Politecnico di Torino

Politecnico di Torino

Corso di Laurea
A.a. 2023/2024
Sessione di Laurea Dicembre 2024

Earth observation satellites for offshore wave and wind applications

Relatori:
Prof. Giuseppe Giorgi
PhD Giulia Cervelli

Candidati:
Eduard Salahoru 305956

Abstract

The Offshore Renewable Energy sector continuously requires accurate data, particularly regarding significant wave height and wind speed, which are key variables necessary to characterize the operational and environmental conditions of offshore wind farms. In this context, accurate data can support decision-making for a specific site to meet these conditions and, most importantly, facilitate continuous energy generation. In-situ platforms can help collect accurate data, but they are quite expensive and frequently experience malfunctions, leading to discontinuous observations.

In this scenario, satellites play an important role as they are capable of acquiring accurate data with discrete temporal resolution.

This thesis aims to evaluate the performance of satellite altimetry data by collocating it with fixed-point positions, between satellite altimetry data and in-situ data, using a spatio-temporal matching method, exploring different criterions.

Moreover, several bias correction techniques are applied to calibrate satellite data against in-situ data to improve the quality of satellite assimilated dataset, aligning it more closely with in-situ and and unlock their potential in providing environmental insights.

List of Tables

List of Figures

List of Acronyms

BC Bias Correction

CDF Cumulative Distribution Function

CNES Centre National d'Etudes Spatiales

EUMETSAT European Organisation for the Exploitation of Meteorological Satellites

FDM Full Dsitribution Mapping

GPS Global Positioning System

IDW Inverse Distance Weighting

L2P Level 2+ product

NRT Near Real Time

NTC Non Time Critical

ORE Offshore Renewable Energy

PDF Probability Density Function

PRF Pulse Repetition Frequency

QM Quantile Mapping

SAR Synthetic Aperture Radar

STC Short Time Critical

TAC Thematic Assembly Center

Contents

Chapter 1

Introduction

1.1 Importance of Data

In the last decades the technological progress has evolved enormously and thanks to this, the knowledge of everything that surround us has grown. New instruments, new types of sensors have been developed and thanks to these new technologies we can gather data from every type of phenomena. In the last years thanks to the arrival of AI and its potential, more than any other things, *Data* have gain an incredible value.
Given this perspective, Metocean Data is the main theme of this thesis. There are different ways to collect information in the marine environment such as saildrones, mooring buoys [1], drifters [2] and the list goes on [3]. These type of instruments are quite important because they are sensitive devices and can get accurate data about temperature, sea level, chlorophille, salinity. Contrary to these aspects, in-situ instrumentation are expensive, depending on the electronics or sensors within, and can be difficult to maintain due to the sever sea and weather conditions. Another disadvantage is that these devices can malfunction for any reason and therefore the temporal resolution is not continous.
Satellite altimetry missions can fill some of those gaps that in-situ instruments have [4]: depending on the type of mission and orbit, a single satellite can map almost entirely the whole globe in some days. Some of the key aspects of these platforms are that can deliver robust and precise measurements of large portions of area, rather than a fixed point measurement. Even satellites are affected by errors, due to the signal loss, orbit type, etc.. Satellite–based measurements have become more central in many sectors, such as the ORE sector, reducing costs [3].
The aim for this work is to investigate the different performance, in terms of spatial and temporal resolution and accuracy, of satellite data for wave and wind applications. This analysis is done firstly by collecting satellite and in-situ data using a time-space matching procedure, then calibrating the satellite dataset by applying some techniques known as *Bias Correction techniques*, in order to adjust more precisely the satellite data respect to in-situ dataset. The main goal is to make use of this data more specifically for offshore engineering applications rather than climate change employment.

1.2 Significant Wave Height and Wind Speed

The amount of information we can gather from our ocean is endless. Depending on which physical quantities we want to investigate, Copernicus Marine Service [5] proposed a 3

"divisions" of the ocean based on the marine type of data [6]:

- the Blue Ocean focuses on the physical state of the ocean (for example sea level)
- the Green Ocean focuses on the green part of the ocean, such as variations in the biological carbon pump or biogeochemical processes
- the White Ocean focuses on the sea ice, in particular monitoring the poles or melting of sea ice

The focus of this work is principally on the Blue Ocean, in particular investigating the importance of **significant wave height** and **wind speed**.
These physical variables play a crucial role into the growth and development of ORE sector. Generating energy in a greener way is becoming the main theme for the next 20 years: depending on the technology, extracting energy from wind and waves can definetely help achieving transition to a greener future. In this scenario satellites are crucial because deliver us robust and continuative data. Despite the incredible contribution, satellite data lack of spatial and temporal resolution [7]: at the moment the goal for future is to improve spatial resolution and revisit time, in order to meet the requirements of the ORE sector [8]. The question is: how this data can become beneficial and important for the industry? The ORE sector has been identified as needing enhanced data in two areas: defining the operational conditions such that a wind farm or any type of power sources can work and environmental constraints (see figure 1.1) [3]. These two aspects can make the difference when the decision has to be made on whether or not an offshore site can be realized and especially the location of that site [3]. An offshore wind farm cannot be placed anywhere in the sea but has to meet operational and environmental constraints so that safety is guaranteed and most importantly energy is generated in a continuous way. Life cycle of a site is important for the decision making and then for profitability of the project. For example turbulence is correlated with early fatigue failures, therefore impacting on longevity of the wind farm [9]. Moreover resource availability and potential energy of a site can also determine profitability of the project.

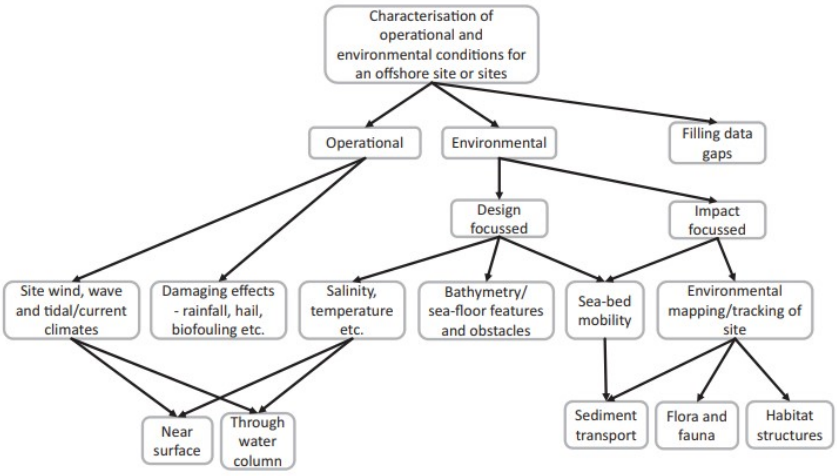
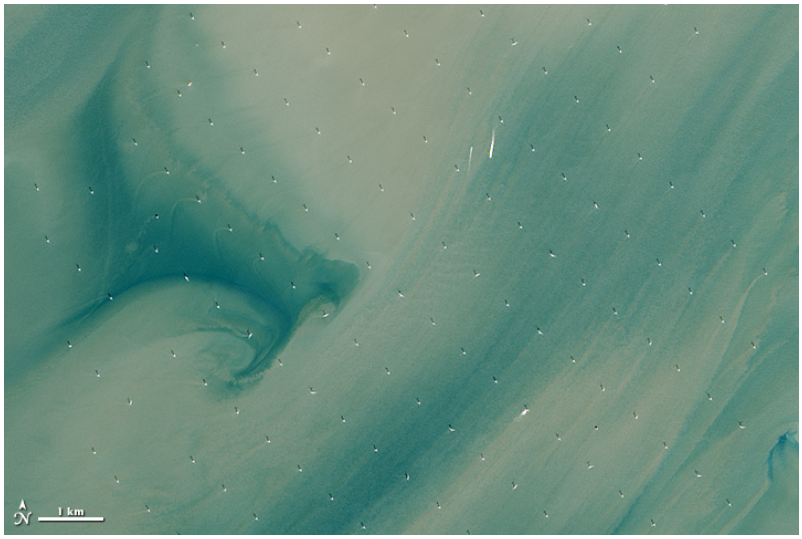


Figure 1.1: Scheme of the different requirements which an off-shore site is operational [3].

Without going into details, the design of a wind farm (see figures 1.2a and 1.2b) takes into account different phase, that were not mentioned, such as technology selection, layout and cabling optimisation and management as well as operational and end-of-life practicalities [3].



(a)

(b)

Figure 1.2: The world's largest offshore wind farm located in the Thames Estuary (in the North Sea) and has an output power of 630 MW.

Chapter 2

Oceanography Data

2.1 In-situ measurements: ocean waves and wind

Significant wave height and wind speed variables are measured differently, depending of the type of instrumentation. As for the in-situ measurements, wind speed is relatively simple to measure, whereas significant wave height is more complex, in particular 2 types of techniques are used.
Waves are disturbances of the equilibrium state in any given body of material, which propagate through that body over distances. Each waves can be described with their respective wavelength and period, which can determine the type of waves.
As an example, tides are generated because of the interaction between oceans and gravitational forces of the Moon and the Sun [10]. Waves can be classified by the period (i.e, time between two consecutive peaks in the wave). Waves that are generated by the interaction of ocean and the atmosphere are called *surface gravity waves* and have a period between 1-4 s and 30 s. Depending on the location of origin, surface gravity waves are then divided into 2 different categories.
Wind sea waves are generated by the local wind and tend to have an irregular pattern while having short crests (typically in the open ocean). While swell are generated as well by the wind but they have travelled long distances from their origin (typical wave we can see right on the beach) [10].
Now is clear that different type of waves exist depending on wave period but it is important to undestand how the motion of waves is captured and recorded. There are two different observation techniques: in-situ platforms and by satellite.
There are several methods to capture and measure waves over time, whether by insitu techniques or satellite altimetry (remote sensing). Regarding to insitu techniques, here are presented 2 methods: *Zero Crossing Method* and *Spectral Analysis* [10], [11] .
Satellites instead, orbit continuously around the earth, following a specific trajectory, and records the characteristics of the waves along the path it takes. This is done thanks to the altimetry technology. An echo impulse is emitted by the instrument, after that it hits the sea surface and it is bounced back to the satellite: this impulse contain itself the information of the sea surface.

2.1.1 Ocean waves: Zero Crossing Method

Now we know that waves are disturbances of the equilibrium state, we need to determine wave characteristics. Usually wave buoys have on-board accelorometers to record vertical acceleration, then it is integrated two times to obatin surface elevation over time [10], [11]. After that, a wave is defined as the profile of the surface elevation between two successive downward zero-crossings of the elevation (zero = mean of surface elevations, see figure 2.1). An alternative approach can be done, by choosing upward zero-crossings.

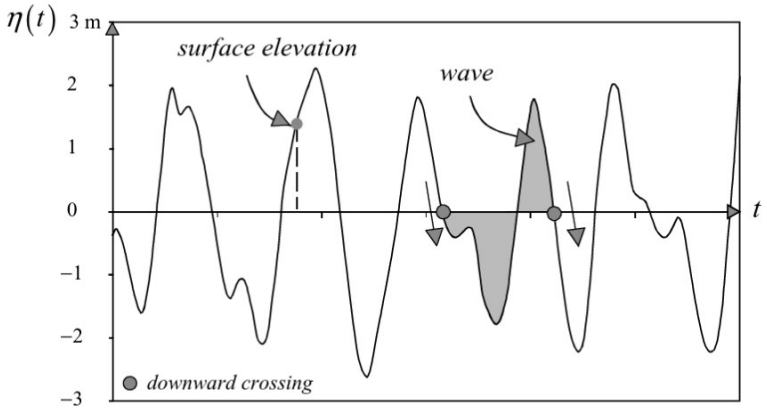


Figure 2.1: Distinction between waves and surface elevation [10].

Depending on the type of method used if the surface elevation $\eta(t)$ follows a typical normal distribution (Gaussian), statistical characteristics are identical, whether upward or downward method is used. Therefore wave is described by taking the wave record and then analyzing all of the individual wave heights and periods in the record. The duration of each measurement has to be short as possible, so that it is stationary but also long enough to describe accurately the sea state. The commonly duration of a time record used is 15–30 min.
Now let us define the wave height, given N waves record, we define the mean height $\overline{H}$:

$$\overline{H} = \frac{1}{N} \sum_{i=1}^{N} H_i \tag{2.1}$$

Instead of using this characterstic, *significant wave height* $H_{1/3}$ is the most used. It is calculated as the mean of the highest one-third of waves:

$$H_{1/3} = \frac{1}{N/3} \sum_{j=1}^{N/3} H_j \tag{2.2}$$

where j is not the sequence number in the record but the rank number of the wave, based on wave height (i.e., $j = 1$ is the highest wave, $j = 2$ is the second-highest wave, etc.) [10, 11]. For the rest of the thesis, when refering to the significant wave height the notation H_s will be used.
Sometimes the mean of the highest one-tenth of waves is used to define $H_{1/10}$:

$$H_{1/10} = \frac{1}{N/10} \sum_{j=1}^{N/10} H_j \tag{2.3}$$

It is clear that after defining the significant wave height, period of the wave has to be defined as well. Period of a wave is the time interval between two crests. The definitions, like mean of zero-crossing wave period or significant wave period, are similar to the previous ones.
The mean of the zero-crossing wave period, denoted as $\overline{T_0}$, is then defined similarly as:

$$\overline{T}_0 = \frac{1}{N}\sum_{i=1}^{N} T_{0,i} \tag{2.4}$$

Mean period of the highest one-third of waves is defined as:

$$T_{1/3} = \frac{1}{N/3}\sum_{j=1}^{N/3} T_{0,j} \tag{2.5}$$

where j is rank number of the wave, based on wave height (it is the same j as in the definition of the significant wave height) [10, 11].
Finally, mean of the highest one-tenth of waves is defined as:

$$T_{1/10} = \frac{1}{N/10}\sum_{j=1}^{N/10} T_{0,j} \tag{2.6}$$

2.1.2 Ocean waves: Spectrum Analysis

In the following sub-section the second methodology is presented, which is used to determine H_s, through a spectrum analysis [10, 11, 12]. The main goal of this analysis is to describe the sea state as a stochastic process (i.e, given the aleatory nature of the process, the goal is to describe all the possible observations that could have occurred under the circumstances of the actual observation).

Random phase/amplitude model

Through this model we look at the wave spectrum in order to describe the surface elevation over time. A more complete description of the surface elevation over time is described by a Fourier summation of different wave components (see figure 2.2):

$$\underline{\eta}(t) = \sum_{i=1}^{N} \underline{a}_i \, cos(2\pi f_i t + \underline{\alpha}_i) \tag{2.7}$$

where N is the number of frequencies, $f_i = 1/D$ i-th frequency, the underscores of amplitude a_i and phase α_i indicate that these are random variables.
In this case phase and amplitude are random variables, that are fully described by their respective probability density functions (see figure 2.3). In this model, the phase at each frequency f_i is uniformly distributed between 0 and 2 π:

$$p(\alpha_i) = \frac{1}{2\pi} \qquad \text{for} \quad 0 < \alpha_i \leq 2\pi \tag{2.8}$$

whereas amplitude $\underline{a_i}$ at each frequency has Rayleigh distribution:

$$p(\underline{a_i}) = \frac{\pi}{2}\frac{a_i}{\mu_i^2}\exp(-\frac{\pi a_i^2}{4\mu_i^2}) \qquad \text{for} \quad a_i \geq 0 \tag{2.9}$$

where μ_i is the expected value of the amplitude $\mu_i = E\{\underline{a}_i\}$. In the probabilistic theory, expected value is the mean of a random variable and in particular this average is weighted by their respective probability.

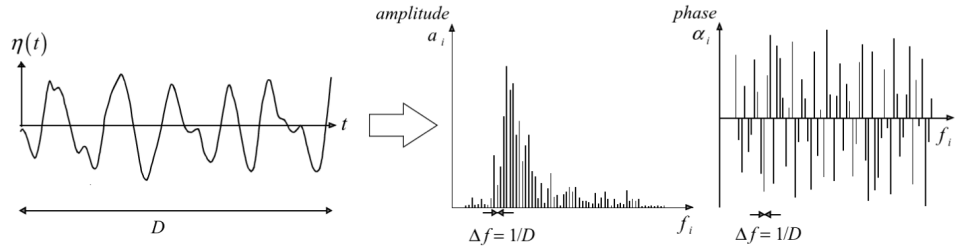


Figure 2.2: Surface elevation decomposed into the wave components spectrum [10].

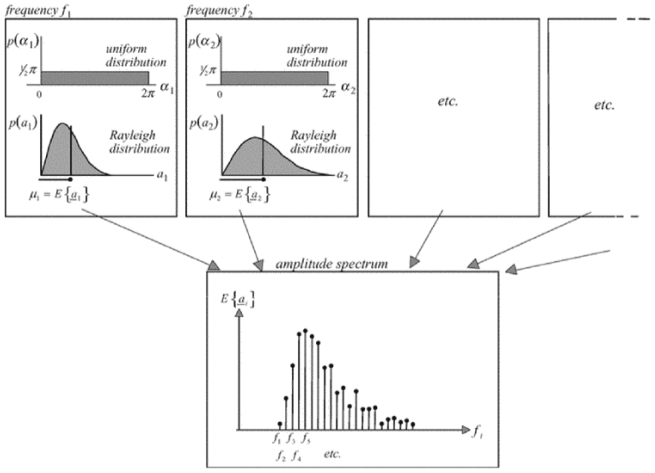


Figure 2.3: Amplitude and phase with their respective PDF [10].

