asymptotic distribution. These three parametric families are (Karathanassi, 2020):

a) Type I - Gumbel Family:  $G_G(x) = \exp\left\{-\exp\left[-\left(\frac{x-a}{b}\right)\right]\right\}, x \in \mathbb{R}, b > 0,$

b) Type II - Fréchet Family:  $G_F(x) = \begin{cases} 0, & x \leq a \\ \exp\left\{-\left(\frac{x-a}{b}\right)^{-\xi}\right\}, & x > a, \xi > 0, \end{cases}$

c) Type III - Weibull Family:  $G_W(x) = \begin{cases} \exp\left\{-\left[\left(\frac{x-a}{b}\right)^\xi\right]\right\}, & x < a, \xi > 0 \\ 1, & x \geq a \end{cases}$

where  $a \in \mathbb{R}$  and  $b > 0$  are the location and scale parameters and  $\xi > 0$  is the shape parameter. The combination of all three aforementioned distributions constitutes the Generalized extreme value (GEV) distribution, as suggested by von Mises (1936) and Jenkinson (1955). The GEV is given by  $G_{GEV}(x; a, b, \xi) = \exp(-\Lambda(x))$  where

$$\Lambda(x) = \begin{cases} \left(1 + \xi \frac{x-a}{b}\right)^{-1/\xi}, & \xi \neq 0, \quad 1 + \xi \frac{x-a}{b} > 0, \\ \exp\left(-\frac{x-a}{b}\right), & \xi = 0, \quad x \in \mathbb{R} \end{cases}$$

with  $a \in \mathbb{R}$  being the location parameter,  $b > 0$  the scale parameter and  $\xi \in \mathbb{R}$  the shape parameter (Kapelonis et al., 2015). Interestingly the shape parameter  $\xi$ , which is also called the extreme value index, directly affects the behavior of GEVs tail (Kapelonis et al., 2015, Karathanasi, 2020). Furthermore, the parameters  $a$ ,  $b$ , and  $\xi$  can be estimated with various techniques such as the Maximum Likelihood method (MLM), the method of moments and the Probability Weighted Moment (PWM) method, amongst which the ML method has a tendency to be preferred for its flexibility in usage (Caires, 2016).

Amongst the various methods of extreme value analysis, the Peak-Over-Threshold (POT) method was chosen and applied in this particular study. This method is a popular one and has the advantage of making use of a larger set of observations per year in comparison with other methods (i.e. the Annual-Maxima), which consequently leads to a more accurate estimation of the parameters of a probability model (Caires, 2016). According to POT approach, if  $(X_1, X_2, \dots, X_n)$  is a sequence of independent and identically distributed random variables with common distribution function  $F(x)$ , and  $u$  is the selected high threshold, then the exceedances over  $u$   $Y_n = X_n - u, (n = 1, 2, \dots, N_u)$  are removed from the sequence.  $N_u$  is the number

of exceedances during a period of  $m$  years. The corresponding distribution of the excesses over  $u$  is described by the following equation (Kapelonis et al., 2015):

$$G^{[u]}(y) := P[Y \leq y | X > u] = \frac{F(u+y) - F(u)}{1 - F(u)}, \quad y \geq 0 \quad (C.5)$$

Furthermore, as stated by the Gnedenko-Pickands-Balkema-de Haan theorem (Gnedenko, 1943; Pickands, 1975; Balkema & de Haan, 1974), the  $G^{[u]}(y)$  distribution of exceedances can be approximated by the Generalized Pareto distribution (GPD) which is as follows:

$$G_{GP}(y; \sigma_u, \xi) = \begin{cases} 1 - \left(1 + \frac{\xi}{\sigma_u} y\right)^{-1/\xi}, & \xi \neq 0, \quad 1 + \frac{\xi}{\sigma_u} y \geq 0 \\ 1 - \exp\left(-\frac{y}{\sigma_u}\right), & \xi = 0, \quad y \in \mathbb{R} \end{cases} \quad (C.6)$$

where  $\sigma_u > 0$  and  $\xi$  are the scale and shape parameters respectively. Depending on the value of  $\xi$ , the GPD shows a distinct type of tail. When  $\xi = 0$ , it resembles a type I tail (and becomes the exponential distribution with mean  $\sigma_u$ ), when  $\xi > 0$  a type II tail (in which case it is the Pareto distribution) and when  $\xi < 0$  a type III tail (and therefore is a special case of the beta distribution) (Caires, 2016). It would be important to note that when  $u$  is high enough then the conditional excess distribution  $G^{[u]}(y)$  is approximately equal to the GPD  $G_{GP}(y; \sigma_u, \xi)$  (Kapelonis et al., 2015).