For a given amplitude spectrum, we can obtain $\overline{\eta}(t)$ by choosing sample values of $\underline{a}_i$ and $\underline{\alpha}_i$ from their respective PDF, at each frequency separately. Summing all these components we get the wave record. Whether we want to obtain a new wave record $\eta(t)$, the sample values of $\underline{a}_i$ and $\underline{\alpha}_i$ are again drawn randomly from their PDF (see figure 2.3).
The applicability of this model is tied to stationary processes. Clearly most of the times the sea state is not stationary. Secondly, this model utilize a summation of wave components at discrete frequencies, where at sea a continuum of frequencies are present [10, 11, 12]. We resolve this problem by adding a new element to this model: we focus on the *variance spectrum* $E\{\frac{1}{2}\underline{a}_i^2\}$ instead of the amplitude spectrum (see figure 2.4).

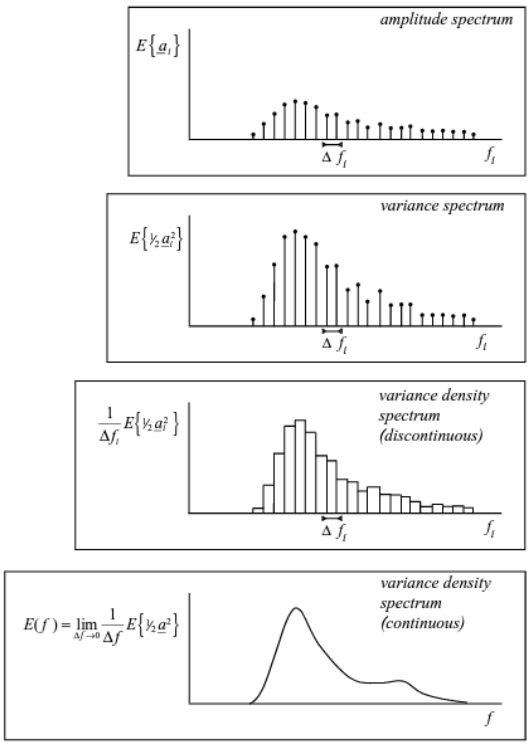


Figure 2.4: Differences between amplitude spectrum and variance spectrum [10].

Variance density spectrum

The variance density spectrum provides a statistical description of the sea-surface elevation of ocean waves, while the amplitude spectrum characterizes sea-surface elevation as a stationary process. This type of process indicates that all statistical variables of the wave field can be represented through this spectrum. It is reminded that realistically never happens.
Instead, the variance density spectrum can be interpred how the variance of the sea surface elevation is distributed over the frequencies. In this context, the focus is shifted towards variance $E\{\frac{1}{2}\underline{a_i^2}\}$ rather than the above introduced expectation of the amplitude $E\{\underline{a_i}\}$ [10, 11, 12].
Both amplitude and variance spectrum are based on discrete frequencies, whereas at sea all frequencies are present, therefore a change on the amplitude/phase model has to be made. Now variance is distributed over the frequency interval Δf_i at frequency f_i. The resulting variance $E^*(f_i)$ is:

$$E^*(f_i) = \frac{1}{\Delta f_i} E\{\frac{1}{2}\underline{a_i^2}\} \qquad \text{for all } f_i \tag{2.10}$$

By having width of frequencies closer to zero, the continous version of the variance is:

$$E(f) = \lim_{\Delta f \to 0} \frac{1}{\Delta f} E\{\frac{1}{2}\underline{a_i^2}\} \tag{2.11}$$

Frequency-direction spectrum

To describe the actual, three-dimensional, moving waves, the horizontal dimension has to be added. Therefore we expand the random-phase/amplitude model by considering

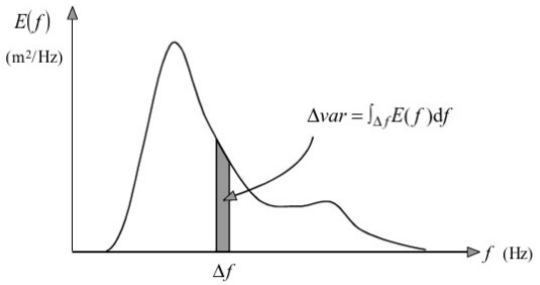


Figure 2.5: Contribution of Δvar variance for each frequency band Δf [10].

a harmonic wave that propagates in x and y-space,in direction θ relative to the positive x-axis:

$$\eta(x, y, t) = a\cos(\omega t - kx - ky\sin\theta + \alpha) \tag{2.12}$$

where k is the wave number, $k_x = kcos\theta$, $k_y = ksin\theta$ and θ is the direction of wave propagation (i.e., normal to the wave crest of each individual component). The corresponding three-dimensional random-phase/amplitude model is quite similar to equation 2.7, whereas in this case y direction is added:

$$\underline{\eta}(x, y, t) = \sum_{i=1}^{N}\sum_{j=1}^{M}\underline{a}_{i,j}\cos(\omega_i t - k_i x\cos\theta_j - k_i y\sin\theta_j + \underline{\alpha}_{i,j}) \tag{2.13}$$

Frequency ω and wave number f are related linearly, through the disperions relationship of the linear theory for surface gravity waves $\omega^2 = gk\tanh(kd)$ (where d = water depth). Now that frequency and wave number are correlated, the three-dimensional model is reduced to two-dimensional model depending on wave number (or frequency) and direction.
By applying the same principles used to determine variance density spectrum in the $1D$ model, the same can be done to the two-dimensional model:

$$E(f, \theta) = \lim_{\Delta f\to 0}\lim_{\Delta\theta\to 0}\frac{1}{\Delta f\Delta\theta}E\{\frac{1}{2}\underline{a}^2\} \tag{2.14}$$

As it was mentioned in the one-dimensional model, variance density spectrum describes variance distribution over frequencies and directions (see figure 2.6.)

Wave period

Finally, here wave period and wave height are calculated.
One of the results random-phase model produced is that surface elevation at an arbitrary moment t_1 assumes a normal distribution and can be written as:

$$p(\eta) = \frac{1}{(2\pi m_0)^{1/2}}\exp(-\frac{\eta^2}{2m_0}) \qquad \text{for a zero-mean: } E\{\underline{\eta}\} = 0 \tag{2.15}$$

where $m_0^{1/2}$ is the standard deviation of the surface elevation, whereas m_0 is the spectral moment of zeroth-order (see Appendix A) [10, 11, 12].
However, this function only gives the probability that $\eta(t)$ is below a certain level η. After

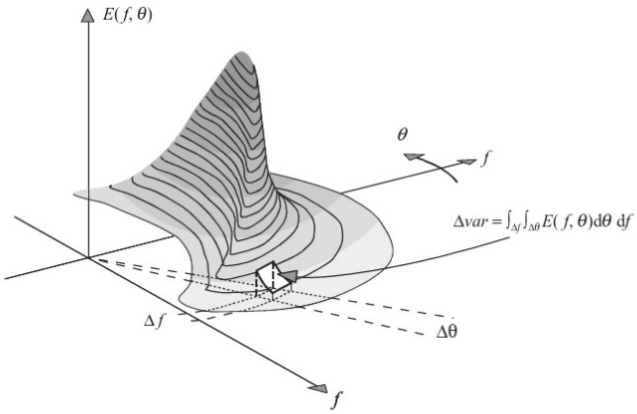


Figure 2.6: Contribution of Δvar variance to the frequency-direction spectrum model [10].

that, the average of the time interval between two crossings can be calculated in terms of the spectrum as (see Fig. 2.7):

$$\overline{T}_\eta = \sqrt{\frac{m_0}{m_2}} / \exp(-\frac{\eta^2}{2m_0}) \tag{2.16}$$

and then the frequency:

$$\overline{f}_\eta = \sqrt{\frac{m_2}{m_0}} \exp(-\frac{\eta^2}{2m_0}) \tag{2.17}$$

Unfortunately, the value of m_2 is sensitive to small errors or variations in the measurement or analysis technique, especially sensitive to noise in high-frequency range spectrum. Less dependent on high-frequency noise is T_s, the significant wave period. A theoretical expression for T_s in terms of the spectrum is available but it is rather complicated and it will not be treated here (see [10]). The following relationships are empirical:

$$T_s \approx T_{peak} \qquad \text{for swell} \tag{2.18}$$

$$T_s \approx 0.95 T_{peak} \qquad \text{for wind sea} \tag{2.19}$$

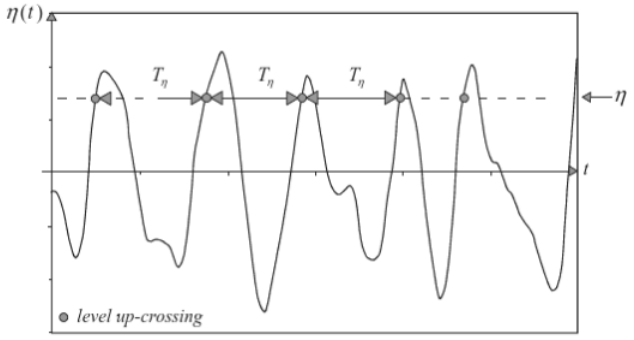


Figure 2.7: The up-crossings of the sea-surface elevation. [10]

Wave height

After defining the wave period, finally *wave height* can be determined. It is much easier for waves in deep water with a narrow spectrum to define wave height as double the crest elevation $\underline{H} \approx 2\underline{\eta}_{crest}$ (see figure 2.8) [10].
The PDF of $\underline{H}$ is determined from the PDF of $\underline{\eta}_{crest}$ by applying a Jacobian transformation

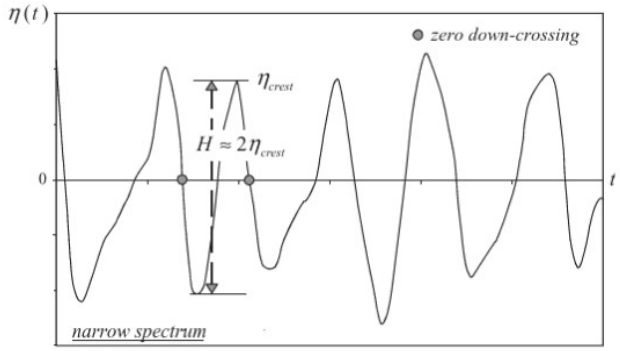


Figure 2.8: Definition of wave height for waves with a narrow spectrum [10].

and at the end the final distribution function is:

$$p(H) = \frac{H}{4m_0} \exp\left(-\frac{H^2}{8m_0}\right) \tag{2.20}$$

which is the Rayleigh distribution. By substituting $\underline{\eta}_{crest} = \underline{H}/2$ into this distribution gives the CDF for the individual wave height $\underline{H}$:

$$\Pr\{\underline{H} \leq H\} = 1 - \exp\left(-\frac{H^2}{8m_0}\right) \tag{2.21}$$

Now all the statistical characteristics of $\underline{H}$ follow the Railegh distribution:

$$\overline{H} = E\{\underline{H}\} = \sqrt{2\pi m_0} \tag{2.22}$$

and

$$H_{RMS} = E\{\underline{H}^2\}^{1/2} = \sqrt{8m_0} \tag{2.23}$$

At the beginning of the section, H_s was indicated as the mean value of the highest one-third of wave heights, using the zero-crossing method. This fraction is related to the Rayleigh distribution, where the wave heights are located in the highest third of this distribution (see figure 2.9):

$$\int_{H^*}^{\infty} p(H)dH = \frac{1}{3} \tag{2.24}$$

The mean value, which is the significant wave height, can be calculated as:

$$H_{m0} = E\{\underline{H}\}_{H \geq H^*} = \frac{\int_{H^*}^{\infty} Hp(H)dH}{\int_{H^*}^{\infty} p(H)dH} \tag{2.25}$$

and by substituting equation 2.24 into this one, gives the following result:

$$H_{m0} \approx 4\sqrt{m_0} \qquad \text{(for deep water)} \tag{2.26}$$

The notation used H_{m0} is different from the original notation statement, it is still reminded that it is the same variable but for the sake of this mathematical model a different notation is used.

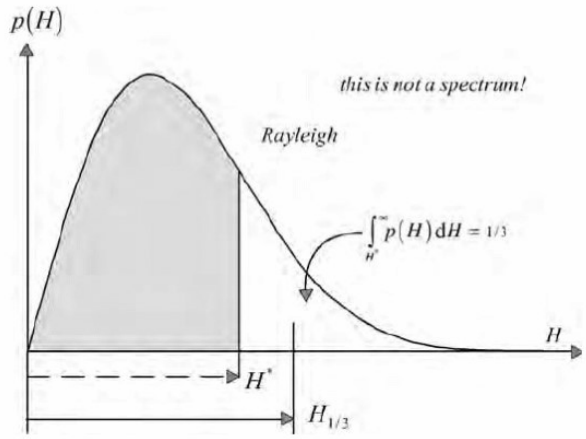


Figure 2.9: Rayleigh PDF of the significant wave height. [10]

2.1.3 Wind: instrumentation and measurements

The basic parameters essential for any wind site's potential assessment are *wind speed* and *wind direction*. These physical quantities are measured primarily by *anemometers*. This sections describes how wind is generated, measured and a brief description of the instrumentation is presented.

Mechanisms of wind's generation

Wind is the motion of air relative to the ground and its origin comes from a series of forces [13]:

1. *Pressure gradient force* that comes from the difference in pressure across a surface, where air moves from high pressure to low pressure areas.

2. *Gravity* causes air to accelerate downward at a rate of 9.81 m/s^2. It does not affect formation of horizontal winds.

3. *Friction,* the force that is generated by air contact with the surface on the ground, can be defined as

$$F = -\mu v \tag{2.27}$$

   where μ is the friction coefficient which depends on the type of surface; v is the wind velocity.

4. *Coriolis force* is caused by the rotation of the Earth: it deflects the wind to the right or the left depending on the hemisphere:

$$F_k = 2\rho v \omega \sin(\phi) \tag{2.28}$$

   where ρ is the density of air, v is the wind velocity, ω is the rate of rotation of Earth and ϕ is the latitude.

5. *Centripetal force* is a force acting on a body in circular motion, that pushes toward the center of curvature:

$$F_B = \frac{mv^2}{r} \tag{2.29}$$

where r is the radius of curvature of the trajectory.

Surface Roughness

Wind speed magnitude is strongly influenced by Earth's surface roughness, indeed in some parts of the World, wind speed is much stronger than others regions [14]. Over the ocean, surface roughness is quite small despite the presence of waves, whereas towards inland wind velocity reduces drastically (going further from the coast, wind speed becomes higher) [14]. In general wind has a height profile, which is directly correlated to surface roughness: at higher heights, wind velocity is not disturbed by uneveness of the surface, indeed a boundary layer is formed and the horizontal velocity is dependent on the height and this relationship can be expressed as this equation, known as *Wind Profile Power Law* (see also figure 2.10):

$$\frac{v_2}{v_1} = \left(\frac{h_2}{h_1}\right)^{\alpha} \tag{2.30}$$

where v_2 and v_1 are respectively wind velocity at heights h_2 and h_1 above the ground, while α is Hellmann height coefficient, which can be expressed in terms of standard ground roughness $z_0 = 0.1\ m$ [14]:

$$\alpha(z_0) = \frac{1}{\ln\left(\frac{15.25\ m}{z_0}\right)} \tag{2.31}$$

Equation (2.30) is quite important because will be used later in the spatio-temporal matching analysis to calculate wind speed measured by moorings: is is a standard procedure to evaluate it at a height of 10 m, indeed in-situ instruments sometimes cannot give a measurement at that height and for this reason wind speed power law comes really helpful. Therefore, in this work when refering to wind speed the notation U_{10} will be used.

Instrumentation

Wind speed is the most important indicator of a site's wind resource and potential of generating energy, therefore obtaining the most accurate measurements of the free-stream wind speed is crucial when a site availability is evaluated. As far for the scope of this thesis, it is reminded that in-situ measurements are labeled as the *ground truth* because it is assumed that they are the most accurate measurements. These type of instruments are precise but in contrary they are quite expensive and can malfunction for many reasons, on the other side satellite can fill those gaps. In order to get better measurements from satellite, corrections techinques on this data are applied to get a better matching with the in-situ measurements.
Most of the instrumentation used to measure wind speed is composed by different types of *anemometers.* Three anemometers types are used for the measurement of horizontal wind speed [15].

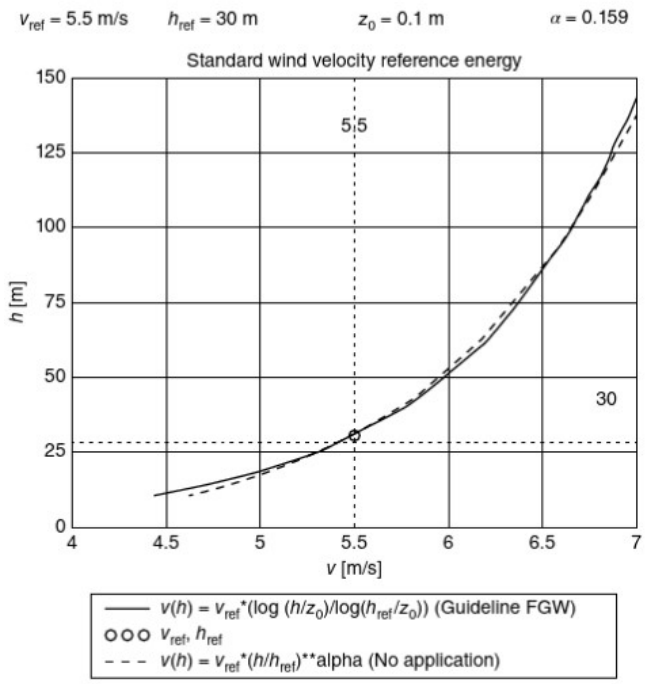


Figure 2.10: Wind profile velocity expressed in terms of height h [14].

The *Cup anemometer* (see figure 2.11) consists of three semispherical cups mounted one on each end of horizontal arms. The axis of rotation is vertical and the velocity of rotation of the cups is proportional to the wind speed, whereas C_1 and C_2 are clutch coefficients of concave and convex surfaces of anemometer cups with air. The forces acting on the opposite cups are defined as [13]:

$$F_y = \frac{1}{2} C_1 \rho S \, (v - v_t)^2 \tag{2.32}$$

$$F_o = \frac{1}{2} C_2 \rho S \, (v + v_t)^2 \tag{2.33}$$

where v is the wind speed, v_t is tangential speed of the cup that rotates,ρ is air density, S is the cross-sectional area of the cup. The equilibrium with air leads to the following equation:

$$C_1 (v - v_t)^2 = C_2 (v + v_t)^2 \tag{2.34}$$

The solution of the equation for v_t leads to the following expression:

$$v_t = v \left(\frac{1 - \sqrt{C_1} C_2}{C_1 - C_2} \right) \tag{2.35}$$

The *Windmill anemometer* combines a 3-4 blade propeller and a tail that leads the propeller orientation toward the direction of the wind (see figure 2.12). The relationship between the velocity blade and wind speed is described as:

$$U_t = \frac{v}{k \tan(\theta)} \tag{2.36}$$

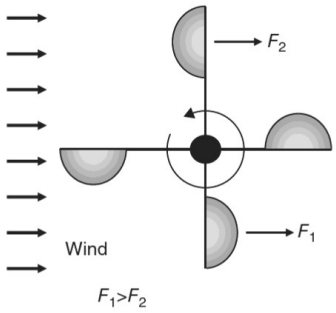


Figure 2.11: Cup anemometer [13].

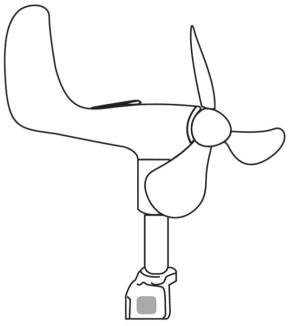


Figure 2.12: Windmill anemometer [13].

where θ is the propeller blade angle relative to the axis rotation and k is the coefficient that depends on the propeller design [13].

The *Sonic anemometer* does not have any rotating parts and measures the wind speed and direction by detecting variations in the speed of ultrasounds transmitted between fixed points. Some sonic anemometers measure wind in two dimensions, whereas others measure in three (see figure 2.13). Because they have no rotational inertia, sonic anemometers are more reactive to rapid speed and direction fluctuations than cup or propeller anemometers. They are also usually more expensive than other types and require more power mainly because they have inside 2-3 pairs of transducers; the path length between them is 10–50 *cm* Along the wind direction the ultrasound propagates faster and along the opposite direction it propagates slower. An ultrasonic anemometer can measure phase shifts in sound waves, which vary based on the orientation of the transducers in relation to the wind direction. [16].

The various pros and cons of these instruments are described in the tables 2.1 and 2.2.

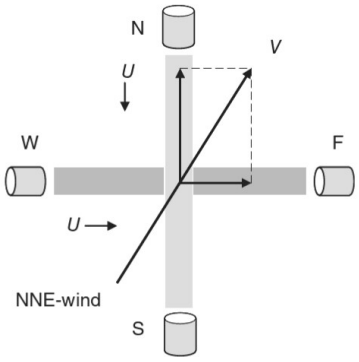


Figure 2.13: Working principle of a sonic anemometer [13].

Instrument	Advantages
Cup Anemometer	- Linear dependence of v_t on the wind speed , if clutch coefficients are constant (this condition is satisfied for cup anemometer). - Wide span operability varies from a range of 90 mm/s to 2.24 m/s. - Portable, simple and cheap.
Windmill anemometer	- Simple and compact; the propeller can be made out of plastic. - Ability of measuring weaker air currents and turbulence - Quick to react to gusts than the cup anemometer. - Relatively low cost.
Sonic Anemometer	- Sensitive and accurate when there are windspeed fluctuations. - Insensitive to icing. - No rotating parts: easy to install in remote locations. - There is less chance of these instruments to malfunction.

Table 2.1: Advantages of three kind of anemometers [13].

Instrument	Disadvantages
Cup Anemometer	- Moving parts wear out. - Slow to react to gusts. - Not sensitive to small wind speeds.
Windmill anemometer	- Moving parts wear out. - Has to be oriented into the wind direction.
Sonic Anemometer	- Requires an accurate calbration: dependence of environmental conditions. - The structure supporting the transducers creates a distortion of the flow itself. - Lower accuracy in precipitation, as raindrops can modify speed of sound. - High complexity and cost due to the electronic parts.

Table 2.2: Disadvantages of three kind of anemometers [13].

2.2 Satellite missions for Oceanography

Nearly 70 years have passed since the first launch of a satellite into space, specifically into orbit around the Earth [17]. On October 4, 1957, the launch of Sputnik was successfully carried out: in effect, a steel sphere the size of a basketball, which was used to send radio signals from the satellite to Earth, enabling communication with the satellite itself [17]. From that moment on, the aerospace sector, particularly the market for satellites orbiting the Earth, has grown enormously in economic and technological terms.
Currently, there are numerous missions dedicated to remote sensing, along with various possible applications we can find. As far as the European territory is concerned, Copernicus [5] currently plays a key role in the field of Earth observation.
The Copernicus Programme [5, 18], is a European Union-led initiative that focuses on Earth observation. Its primary goal is to monitor the environment and provide data for climate change, natural disasters, and various environmental challenges. Launched in 2014, the programme delivers free and open data through satellite systems (such as the Sentinel satellites [19]) and ground-based observations. There are several world-wide programmes with their own respectively satellites: in the figure 2.14 is illustrated the timeline of radar altimetry missions with the future missions planned.

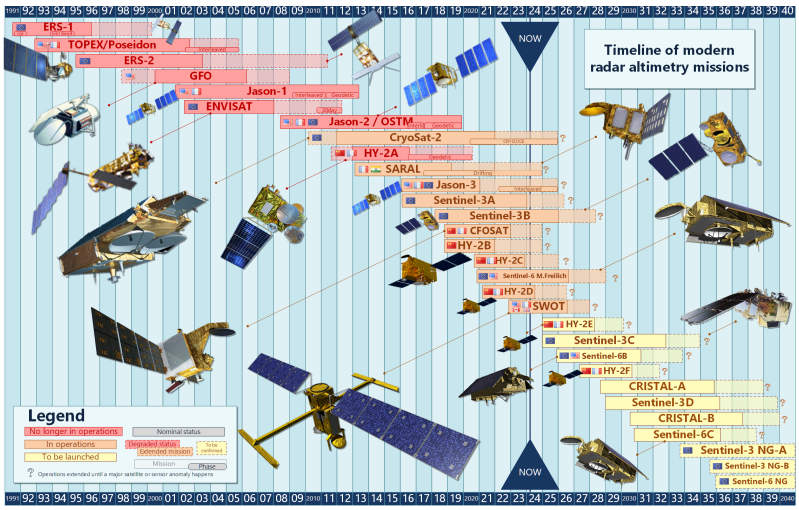


Figure 2.14: Timeline of modern radar altimetry missions [20].

2.2.1 Orbits

In the figure 2.14 is illustrated the timeline of the altimetry missions but also the modern ones that are still operating. Other than the instruments on-board, the orbital characteristics are the parameters that differentiate one from each other. The type of orbit is crucial for altimetry mission because it can make the difference on the temporal aspect. Now we shall present some definitions.
An *orbit* is defined as a single revolution around the Earth by the satellite.A satellite *pass* or *track* is half a revolution of the Earth by the satellite from one extreme latitude to the opposite extreme . *Repeat cycle* is the time period that elapses until the satellite passes over the same location again [20].
In addition to that satellites can operate in many types of orbits, generally speaking for the remote sensing applications satellites orbit in a Low Heart Orbit (LEO).

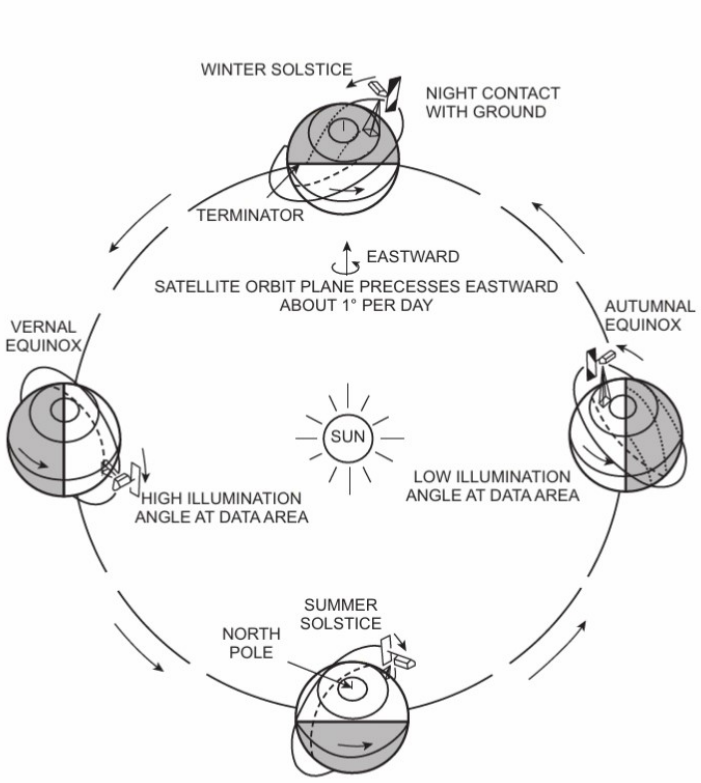


Figure 2.15: Rotation of the SSO plane [17].

A Low Earth Orbit (LEO) orbit is a path that satellites follow at an altitude between 160 and 2000 km. Weather satellites and some Earth observation satellites operate in the 500 e 900 km range [17]. Many of these satelites operates in *Sun-synchronous orbits*. A sun-synchronous orbit (SSO) is a type of near-polar orbit that allows a satellite to pass over the same region of the Earth's surface at approximately the same local solar time on each pass. For example, a satellite operating in this orbit might cross the equator 12 to 15 times a day (depending on the orbital altitude) each time at about 3:00 pm local time. Another key aspect is that the orbital plane of a SSO must also rotate approximately one degree eastward each day to keep pace with the Earth's revolution around the Sun [17].
Beyond SSO and LEO, other types of orbits exist such as GEO, MEO (Medium Earth Orbits) and GTO (Geostationary Transfer Orbits). It is important to mention briefly what is a GEO orbit. Basically, a satellite orbits around Earth at an altitude of 35,786 km and at this height the satellite moves with such an high speed that allows continuously observe the same location on the surface (i.e,the satellite rotates in sync with Earth's rotation, making it appear stationary in the sky) [17]. This orbit is used by both communications and weather satellites. One of the advantages is that antennas do not need to track satellite's motion, but can be pointed in a fixed postion towards the satellite of interest [17].
From this brief explanation of the types of orbits. it is clear that altitude affects thy type of orbit, therefore changing satellite mission (it is meant to be used for weather, communications, defense or remote sensing applications). In the table 2.3, a brief presentation of orbital characteristics of active remote sensing satellites missions: most of them are SSO orbits, for the reason explained before. Another important aspect of this table is Repeat Cycle: depending on the satellite's altitude, it can varies drastically. For example *Jason-3* and *Sentinel-6* have been engineered in order that they have the same orbit (*Sentinel-6* is meant to be the successor of *Jason-3*). Beyond spatial resolution, also temporal resolution

Table 2.3: Brief summary of the orbital characteristics of the satellites.

Satellite	Orbit type	Altitude	Repeat Cycle	Launch Date	End Date
Saral	SSO	800 km	35 days	25/02/2013	On-going
Cryosat-2	SSO	717 km	30 days	08/04/2010	On-going
Jason-3	Non Polar	1336 km	10 days	17/01/2016	On-going
Sentinel-3A	SSO	814 km	27 days	16/02/2016	On-going
Sentinel-3B	SSO	814 km	27 days	25/04/2018	On-going
Cfosat	SSO	500 km	13 days	29/10/2018	On-going
HaiYang	SSO	973 km	14 days	25/10/2018	On-going
Sentinel-6MF	Non Polar	1336 km	10 days	21/11/2020	On-going
Swot	Non Polar	890 km	21 days	16/12/2022	On-going

is the key factor that makes the difference: the goal is to obtain data with high resolution in the lowest time-frame possible.

2.2.2 Altimeter principle

Missions of Earth observation focus are based on a rather simple principle. The satellite emits an impulse and the analysis of the returned signal allows the calculation of the time needed by the signal to go and come back, i.e. the distance satellite-sea surface (fig 2.17).This analysis can be done on the wave form graph (see figure 2.16) which describes the energy of the signal over time [20, 21]. The time that elapses between when the pulse

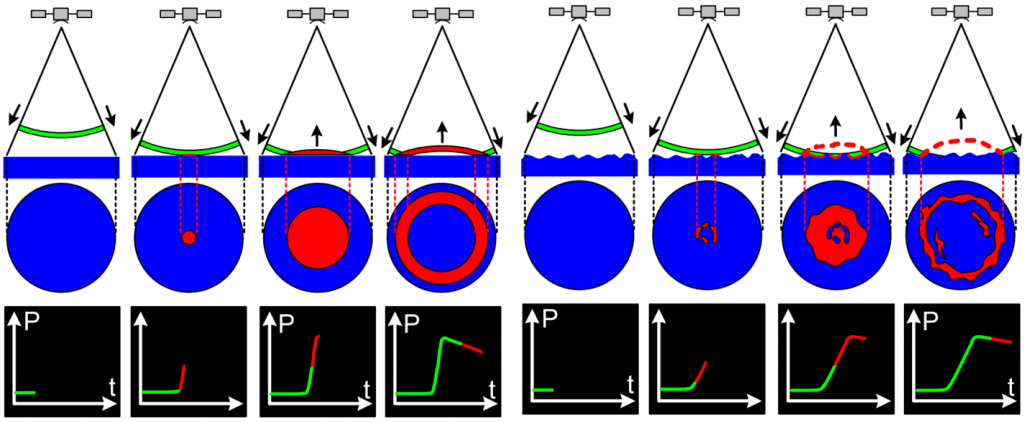


Figure 2.16: Surface elevation distribution effects on the wave form [20].

is emitted and when the pulse is reflected by the Earth is proportional to the altitude of the satellite [20, 6, 21]. The satellite altitude refers to the distance from the satellite's center of mass above a reference point. This reference point is often indicated on a reference ellipsoid or the center of the Earth. For example, to measure the sea surface level, we need to apply the following equation ([20, 6]):

$$\text{Sea Surface Height} = \text{Satellite Altitude} - \text{Altimeter Range} - \text{Corrections} \quad (2.37)$$

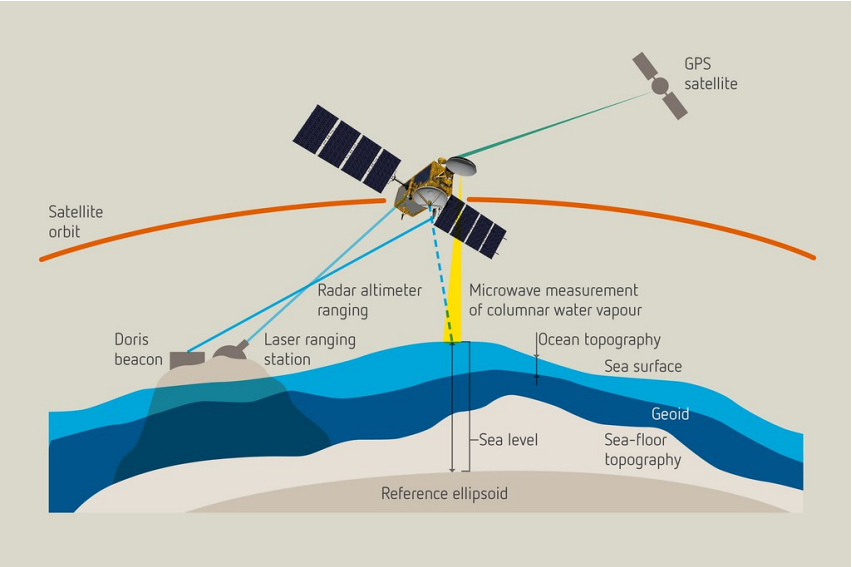


Figure 2.17: Altimeter emits a signal to Earth, and receive the echo from the sea surface, after its reflection [22].