Another point of interest in the context of extreme value theory is the estimation of the  $m$ -year return level  $z_m$ , which essentially refers to the level which is exceeded once every  $m$  years (Davison & Smith, 1990). The expression which calculates  $z_m$  for the POT method is the following (Caires, 2016):

$$z_m = \begin{cases} u + \frac{\hat{\sigma}_u}{\xi} [(\lambda_u m)^\xi - 1], & \xi \neq 0 \\ u + \hat{\sigma}_u \ln(\lambda_u m), & \xi = 0 \end{cases} \quad (C.7)$$

where  $\lambda_u = \frac{N_u}{N}$  is the threshold crossing rate. Another important procedure is the selection of  $u$ , which if done correctly might improve the tail estimation to a great extent (de Haan & Zhou, 2009). The selected threshold should be such that maintains a balance between a sufficient sample size for the calculation of the GP parameters and the convergence of the excess to the GP distribution (Abild et al., 1992).

On a final note, special attention should be given to the independence between extreme events as it can be compromised in the case where the time step of the timeseries is smaller than the usual duration of an extreme event (Kapelonis et al., 2015; Karathanasi, 2020). For this reason, as a general rule a declustering methodology is applied at the timeseries before using the POT method.

## References

- Abild J, Andersen EY, Rosbjerg D, 1992. The climate of extreme winds at the Great Belt. *Denmark Journal of Wind Engineering and Industrial Aerodynamics* **41** (1-3): 521-532.
- Athanassoulis GA, Belibassakis KA, 2002. Probabilistic description of metocean parameters by means of kernel density models, 1. Theoretical background and first results. *Applied Ocean Research* **24**: 1-20.
- Balkema AA, de Haan L, 1974. Residual Life Time at Great Age. *The Annals of Probability*, **2**(5): 792-804.
- Belibassakis KA, Karathanasi FE, 2017. Modelling nearshore hydrodynamics and circulation under the impact of high waves at the coast of Varkiza in Saronic-Athens Gulf. *Oceanologia* **59**: 350-364.
- Booij N, Ris RC, Holthuijsen LH, 1999. A third-generation wave model for coastal regions, 1. Model description and validation. *Journal of Geophysical Research* **104** (C4): 7649-7666.
- Caires S, 2016. A Comparative Simulation Study of the Annual Maxima and the Peaks-Over-Threshold Methods. *Journal of Offshore Mechanics and Arctic Engineering* **138**: 051601-1:051601-14.
- Cavaleri L, Rizzoli PM, 1981. Wind wave prediction in shallow water: theory and applications. *Journal of Geophysical Research* **86**: 10961-10973.
- Davison AC, Smith RL, 1990. Models for Exceedances Over High Thresholds (With Discussion). *Journal of the Royal Statistical Society Series B* **52**(3): 393-442.
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bodlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, KållbergP, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F, 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of Royal Meteorological Society* **137** (656):553-597.
- EMODnet Bathymetry Consortium, 2018.
- Fisher RA, Tippett LHC, 1928. Limiting forms of the frequency distributions of the largest or smallest number of a sample. *Proceedings of the Cambridge Philosophical Society* **24** (Part 2): 180-190.

- Fréchet M, 1927. Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématiques* **6**: 93-116.
- GEBCO Compilation Group 2020, GEBCO 2020 Grid.
- Gelci R, Cazalé H, Vassal J, 1957. Prediction of waves. The method of frequency-directional spectral densities. *Bulletin d'Informacion - Comité central d'océanographie et d'études des côtes* **9**: 416-435 (in French).
- Gnedenko B, 1943. Sur la distribution limite du terme maximum d'une série aléatoire. *The Annals of Mathematics*, **44** (3): 423.
- Günther H, Rosenthal W, Weare TJ, Worthington BA, Hasselmann K, Ewing JA, 1979a. A hybrid parametrical wave prediction model. *Journal of Geophysical Research* **84**: 5727-5738.
- Günther H, Rosenthal W, Richter K, 1979b. Application of the parametrical surface wave prediction model to rapidly varying wind fields during JONSWAP 1973. *Journal of Geophysical Research* **84**: 4855-4864.
- Günther H, Behrens A, 2012. The WAM model. Validation document Version 4.5.4. Institute of Coastal Research Helmholtz-Zentrum Geesthacht (HZG), Geesthacht, Germany.
- de Haan L, Zhou C, 2009. Notes for the Discussion Deltares 15/10/08. Evaluation of the Statistical Methods Currently Used to Determine Boundary Conditions, Proceedings of a Workshop, Report No. H5098.70.
- Hasselmann K, 1962. On the nonlinear energy transfer in a gravity wave spectrum. Part 1. General theory. *Journal of Fluid Mechanics* **12**: 481-500.
- Hasselmann K, Ross DB, Müller P, Sell W, 1976. A parametric wave prediction model. *Journal of Physical Oceanography* **6**: 200-228.
- Hemer MA, Wang XL, Weisse R, Swail VR and COWCLIP Team, 2012. Advancing wind-waves climate science - The COWCLIP project, *Bulletin of the American Meteorological Society*, **93** (6): 791-796.
- Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut JN, 2020. The ERA5 global reanalysis. *Quarterly Journal of Royal Meteorological Society* **146** (730): 1999-2049.