Significant wave height and wind speed

To accurately calculate the sea surface height, it is necessary to have proper calibration of the satellite's trajectory relative to the Earth. Systems like GPS are typically used to accurately calibrate the satellite with respect to ground stations with millimetric precision [23].
After deriving the satellite altitude, we need to address the measurement of the range: this is simply the distance from the moment of signal transmission to its epoch (e.g., the received signal has reached the midpoint of the leading edge of the waveform). Clearly, not all the received signal is used, but it is necessary to correctly associate the emitted pulse with the received signal to extract the correct data from the associated waveform. The antenna must be able to "open" to receive the signal at an exact moment, which is associated with the time the signal arrives at the leading edge; this process is called *retracking* [20, 21]. There are different approaches to deriving the associated waveforms, depending on the altimeter technology. The figure 2.18 shows us the wave form of the

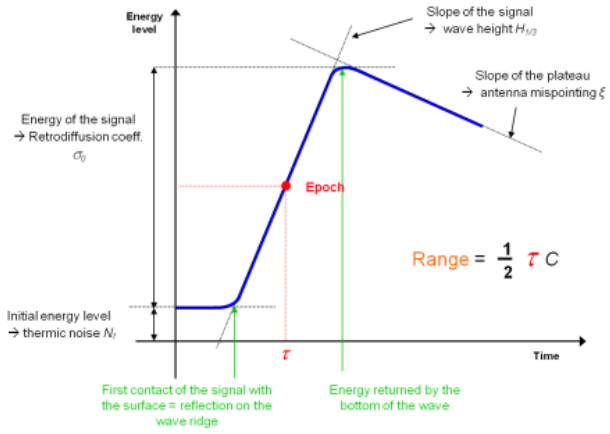


Figure 2.18: Wave form graph [6, 20].

echo reflected by the waves. There are some several characteristics of this graph, such as the leading edge slope of the curve relates to the significant wave height, the trailing edge slope relates to the mispointing of the antenna. The skeweness describes the curvature of the leading edge slope and then the epoch at mid-height gives us an indication of the time delay of the radar echo that is completely reflected back. Finally the P parameter is the backscatter coefficient which relates to the wind speed and then P_0 related to the thermal noise of the signal.
Wave information is determined from the slope of the leading edge of the altimetry waveform (figure 2.18). The approach to estimating significant wave height is quite simple. The difference in delay for reflections from wave crests and wave troughs affects the slope of the leading edge of the waveform, σ_c. This measure includes both the width of the original pulse σ_p and the standard deviation of the delays associated with reflecting facets, $2\sigma_h/c$ (where c is the speed of electro-magnetic waves) [24]. As the standard deviation of surface elevation is equivalent to one quarter of the H_s, the wave height can be determined via:

$$(H_s)^2 = 4c^2(\sigma_c^2 - \sigma_p^2) \tag{2.38}$$

The backscatter coefficient P is quite important because is related to the wind speed: an algorithm has to be applied in order to calculate the wind speed. Here is presented a simple one [25]:

$$U_{10} = U_m + 1.4U_m^{0.096}exp(-0.32U_m^{1.096}) \tag{2.39}$$

where:

$$U_m = \begin{cases} \alpha - \beta\sigma_0 & \text{for } \sigma_0 \leq \sigma_b \\ \gamma exp(-\delta\sigma_0) & \text{for } \sigma_0 > \sigma_b \end{cases} \tag{2.40}$$

The values in the equation 2.40 are $\alpha = 46.5$, $\beta = 3.6$, $\gamma = 1690$, $\delta = 0.5$ and $\sigma_b = 10.917$ dB.This can be a typical algorithm which can be used to determine the wind speed from the backscatter coefficent [25].

Altimetry waveforms and operating modes

Depending on the type of surface which reflected back the echo pulse, the waveform it is different. The sea state surface elevation distribution impacts the speed at which the return signal is fully returned to the satellite [20, 19, 26, 21]. Hence, the significant wave height H_s over ocean surfaces is determined from the slope of the front in the radar altimeter wave form. The higher the waves, the more the returned signal is spread in time. Hence, a long delay between the first returns and a full signal return will result in a long shadow in the wave form, which then indicates a high sea state (fig 2.16) [20, 19, 26, 21].
Figures 2.19 and 2.20 demonstrate how the type of surface can affect the final waveform. A rough sea will not reflect back the impulse and then the energy of the signal is diffused. On the contrary, a flat surface like sea ice will bounce back the signal without dispersing it due to the reflection of the waves.

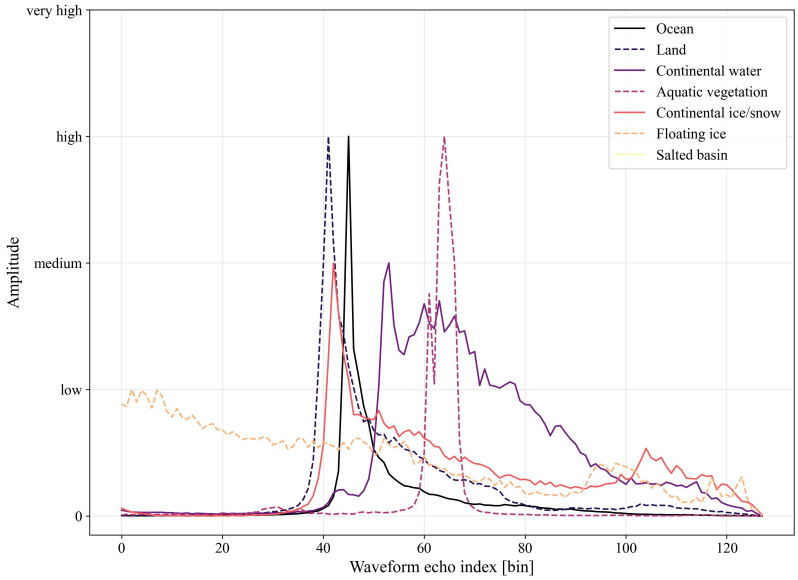


Figure 2.19: Type of surface reflection affects the wave form [24].

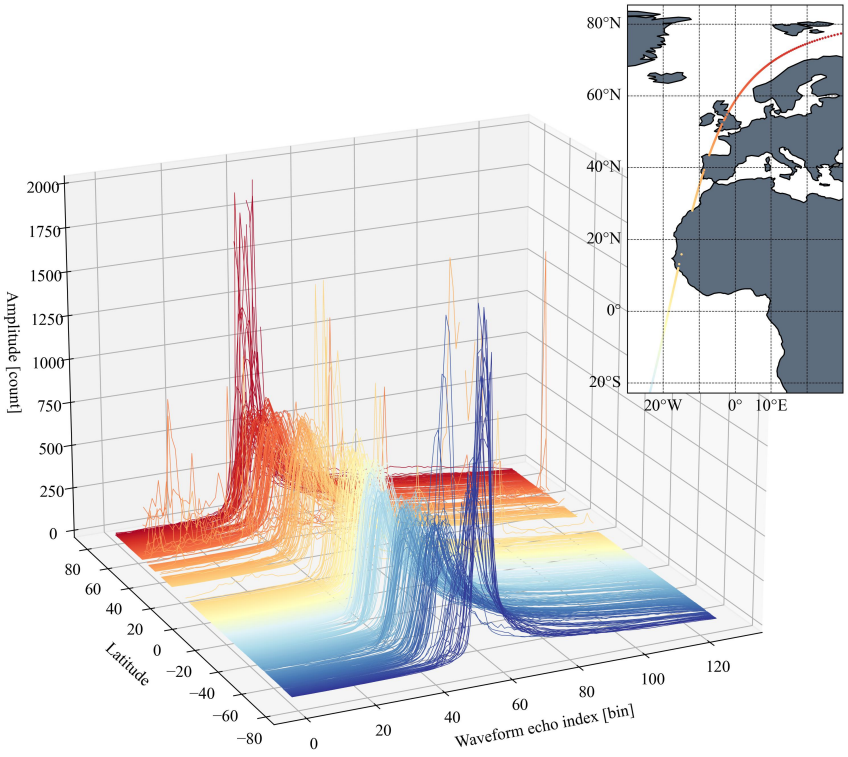


Figure 2.20: Waveform changes shape along the track, based on the type of surface [24].

The newest satellites (the last 10 years at least) present altimeters on board that can operate in more than one mode. This is the case for Sentinel-3 and Sentinel-6, which can operate in in two modes [27, 19]:

- High Resolution Mode, known as Synthetic Aperture Radar mode (SAR)

- Low Resolution Mode (LRM)

Depending on the mode, a different waveform can be achieved.
The LRM is the conventional altimetry mode, which operates by emitting a rapid series of pulses, each with a duration of 3.125 ns and based on 3 different pulse pattern 3 Ku,1 C,3 Ku (Ku and C are the bands on which pulses are emitted) [27, 19]. Every pulse is emitted at $\approx$ 2 kHz Pulse Repition Frequency (PRF). The necessary content energy cannot be squeezed into a single short pulse (could be some problems with electronics), instead the signal is emitted as a chirp with the same frequency content as the intended pulse (320 MHz) and the reflected signal 'deramped' by mixing with an identical chirp to localise the echo of the supposed pulse [24, 27, 19]. Lastly, in LRM mode the satellites emittes signals in nadir pointing (perpendicular to the Earth's surface).
SAR mode exploits the Doppler shift caused by the satellite's motion relative to Earth, by emitting a bursts of pulses: this technique was used for the first time on Cryosat [28] for detecting ice and then was introduced on Sentinel-3 [29, 30]. By moving along his track, with this technique we can 'fool' the sensor by virtually synthesizing a large antenna (aperture) from many small pulses [19].
The Sentinel-6A [31] has an improved SAR mode that is called SAR Interleaved [32]: basically the instrument can send bursts countinously (see figure 2.21 for comparing the different modes).This technology is quite important because can produce images with high resolution compared to the LRM mode. Another diffence between the two modes can be seen in figure 2.22, in terms of footprint and how the pulses are emitted,

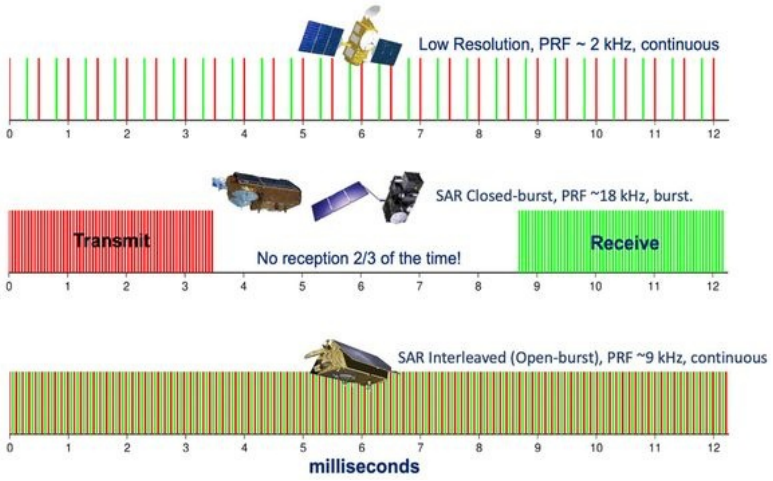


Figure 2.21: Comparison of measurement modes between TOPEX/Poseidon and Jason LRM, Sentinel-3/Cryosat SAR Closed-burst and Sentinel-6 SAR interleaved (Open-burst) [24, 19].

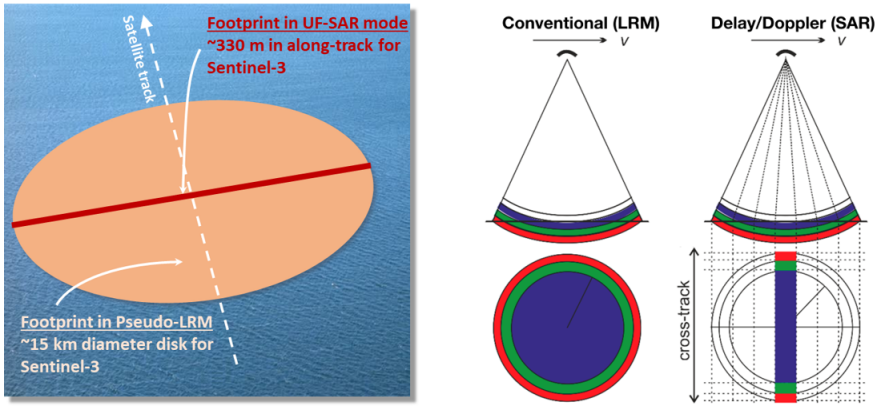


Figure 2.22: Illustration of Sentinel-3 SAR mode and P-LRM footprints over a sea surface [24, 19].

2.2.3 Processing Data Levels

Satellite data is released based on different processing levels, since we are still talking about the Copernicus programme, the processing levels are illustrated in table 2.4 (depending also on the type of mission/organisation, there could be also intermediate levels of processing).

Processing level	Description
Level 0	L0 is the raw telemetered data, dated and with initial geolocation
Level 1	L1 is the level 0 data corrected for instrument and geometric effects. This is broken into multiple steps, with associated products, eg level 1A, level 1B-S, and level 1B.
Level 2	L2 is the level 1 data corrected for geophysical effects. and including derived geophysical products
Level 2P	L2P homogenises the format of the data to match a set of consistent standards between all altimetry missions.
Level 3	L3 accumulates the level 2P products globally.
Level 4	L4 data have undergone extensive processing, potentially incorporating model outputs along with measurements from multiple satellites across several days. All input data have been validated.

Table 2.4: Levels of processing satellite data [33].

2.2.4 Future trends

In the last 30 years a large number of satellites have been deployed into the outer space, especially technlogy has evolved dramatically [7, 34]. Satellite altimetry keeps providing a large quantity of data that is also incredibly valuable for a wide range of applications.

If we look at long term period, future satellites will need to provide better spatial and temporal coverage so that we can study further near the coasts or mesoscale variations and other phenomena more closely. At the moment the best temporal resolution obtainable is 10 days by Sentinel-6A [19].
One of the ways to improve altimetry spatial resolution is to use several satellites at the same time. Until now, this has been done with very different types of satellite. The use of several satellites togheter and a cross calibration between all satellite missions could to reducing costs and improving quality of measurements (see figure 2.23 where having several satellite operating can help mapping out Earth surface in less time) [20]. This

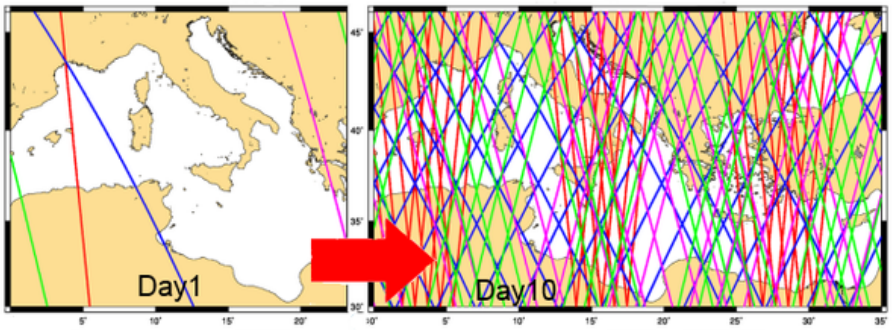


Figure 2.23: Grounds tracks for the 4 operational satellites on June 2015, over the Mediterranean Sea: Jason-2 ground tracks are blue, Saral/altiKa are green, Cryosat-2 are red and HY-2A are pink [20].

solution can increase drastically the temporal resolution, on the other hand the last decade space agencies had the urgency to test and build new instruments in order to obtain a higher spatial resolution on ground. Copernicus Sentinel-6 Michael Freilich was launched on 21 November 2020 and almost ready to "replace" the Jason-3, with big improvements on the altimetry side, by having the latest altimetry technology: figure 2.24 shows an improvement on the standard deviation of H_s with respect to *Jason-3* (i.e, statistical parameter that measure dispersion of a variabile relative to its mean). These two satellites they have the same orbit, so they have the same "view" of the ocean.

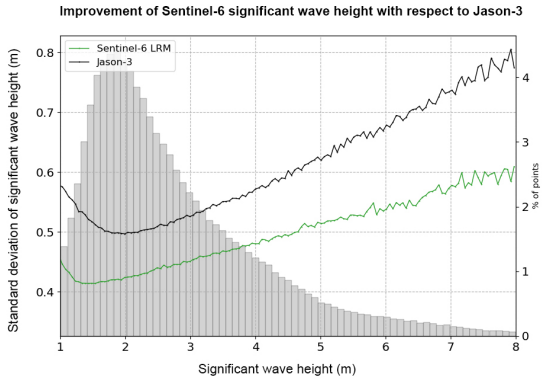


Figure 2.24: Improvement of S6 significant wave height with respect to J3 [26].

Chapter 3

Datasets description

In order to understand how calibration process works, satellite and in-situ datasets have to be discussed extensively: understanding how data is selected and preprocessed, the location of in-situ platforms, time domain used for the analysis, the reasons behind why a specific site has been selected. As for the satellite datasets, a brief description is presented regarding the temporal extent, which satellite data are used.

3.1 In-situ data

In-situ data are often used as a reference value, for accuracy puproses, for the calibration and quality validation of the altimetry data. In the scientific reasearch world, buoy data is identified as *ground truth* [35] because can be seen that they deliver the most accurated data, compared to other instruments. For this work, *mooring* platforms have been used as the main source of in-situ data delivering.
In the last years the network of in-situ platforms has grown enormously: there are many providers which give access to all type of in-situ instruments's data. Many scientific papers [36, 37] decided to make use of the U.S. National Data Buoy Center (NDBC) and the Copernicus Marine Service (CMEMS) [5] as providers for in-situ data. For this work all the mooring platforms have been selected on the In-Situ TAC,which is is the component of the Copernicus Marine Service and guarantee a reliable access to a wide range of in situ data [38], [39].
CMEMS website gives access to plenty of products, in terms of which data can accessed and downloaded for any type of purpose, but for this thesis significant wave height and wind speed are the variables of interest [6]. As for the calibration and validation of satellite data, often in the scientific world reasearchers decided to make their analysis by selecting in-situ platforms located, most of the time in the Northern part of America [35].
In this work a different path has been taken, instead mooring platforms selected for the calibration process are located in the North-West-Shelf region. As it is mentioned in the quality information document [6], from 1973 to 2023 the number of wave platforms has grown enormously, especially in the last 20 years (see figure 3.1, whereas figure 3.2 to see spatial coverage of all platforms and their time coverage), indeed most of these platform are recentely deployed.
As mentioned before most of the past studies were made on the American continent therefore due to the lack of studies made on the European seas, a different path was taken.

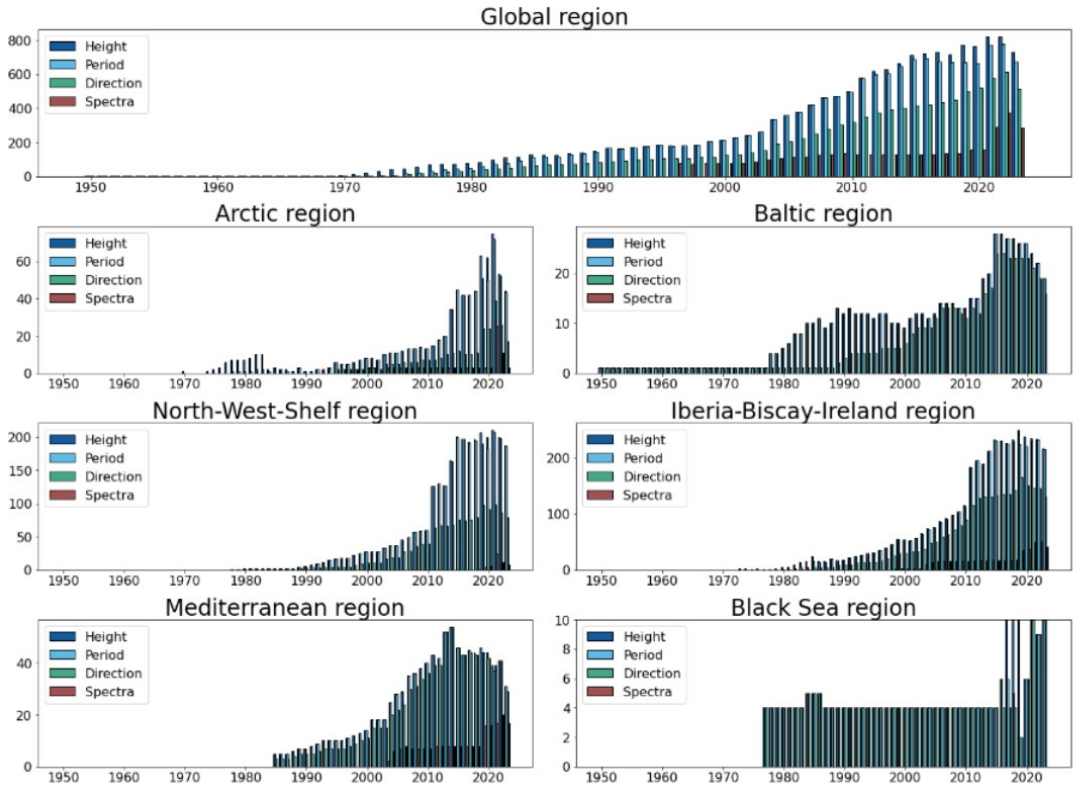


Figure 3.1: Evolution of the number of wave platforms from 1970 to 2023 at global scale and European seas [40].

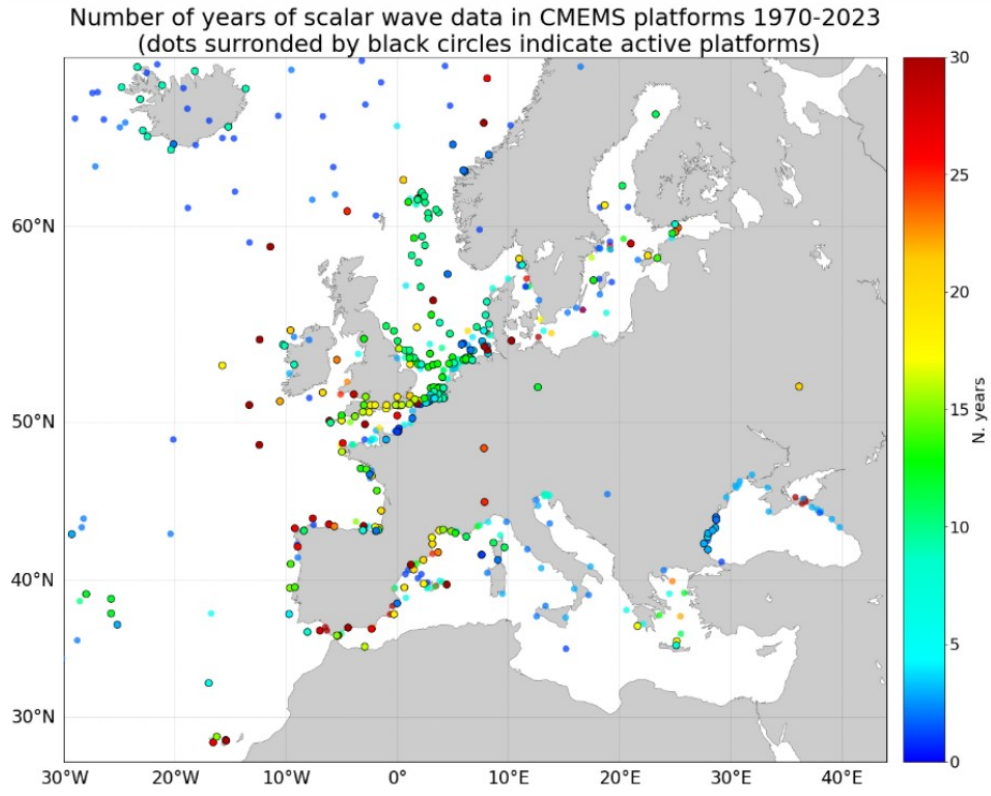


Figure 3.2: Wave data geospatial coverage at European seas, coloured by time coverage [40]

3.1.1 Description of the buoys

In the figure 3.3a are illustrated all the 200+ platforms which deliver data of wave and wind with a temporal coverage from 1st January 2021 to 31th December 2023: all platforms that have been downloaded are part of the following product from CMEMS Datastore INSITU_GLO_PHYBGCWAV_DISCRETE_MYNRT_013_030 [40]. Clearly not all the mooring are selected: only platforms that distance 100 km from the coast and have produced data at least in the last 5 years (from 2019 to 2024). This distance value ([41], [35],[6]) has been used in order to filter out satellite data near the coast. At the end of the filtering process,only 18 mooring platforms are filtered [42](see figure 3.3b). Not all

Platform	Start date	End date	Latitude	Longitude	Dcoast [km]	D. closest platf. [km]	Platform
Gullfaks-C	01/01/2014	On-going	2.26	61.20	124	19.94	Visundfeltet
Oseberg-A	05/01/2014	On-going	2.82	60.49	105	11.37	Oseberg-SOR
Oseberg-SOR	02/06/2014	On-going	2.79	60.39	109	11.37	Oseberg-A
6300112*	19/05/2011	On-going	1.0	61.1	100	49	Statfjord-A
Snorre-A	06/01/2014	On-going	2.14	61.44	131	9.20	Snorre-B
Snorre-B	06/01/2014	On-going	2.20	61.52	129	9.20	Snorre-A
Statfjord-A	01/01/1978	On-going	1.85	61.25	146	23.08	Gullfaks-C
Visundfeltet	24/04/2014	On-going	2.43	61.36	115	18.11	Snorre-A
A122*	06/07/2015	On-going	3.81	55.41	241	85.53	F3platform
F3platform*	06/01/2014	On-going	4.72	54.85	164	85.53	A122
J61*	19/05/2011	On-going	2.95	53.81	142	69.01	K13a3
K13a3*	19/05/2011	On-going	3.22	53.21	101	69.01	J61
Ekofisk	11/01/1980	On-going	3.22	56.54	262	100.23	6200146
Sleipner-A	01/01/2014	On-going	1.90	58.37	199	50.99	6200130
Granefeltet	09/04/2014	On-going	2.48	59.16	135	67.44	6300110
6200130*	07/07/2014	On-going	1.3	58.7	190	50.99	Sleipner-A
6200146*	07/07/2014	On-going	2.1	57.2	234	100.23	Ekofisk
6300110*	19/05/2011	On-going	1.5	59.5	157	67.44	Granefeltet

Table 3.1: Informations of each mooring filtered.

the mooring deliver data about the wind speed, in particular those with the asterisk are the ones that do not give that information while the others deliver both significant wave height and wind speed. Another difference between some of the moorings it is pointed out: some instruments deliver data in time differently compared to the other. More specifically "*6200130*", "*6200146*", "*6300110*" and "*6300112*" give data every hour whereas the others every 15 minutes (see figure 3.4 and 3.6a, 3.6b and 3.6c). Therefore temporal continuity was checked, whether big time gaps are present or not in the time series. As it was mentioned at the beginning, mooring platforms are not free of malfunctioning and for this reason some moorings have time gaps where they did not deliver data. In the figures 3.5a, 3.5b, 3.5c and 3.5d are represented the ΔT in days, between a i_{th} and $i_{th} + 1$ measurements in the time series: if the $\Delta T = 160$ days it means that 160 days occured from the i_{th} and $i_{th} + 1$ mooring index file and in that time gap there are not present any available measurements. Only 1-2 instruments presented big time gaps in the time series (figure 3.5a) whereas the others had continous time series with little time gaps. Just to remark that in-situ instruments are more vulnerable to periods of malfunctionalities.

(a) (b)

Figure 3.3: From left to right: all the 200+ mooring buoys have been filtered to 18.

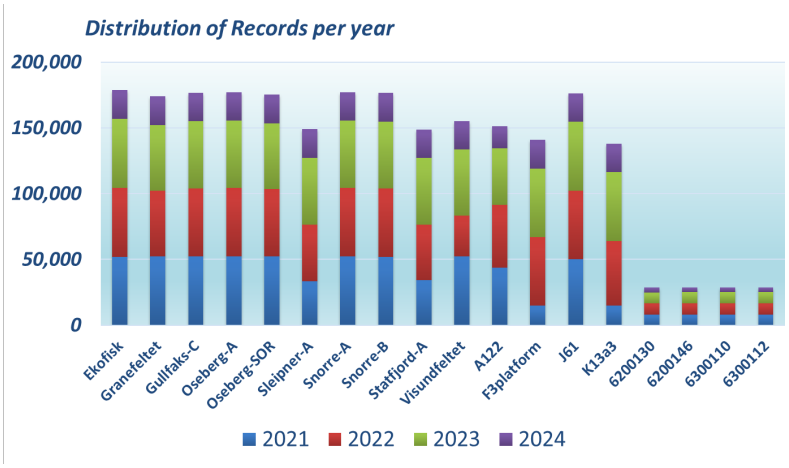
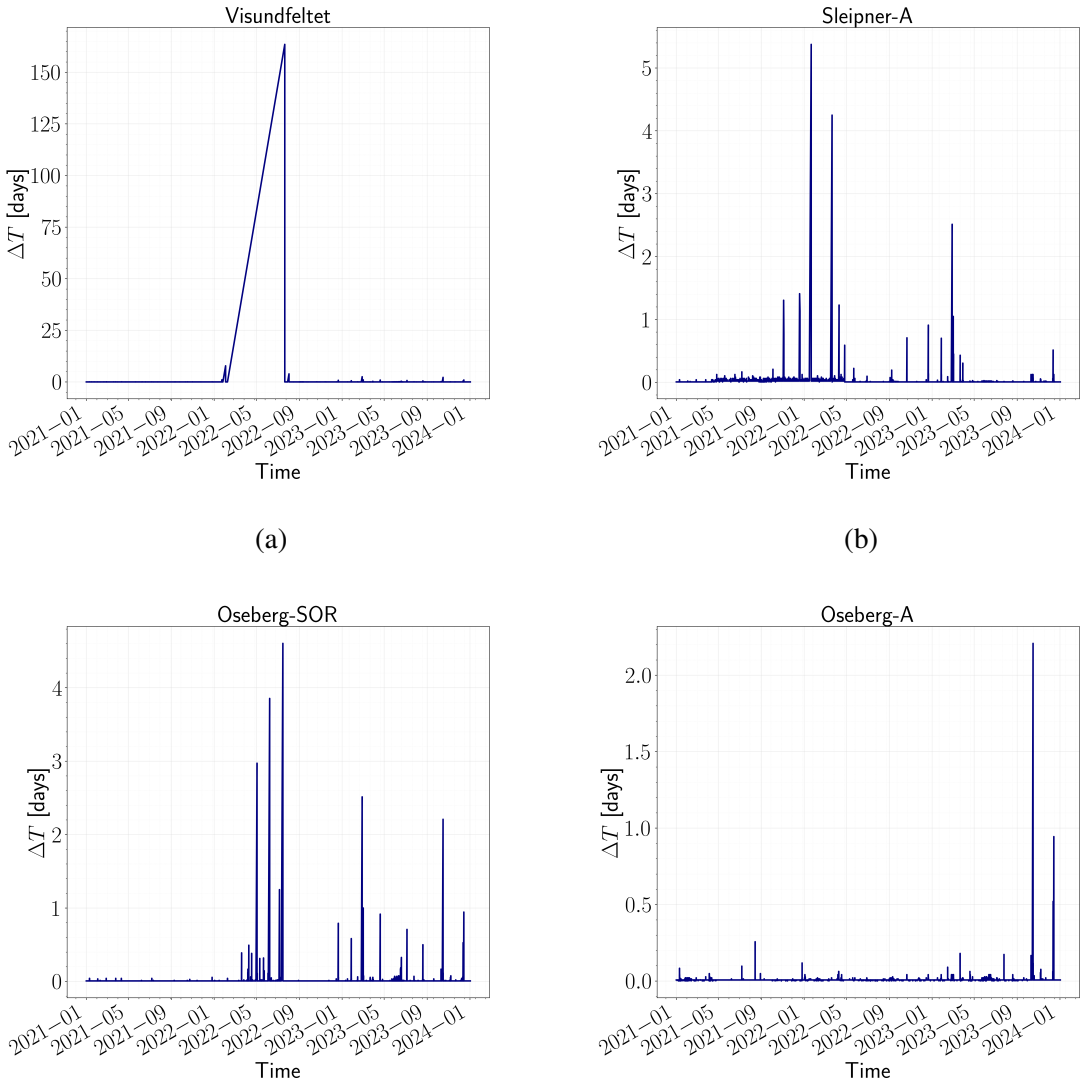


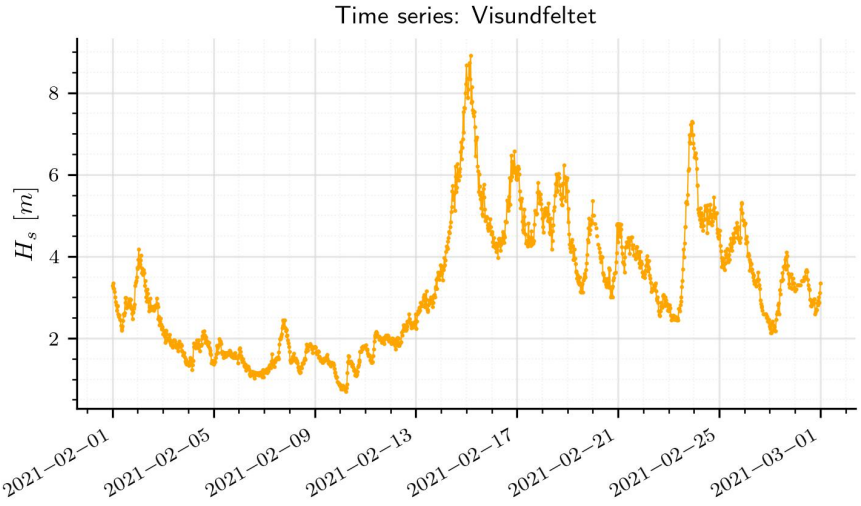
Figure 3.4: Distribution of the number's records per year for each mooring.