- Jenkinson AF, 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society* **81**(348): 158-171.
- Kapelonis ZG, Gavriliadis PN, Athanassoulis GA, 2015. Extreme value analysis of dynamical wave climate projections in the Mediterranean Sea. *Procedia Computer Science* **66**: 210-219.
- Karathanasi F, 2020. Probabilistic modeling of linear and directional wind and wave data with applications to the marine environment, PhD Thesis.
- Komen GJ, Hasselmann S and Hasselmann K, 1984. On the existence of a Fully Developed Wind-Sea Spectrum, *Journal of Physical Oceanography* **14** (8): 1271-1285.
- Komen GJ, Cavaleri L, Donelan M, Hasselmann K, Hasselmann S and Janssen PAEM, 1994. Dynamics and Modelling of Ocean Waves, Cambridge University Press, 560 pp.
- Konishi S, 2014. Introduction to Multivariate Analysis: Linear and Non Linear Modeling, CRC Press, Taylor & Francis Group, 307 pp.
- Kontoyiannis H, Panagiotopoulos M, Soukissian T, 2015. The Euripus tidal stream at Halkida/Greece: a practical, inexpensive approach in assessing the hydrokinetic renewable energy from field measurements in a tidal channel. *Journal of Ocean Engineering and Marine Energy* **1**(3): 325-335.
- Lavidas G, Venugopal V, 2017. A 25 year high-resolution wave atlas for nearshore energy production and economics at the Aegean Sea. *Renew Energy* **103**: 401-417.
- Mardia KV, 1972. Statistics of Directional Data. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **37**(3): 349-393.
- Massel SR, 2017. Ocean Surface Waves: Their Physics and Prediction 3rd edition in Advanced Series on Ocean Engineering Volume 45. World Scientific, 776 pp.
- MEDATLAS project,2004. Wind and Wave Atlas of the Mediterranean Sea.
- Miles JW, 1957. On the generation of surface waves by shear flows. *Journal of Fluid Mechanics* **3**: 185-204.
- von Mises R, 1936. La distribution de la plus grande de n valeurs. Reprinted in *Selected Papers II*, American Mathematical Society, Providence R.I: 271-294.

- Phillips OM, 1957. On the generation of waves by turbulent wind. *Journal of Fluid Mechanics* **2**: 417-445.
- Phillips OM, 1977. Dynamics of the Upper Ocean, Second edition, Cambridge University Press.
- Pickands JI, 1975. Statistical Inference Using Extreme Order Statistics. *The Annals of Statistics*, **3** (1): 119-131.
- Ravdas M, Zacharioudaki A, Korres G, 2018. Implementation and validation of a new operational wave forecasting system of the Mediterranean Monitoring and Forecasting Centre in the framework of the Copernicus Marine Environment Monitoring Service. *Natural Hazards and Earth System Sciences* **18**: 2675-2695.
- Soukisian T, Prospathopoulos A, Korres G, Papadopoulos A, Hatzinaki M, Kambouridou M, 2008. A new wind and wave atlas of the Hellenic Seas. Proceedings of the *American Society of Mechanical Engineers 27th International Conference on Offshore Mechanics and Arctic Engineering*, June 15-20, Estoril, Portugal.
- Sverdrup HU, Munk WH, 1947. Wind, Sea and Swell: Theory of Relations for Forecasting, Hydrographic Office Publication No 601, U.S Navy Hydrographic Office, Washington DC: 44p.
- SWAMP Group, 1985. Ocean Wave Modeling. Plenum Press, New York, 256 pp.
- Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Bechtold VDC, Florino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg LVD, Bidlot J, Bormann N, Cairns S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J, 2005. The ERA-40 re-analysis. *Quarterly Journal of Royal Meteorological Society* **131** (612): 2961-3012.
- WAMDI Group, 1988. The WAM model - a third generation ocean wave prediction model. *Journal of Physical Oceanography* **18**: 1775-1810.
- Wessel P, Smith WHF, 1996. A Global Self-consistent, Hierarchical, High-Resolution Shoreline Database. *Journal of Geophysical Research* **101**: 8741-9743.
- Young IR, 1999. Wind generated ocean waves, Elsevier Ocean Engineering Book Series, Volume 2.

Zacharioudaki A, Korres G, Perivoliotis L, 2015. Wave climate of the Hellenic Seas obtained from a wave hindcast for the period 1960-2001. *Ocean Dynamics* **65**: 795-816.