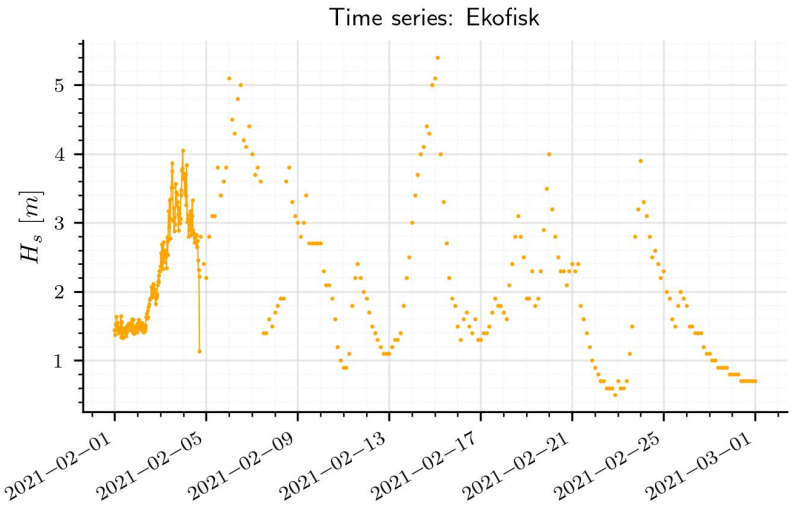
(a) (b)

(c) (d)

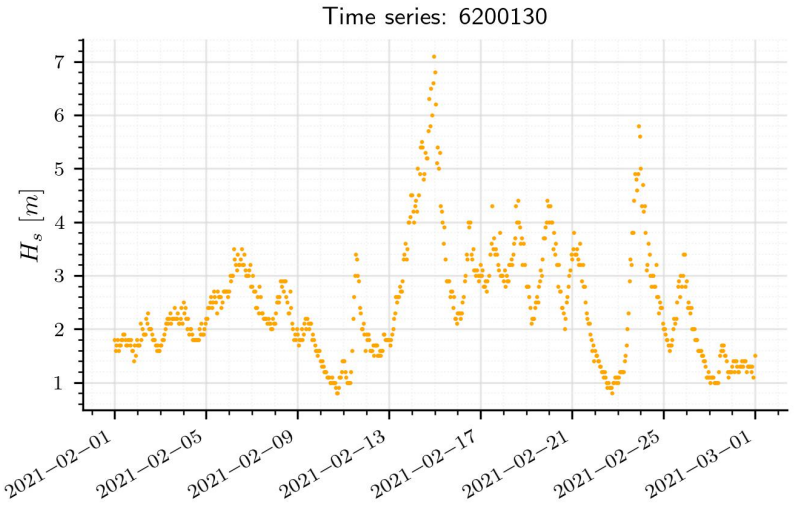
Figure 3.5: Delta times between observation t_{i+1} and t_i, for 4 different mooring buoys.



(a) Time series of H_s: the buoy delivers data every 10 minutes.



(b) Time series of H_s: the buoy delivers data every 10 minutes but after a specific time it starts to deliver data every 3 hour.



(c) Time series of H_s: the buoy delivers data every 1 hour.

3.1.2 Clustering

In this subsection the concept of *clustering* is introduced but before describing it, we anticipate that this approach it is not used in the following work and we are going to explain why.
The initial idea was to divide into smaller groups the 18 moorings, mainly because the arrangement of the platforms was quite unique. They are distributed homogeneously within the North Sea and in this was the sea state could have been described more precisely by the clusters. In a similar way, in the following paper [35] sea state has been categorized in 9 states based on different wave levels. The data is calibrated in every segment and then combined calibrated segments as the overall calibration results [35].

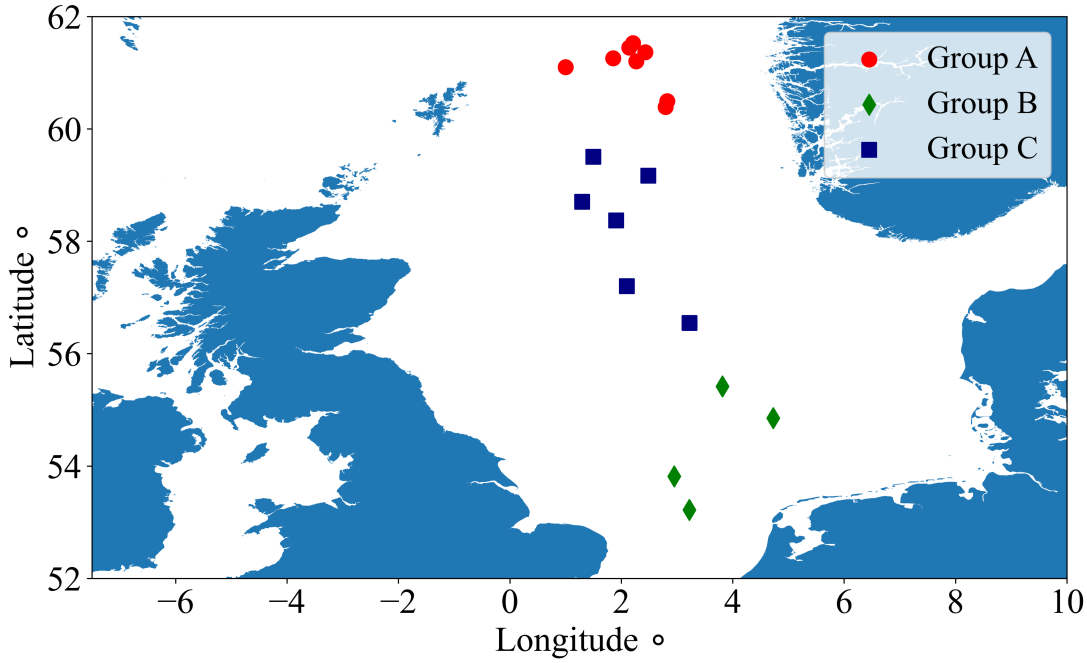


Figure 3.7: Clustering of the 18 buoys into 3 groups.

The 3 clusters can be seen in the figure 3.7.After making the division, the mean of every cluster is computed and compared to the moorings of the cluster: the difference in the time series are illustrated in the figure 3.8.
Then for every cluster the mooring in the middle of the group is picked: $\Delta H_s = H_{s_{buoy}} - H_{s_{ref}}$ is computed for every platform. As we can see from the figure 3.9, for the clusters A and B the differences are $\Delta H_s < 1$ but for the cluster C they are $\Delta H_s > 1$, reaching peaks where $\Delta H_s = 4\ m$-$5\ m$. After that for every cluster then correlation coefficient CC and root-mean-squared error $RMSE$ are computed (see Appendix A for metric coefficients). Figures 3.10a, 3.10b, 3.10c, 3.10d, 3.10e and 3.10f show those coefficients computed respect to their reference cluster buoy: from these figures we can observe that $RMSE$ and CC are quite different among buoys from their respective cluster, indicating that this path is not applicable. Only for cluster A (figures 3.10a and 3.10b) we have comparable values among all buoys.

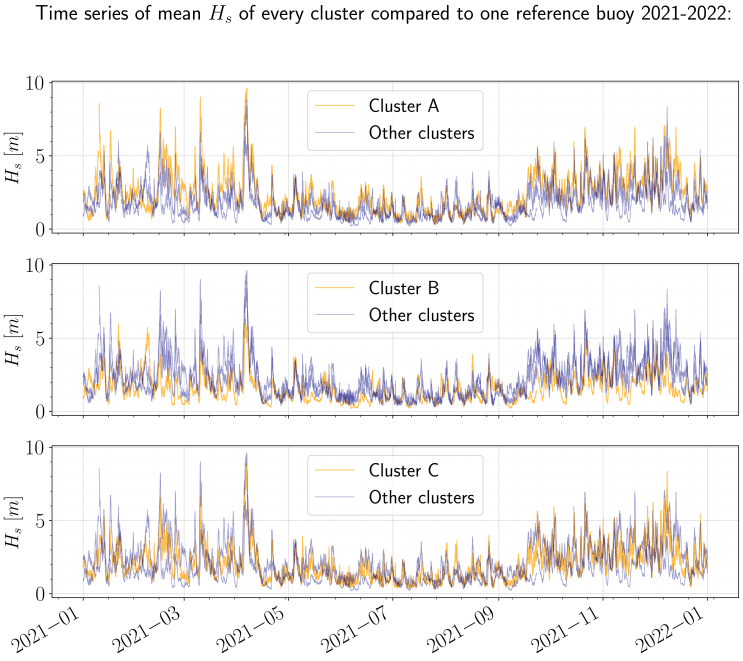


Figure 3.8: Comparison of the means of every cluster to others.

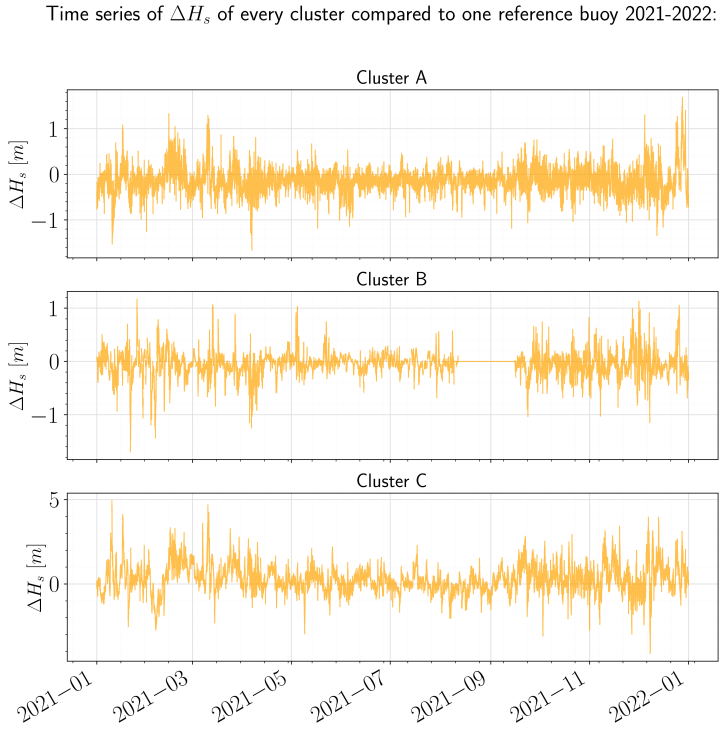


Figure 3.9: 3 subplots of the clusters of the difference between the i-th buoy and reference buoy.

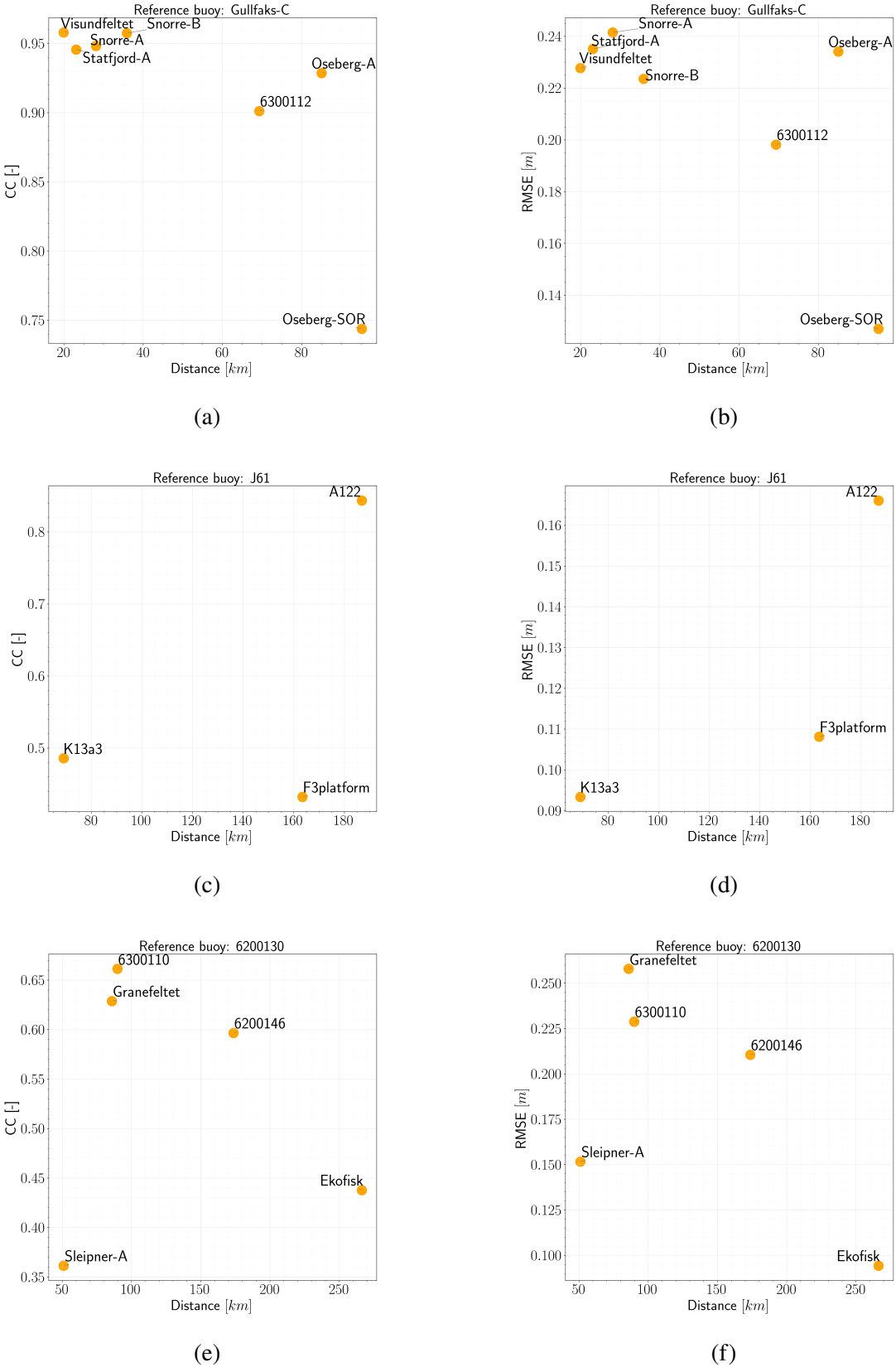


Figure 3.10: Correlation coefficient and RMSE computed for every cluster regard to the reference buoy.

3.2 Satellite data

3.2.1 Description of the dataset

Now that we have introduced the in-situ dataset, the satellite datasets will be presented. For the following spatial analysis, the product WAVE_GLO_PHY_SWH_L3_NRT_014_001 [42] has been selected from the CMEMS data store. This product is based on near-real-time (NRT) measurements with processing level L3 of significant wave height and wind speed. The following tables summarize the most important information about the entire dataset and the temporal availability for each satellite mission. More details can be found on the product page in the Copernicus data store [42].

Description	
Full name	Global Ocean L 3 Significant Wave Height From Nrt Satellite Measurements
Product ID	WAVE_GLO_PHY_SWH_L3_NRT_014_001
Variables	Significant Wave Height and Wind Speed at 10-m
Spatial extent	Global Ocean
Missions	Sentinel-6A; Jason-3; Sentinel-3A; Sentinel-3B; SARAL/AltiKa; Cryosat2; CFOSAT ; HaiYang-2B ; HaiYang-2C, SWOT nadir
Spatial resolution	Along-track ~ 7 km (full 1 Hz resolution)
Temporal resolution	Instantaneous

Table 3.2: Briefly description of the satellite data product [42].

Temporal availability			
Satellite	**Begin date**	**End date**	**Cross-over points***
Altika	01/01/2021	on-going	2231
C2	01/01/2021	on-going	1920
Cfosat	01/01/2021	on-going	2615
H2b	01/01/2021	on-going	2142
H2c	01/12/2022	on-going	1623
J3	01/01/2021	on-going	3407
S3a	01/01/2021	on-going	2556
S3b	01/01/2021	on-going	2757
S6a	21/09/2021	on-going	2632
Swon	01/08/2023	on-going	680

Table 3.3: Temporal availability of the different satellite missions.* Cross-over points have been calculated using a 50 km cross-radius [42].

As we can see from table 3.3, *H2c*, *S6a* and *Swon* have different temporal availability, but for the scope of this work, only *Swon* dataset will not be used, because has a limited time-frame.

Chapter 4

Methodology

In the following chapter the methodology used for the calibration process is presented: the idea is to calibrate the satellite data against in-situ platforms by following a common approach used in the literature which is *time-space matching* method that consists on making sure that satellite data and in-situ data are cross-matched [25, 35, 43, 41].
The scope of this part is to describe the calibration process and how different criterias (both spatial and temporal) can affect the final result. It is reminded that in this work, in-situ measurements are referred as ground truth data, as for the satellite measurements are referred as the satellite data.
In the first two sections, spatial and temporal analysis are illustrated and different criteria are used in these studies.Then these two approaches are combined to get a final spatio-temporal analysis, in order to select the most reliable data, base on the different criteria. During this process all the data is filtered also by checking quality flags: only good data was selected during this process (within the various datasets, both satellite and in-situ) in order to build the most robust and reliable dataset.
After showing the results of these analysis, 4 bias correction techniques are used to calibrate and validate data.
In the figure 4.1 is presented the flow-chart of the methodology used in this case study .

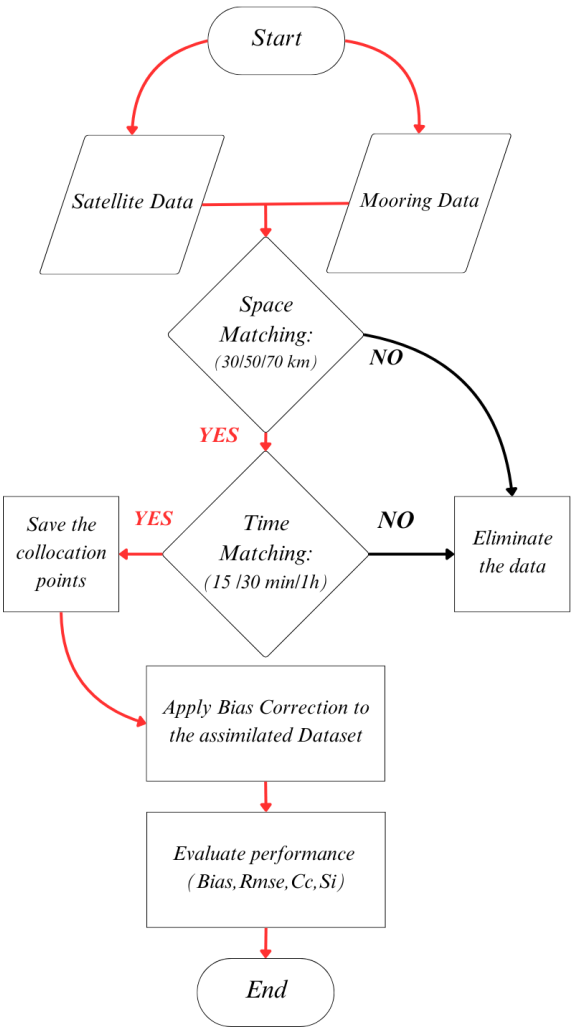
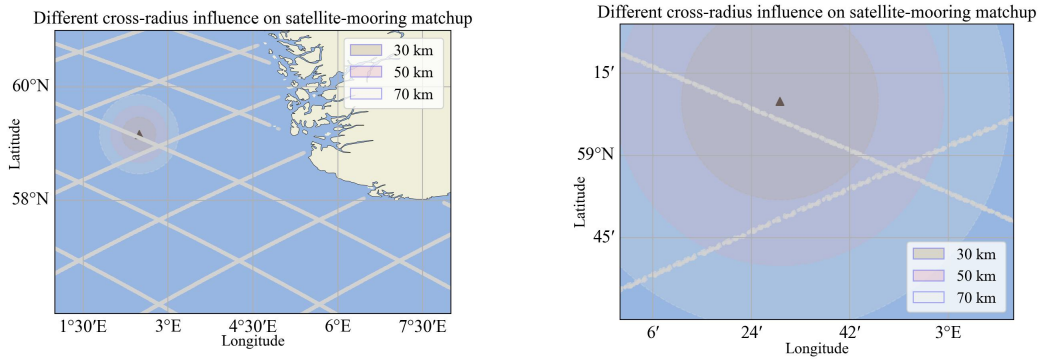


Figure 4.1: Flow-chart of the methodology [35].

4.1 Spatio-temporal matching

4.1.1 Spatial matching

Spatial matching between satellite and mooring is searched. In the scientific literature the reference cross-radius which to consider a matchup is to consider a 50 km radius ([35],[41], [43]). Different cross-radius values are used in this analysis to look for any improvements in the calibration process: 30-50-70 km cross-radius are used in the analysis to examine accuracy of data close and distant from coast (see figure 4.2a). Having a bigger area of influence within the mooring, can result in more matchup cross-over points between the platform and satellite (see 4.2b).



(a) (b)

Figure 4.2: Sentinel-6A tracks points over the mooring, with different cross-radius.

Whether satellite passes within the mooring buoy influence, all the points of the satellite track are saved.
After this procedure, the collocation point from the track has to be selected. There are two criteria that have been used in this study: the closest track point to the mooring and Inverse Distance Weighting [43]. The first one is self-explanatory: within the track points the closest to the mooring is selected, whereas the second consists on evaluating the value z at postion x by doing a weighted mean of nearby n observations:

$$\hat{z}(x) = \frac{\sum_i^n w_i z_i}{\sum_i^n w_i} \tag{4.1}$$

where w_i are the weights at the position x_i:

$$w_i = |x - x_i|^{-\beta} \tag{4.2}$$

and β can equal to $\beta = -1$ or $\beta = -2$. For this analysis $\beta = -2$ has been selected. For bigger cross-radii therea are more track points within the influence of the mooring buoy (see figure 4.3), so can be beneficial when using the IDW approach.

4.1.2 Temporal matching

After space matching was checked, temporal matching between satellite and mooring measured has to be checked as well. As it was mentioned in the chapter 3, several mooring platforms deliver data in different timespan and based on this information different time

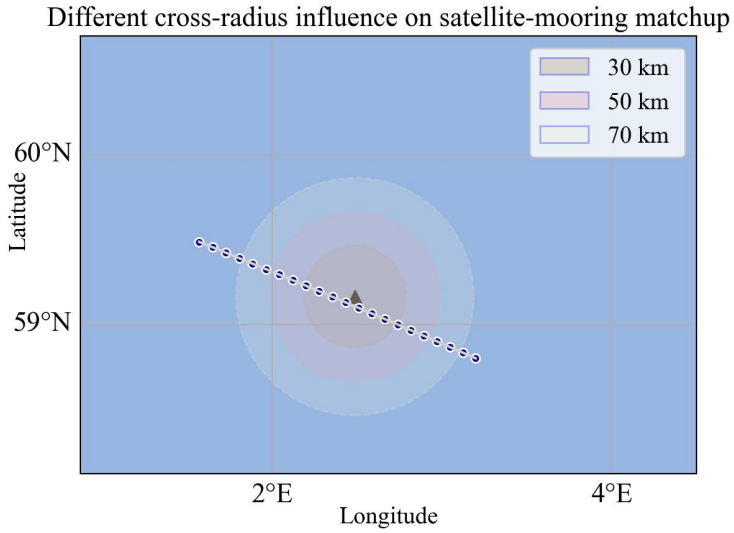


Figure 4.3: Satellite trajectory.

frames are used in this analysis for 2 reasons.
First of all, to compensate the different time-frames, using a wider time-frame can increase the number of collocation points. Second reason has to do with the problem of lack of data, as it was mentioned before in-situ platforms can malfunction quite often for many reasons. Increasing the time frame in this perspective can help but the quality of the produced data has to be checked. The analysis are done with timeframes of *15 min-30 min-60 min* [35, 43].
A wider time interval can produce more collocation points (since satellite and mooring crossings are matched in time), but this comes at the cost of less accurate data.
In the previous subsection, have been explained the criterions for the spatial matching between satellite and mooring platform, in a similar way has been done for the temporal matching. After the spatial matching, satellite measurement has to be 'synchronyzed' against in-situ time-series. It is checked that 30 (same for 15 minutes/ 1 hour) minutes before or 30 minutes after my satellite measurement, there is indeed a mooring measurement that is 'synchronized' [35, 43].
After that, two possible temporal criteria are applied: if a mooring measurement is present within the 30-minute window, the closest measurement in time is selected. With the second criterion, all mooring measurements within those 30 minutes are taken, and an average is calculated.
After the spatio-temporal matching analysis has been done, all the collocation points are saved and the final assimilated dataset is build. This procedure is done for every satellite dataset, as it was mentioned in the section *'Datasets description'*.
In order to assess the performance of every dataset [25, 35, 43, 41], based on different spatio-temporal cirteria that have been used, statistical parameters are used to firstly evaluate performance of every assimilated dataset (see Appendix A).
As shown in the table 3.3 are listed temporal availability for every satellite, and most of them have the same time span. In order to maintain a consistency in this work, spatio-temporal matching has been done from 01-01-2021 to 01-01-2024.

4.2 Bias-correction

Now that the assimilated datasets have been built, bias correction techinques are applied. These are just statistical methods that are used to adjust satellite outputs to match the observed data: these techniques reduce systematic errors (biases) when comparing them to observed or reference data [44, 45]. Usually they are used in fields weather, climate, or hydrology, where predictive models contain quite often biases due to simplifications and assumptions made during model development [46, 47].
These techniques can be applied to any variable and do not require any fundamental knowledge on the physics of the models or data assimilation methods. It is important to remark that the quality of the corrected datasets depends on the quality of the reference dataset considered as the ground truth (mooring observations) [44, 45].
For this work 4 different bias correction techniques are used: *Delta-Change*, *Linear Regression Calibration*, *Quantile Mapping* and *Full Distribution Mapping* [44, 45].

4.2.1 Delta-Change

Delta-change is the simplest among the BC techniques. It consists in adjusting the distribution of the assimilated dataset by summing a constant correction factor (the Delta factor) [44, 45]. Any data of the satellite dataset is corrected by this constant which is computed as the difference between the average values of the satellite dataset $\hat{y}^{assim}$ and the in-situ datasets $\hat{y}^{obs}$. Therefore, the dataset corrected via this technique BC is given as:

$$y_i^{BC} = y^{assim} + (\hat{y}^{obs} - \hat{y}^{assim}) = y^{assim} + \Delta \tag{4.3}$$

where $i = 1, .., N$ and N are the number of collocation points assimilated. It is reminded that for this work y represent H_s or U_{10}.
In the research world there are different version of the Delta-Change, such as Additive Delta-Change, Quantile-based Delta-Change and many others [44, 45].

4.2.2 Linear Regression

The following method it is not a conventional BC technique but is based on the linear regression, which is a statistical model used for estimating a linear relationship between a dependent variable (in this case the *ground truth*) and a regressor or independent variable (satellite data to adjust) [44, 45, 35].
In the literature there are plenty of different versions of this method, such as Multi Variate Linear Regression, Quadratic Regression.
Considering a simple linear regression:

$$y = a + bx \tag{4.4}$$

where a is the intercept and b is the slope of the straight. In this case we have N collocation points, where x_i are the true observations (in-situ) data and y_i are the satellite data. The following relationship can be established between those two, involving an additional term ϵ_i which is an error term:

$$y_i = a + bx_i + \epsilon_i \tag{4.5}$$

where:

$$\hat{\epsilon}_i = y_i - a - bx_i \tag{4.6}$$

is the residual error. The goal of this mathematical model is to find the best coefficients $\hat{a}, \hat{b}$ (are the estimated coefficients) so that the sum of squared residuals are minimazed. Given the objective function Q:

$$Q(a,b) = \sum_{i=1}^{N} \hat{\epsilon}_i^2 = \sum_{i=1}^{N} (y_i - a - bx_i)^2 \tag{4.7}$$

the main objective is to solve this minimization problem:

$$(\hat{a}, \hat{b}) = \arg\min(Q(a,b)) \tag{4.8}$$

For deriving $\hat{a}, \hat{b}$:

$$\hat{a} = \overline{y} - (\hat{b}\overline{x}) \tag{4.9}$$

$$\hat{b} = \frac{\sum_{i=1}^{N} \Delta x_i \Delta y_i}{\sum_{i=1}^{N} \Delta x_i^2} \tag{4.10}$$

where $\Delta x_i, \Delta y_i$ are the devations from the mean and $\overline{x}, \overline{y}$ are the mean values.
After finding these coefficients, bias corrected values are calculated as it follows:

$$y_i^{BC} = a + b y_i^{assim} \tag{4.11}$$

4.2.3 Quantile Mapping

Third BC technique used in this work is Quantile Mapping. This method is part of the distribution mapping techniques, which aim at alligning distribution of two datasets, shown in figure 4.4). In this case for every quantile q_j, the distribution of satellite data is alligned with the distribution of the observation dataset. This is done by adding a correction factor X^{QM} to each quantile, but before every dataset has to be divided into quantiles: in this work datasets have been divided into 10 linearly spaced quantiles [44, 45].
Therefore the correction factor is calculated as:

$$X^{QM}(q_j) = CDF_{assim}^{-1}(q_j) - CDF_{obs}^{-1}(q_j) \tag{4.12}$$

then this factor is applied at each quantile:

$$y^{BC}(q_j) = y^{assim}(q_j) + f(X^{QM}(q_j), n) \tag{4.13}$$

Equation 4.12 is then fitted with an n-order polynomial function in order to transform the correction factor X^{QM} into time-domain correction factors to be applied to the raw assimilated dataset [44, 45].

4.2.4 Full Distribution Mapping

Full Distribution Mapping technique is part of the distribution mapping techniques: this method, similarly to the Quantile Mapping, aims at alligning satellite dataset to the

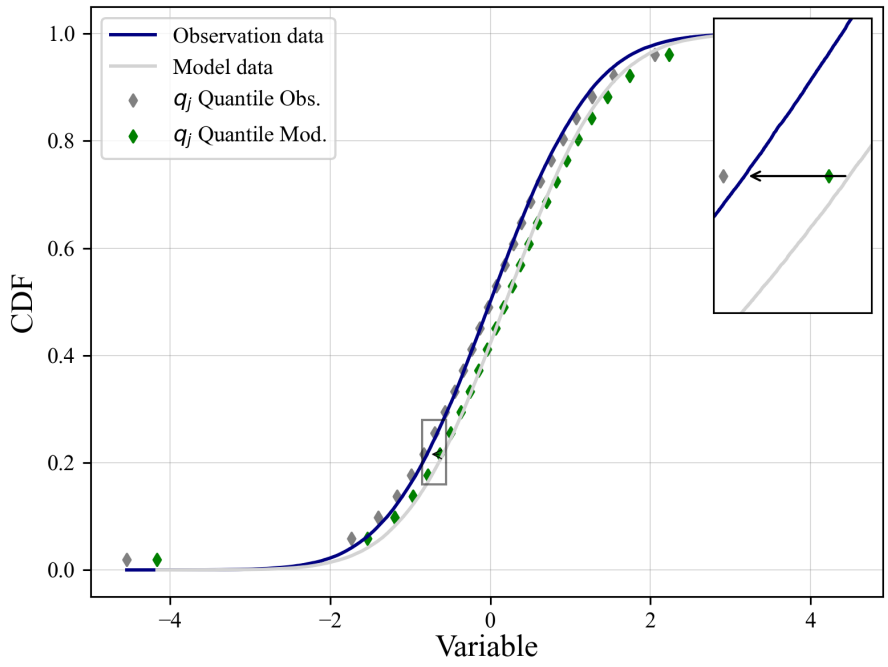


Figure 4.4: Schematic represantion of Quantile Mapping technique.

observation dataset by utilizing the whole CDF [44, 45]. Here the correction factor is computed as:

$$X^{FDM} = CDF^{-1}_{assim} - CDF^{-1}_{obs} \tag{4.14}$$

After calculating X^{FDM}, then it is interpolated with a polynomial of $n-$grade. Therefore this factor is applied to the full dataset [44, 45]:

$$y^{BC} = y^{assim} + f(X^{FDM}, n) \tag{4.15}$$

Chapter 5

Results

5.1 Spatio-temporal analysis results

5.1.1 Spatial criterions

Here are presented the results of the spatio-temporal matching analysis, with all the different criterions. This part is crucial because it shows how satellite datasets perform and how the different criteria affect these performance.
In the table 5.1 results of the spatial analysis with differents cross-radii are shown, for both H_s and U_{10}. For every cross-radii case, each metric coefficient is compared between Minimum Distance and IDW criterions, and the best metric has a colored cell (the different colors is to differentiate visually from each metric coefficient).
Increasing spatial gap between mooring platform and satellite tracks can bring more cross-over points but do not mean that the quality of satellite data improve (see figures 5.1a, 5.1b, 5.1c, 5.1d, 5.1e and 5.1f). Furthermore number of cross-over points at different cross-radii are shown in the table 5.2, where the increase of points from 30 km to 50 km is much higher than going from 50 km to 70 km. Besides that it is reminded that the number of platforms delivering H_s data is 18 platforms, while those delivering data of U_{10} are 10, indeed this is the reason behind the difference of number of cross-over points at same cross-radii.
We can see that the differences on 4 metrics parameters are not so big, but choosing a 50 km cross-radius bring the most accurate data, which is the standard value used in most of the calibration and validation studies [43, 25]. Other then cross-radius, spatial criterions such as IDW and Minimum Distance were proposed. It is reminded that for the Minimum Distance, closest satellite's track point to the mooring buoy is selected, whereas for IDW all satellite's track points are selected and an inverse averaging based on the distance from the mooring buoy is computed. It is interesting to see how IDW performs at different cross-radii because the number of track points within the influence of the mooring buoy is greater and having more points can influence averaging method results.
We can notice that at 30 km and 70 km, IDW performs better across almost all metric parameters, whereas Minimum Distance criterion is the best option when choosing a 50 km radius. As for the wind speed variable, IDW criterion works fine at almost all cross-radii, peforming quite good in all the metric coefficients.

Significant Wave Height: Jason-3								
	Minimum Distance				**IDW**			
Radius	**RMSE** $[m]$	**Bias** $[m]$	**CC**	**SI**	**RMSE** $[m]$	**Bias** $[m]$	**CC**	**SI**
30 km	0.3776	0.0971	0.9627	0.1712	0.3592	0.0900	0.9655	0.1635
50 km	0.3708	0.0834	0.9649	0.1675	0.3736	0.0832	0.9641	0.1690
70 km	0.3912	0.0886	0.9610	0.1749	0.3899	0.0867	0.9609	0.1744
Wind Speed: Jason-3								
Radius	**RMSE** $[m/s]$	**Bias** $[m/s]$	**CC**	**SI**	**RMSE** $[m/s]$	**Bias** $[m/s]$	**CC**	**SI**
30 km	2.5788	0.5737	0.8338	0.3125	2.6012	0.5487	0.8286	0.3120
50 km	2.6015	0.6077	0.8287	0.3143	2.5444	0.6262	0.8342	0.3076
70 km	2.6586	0.7262	0.8182	0.3267	2.6251	0.7459	0.8207	0.3222

Table 5.1: Results of the different cross-radii.

N. cross-over points: Significant wave height			
Dataset	**30 km**	**50 km**	**70 km**
Altika	773	1316	1870
C2	615	1061	1501
Cfosat	836	1605	2202
H2b	906	1515	1831
H2c	679	1149	1634
Jason3	1676	2919	4226
S3a	870	1488	2214
S3b	989	1434	2155
S6a	1107	2213	3168
N. cross-over points: Wind speed			
Dataset	**30 km**	**50 km**	**70 km**
Altika	485	850	1214
C2	377	677	963
Cfosat	-	-	-
H2b	718	1021	1224
H2c	456	751	1085
Jason3	1091	1930	2875
S3a	511	929	1378
S3b	603	915	1370
S6a	693	1490	2177

Table 5.2: Number of cross-over points at different cross-radii.

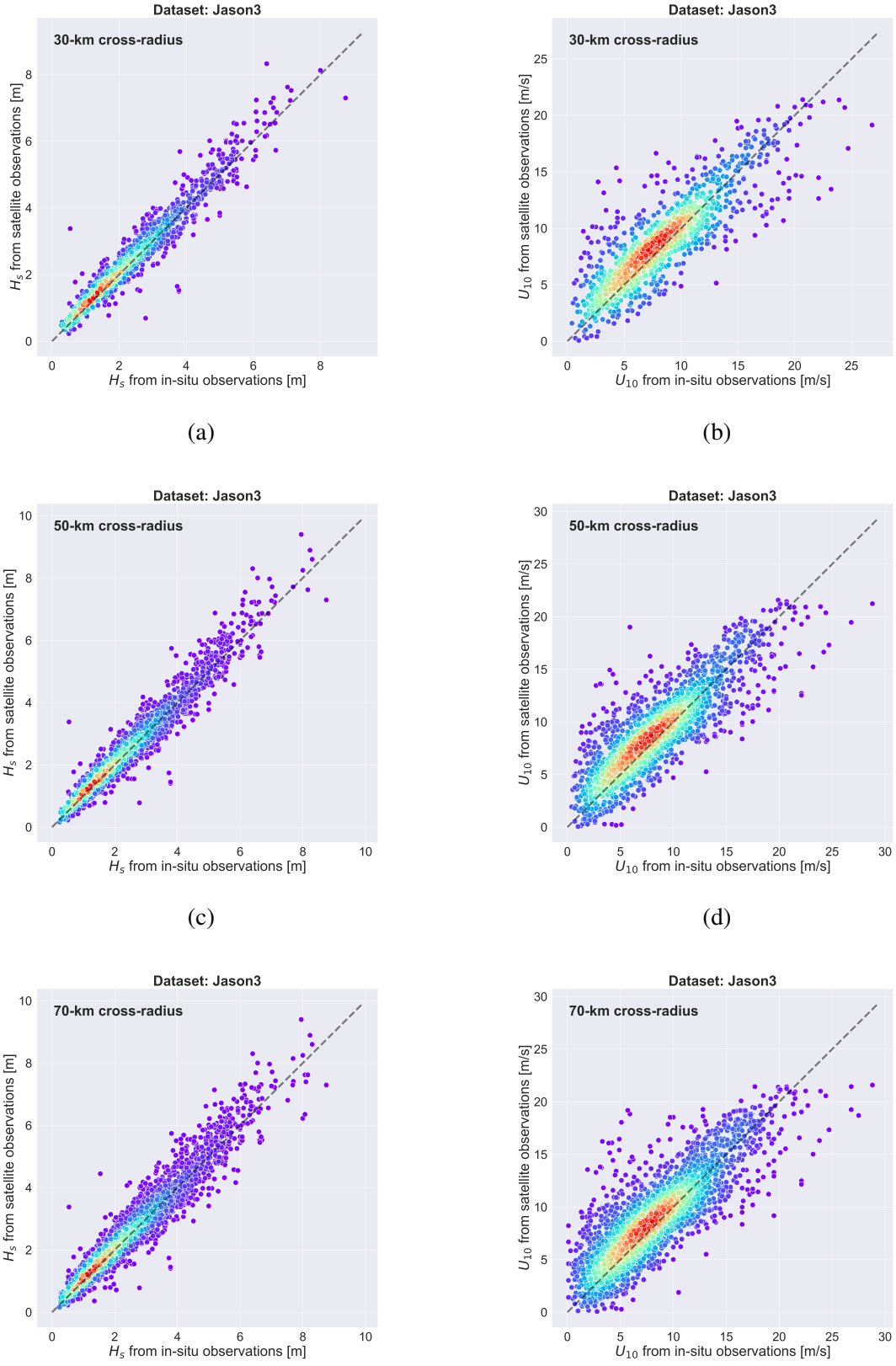


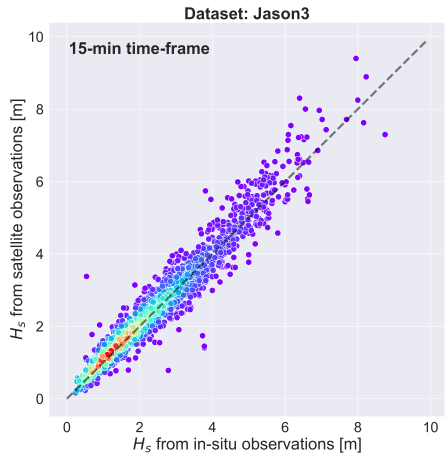
Figure 5.1: Scatter plots of *Jason-3* for different cross-radii, both H_s and U_{10}: hot colors in the graph show areas of high density points.

5.1.2 Temporal criterions

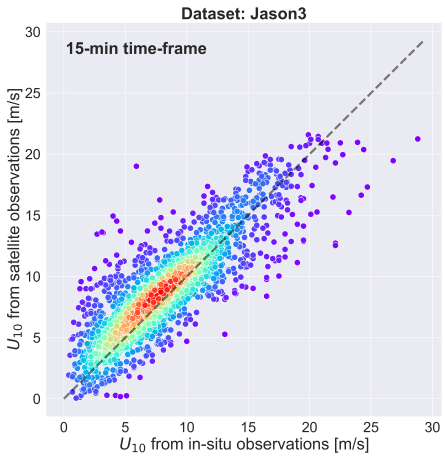
Similarly on what has been done for the spatial analysis, for the temporal analysis different time-frames are used to investigate data accuracy. In the table 5.3, for both H_s and U_{10}, 30 *min* time-frame seems to be the best option for this analysis, proving again to be the best criteria, similarly on what has been done in these studies [43, 35, 25]. For each time-frame, every metric coefficient for Clostest Observation case is compared to those of Mean Averaging case and the colored cells represent best ones for each time-frame (as mentioned before, color diversity is to visually divide results from all metric coeffcients). It is reminded Closest Observation means that first mooring buoy measurement available for the matchup is selected, whereas Mean Averaging takes all mooring buoy measurements available in that time-frame and makes an average of those values.
Clostest Observation criterions performs well in both H_s and U_{10}, whereas Mean Averaging tends to give better results for the bias metric.

	Significant Wave Height: Jason-3							
	Closest Observation				**Mean Averaging**			
Time-frame	**RMSE** $[m]$	**Bias** $[m]$	**CC**	**SI**	**RMSE** $[m]$	**Bias** $[m]$	**CC**	**SI**
15 min	0.3829	0.0820	0.9626	0.1747	0.7749	0.0754	0.8460	0.3478
30 min	0.3708	0.0834	0.9649	0.1675	0.6458	0.0757	0.8969	0.2865
60 min	0.3721	0.0855	0.9648	0.1687	0.4797	0.0788	0.9431	0.2121
	Wind Speed: Jason-3							
Time-frame	**RMSE** $[m/s]$	**Bias** $[m/s]$	**CC**	**SI**	**RMSE** $[m/s]$	**Bias** $[m/s]$	**CC**	**SI**
15 min	2.5917	0.6633	0.8308	0.3116	4.0030	0.6941	0.5945	0.4746
30 min	2.5444	0.6262	0.8342	0.3076	3.9842	0.5582	0.5828	0.4741
60 min	2.5661	0.6395	0.8320	0.3107	5.2583	0.1477	0.2866	0.6317

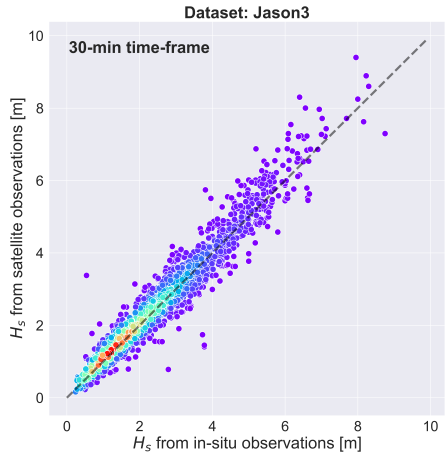
Table 5.3: Results of the different time-frames.



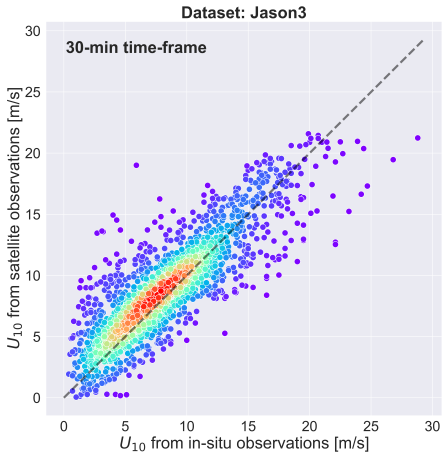
(a)



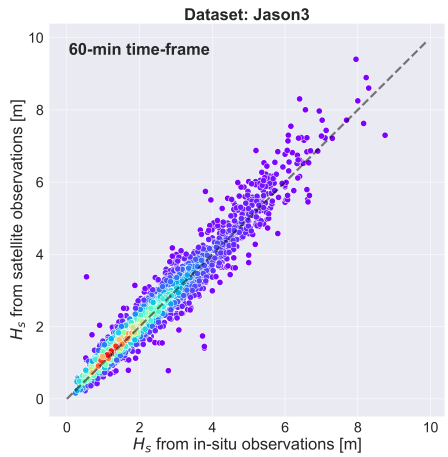
(b)



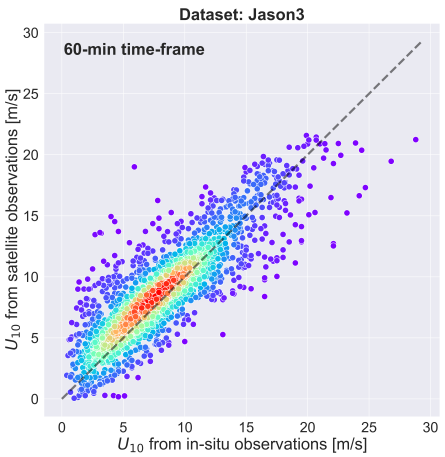
(c)



(d)



(e)



(f)

Figure 5.2: Scatter plots of *Jason-3* for different time-frames, both H_s and U_{10}: hot colors in the graph show areas of high density points.

5.1.3 Final matching metrics

All the results prove that best combination for the spatio-temporal matching analysis is 50 km and 30 min time-frames, according also to other studies [43, 25, 35]. In the tables 5.4 and 5.5 are shown final results of this methodology using best spatio-temporal criterions combination: same logic of colored cells as explained before in other tables is applied. In addition to that, Clostest Observation is the better option for temporal criterion (both wave and wind), whereas Minimum Distance seems to work well for wave height and IDW for wind.
Lastly, metrics performance for the wind speed are less good compared to the significant wave height, indicating the strong variability of such variable.

Significant Wave Height: spatial criterions

50 km / 30 min	**Minimum Distance**				**IDW**			
Dataset	**RMSE** $[m]$	**Bias** $[m]$	**CC**	**SI**	**RMSE** $[m]$	**Bias** $[m]$	**CC**	**SI**
Altika	0.3279	0.0925	0.9755	0.1509	0.3328	0.0917	0.9747	0.1533
C2	0.3838	0.0448	0.9650	0.1710	0.3805	0.0538	0.9659	0.1694
Cfosat	0.3285	0.1125	0.9693	0.1508	0.3350	0.1115	0.9675	0.1538
H2b	0.3175	0.0816	0.9781	0.1348	0.3267	0.0724	0.9763	0.1383
H2c	0.4444	0.1442	0.9488	0.2028	0.4487	0.1473	0.9482	0.2041
Jason-3	0.3708	0.0834	0.9649	0.1675	0.3736	0.0832	0.9641	0.1690
S3a	0.3701	0.0998	0.9711	0.1619	0.3657	0.0937	0.9716	0.1599
S3b	0.3030	0.0938	0.9767	0.1385	0.3062	0.0886	0.9758	0.1401
S6a	0.3999	0.1016	0.9594	0.1720	0.4039	0.1069	0.9589	0.1736

Wind Speed: spatial criterions

50 km / 30 min	**Minimum Distance**				**IDW**			
Dataset	**RMSE** $[m/s]$	**Bias** $[m/s]$	**CC**	**SI**	**RMSE** $[m/s]$	**Bias** $[m/s]$	**CC**	**SI**
Altika	2.2615	0.2386	0.8687	0.2856	2.2556	0.2396	0.8693	0.2843
C2	2.4853	0.0924	0.8561	0.2931	2.4700	0.1119	0.8576	0.2910
Cfosat*	-	-	-	-	-	-	-	-
H2b	2.7020	1.0288	0.8470	0.3196	2.6478	0.9841	0.8516	0.3126
H2c	2.6308	1.4285	0.8220	0.3575	2.5647	1.4109	0.8316	0.3469
Jason-3	2.6015	0.6077	0.8287	0.3143	2.5444	0.6262	0.8342	0.3076
S3a	2.6191	0.7514	0.8425	0.3128	2.6429	0.6394	0.8374	0.3162
S3b	2.3273	0.4209	0.8500	0.2706	2.3205	0.3427	0.8492	0.2708
S6a	2.6971	0.4926	0.8243	0.3146	2.6555	0.5300	0.8332	0.3091

Table 5.4: Results for different spatial criterions.

Significant Wave Height: temporal criterions

50 km / 30 min	**Mean Averaging**				**Closest observation**			
Dataset	**RMSE** $[m]$	**Bias** $[m]$	**CC**	**SI**	**RMSE** $[m]$	**Bias** $[m]$	**CC**	**SI**
Altika	0.6706	0.0837	0.8993	0.3019	0.3279	0.0925	0.9755	0.1509
C2	0.3523	0.0250	0.9720	0.1523	0.3838	0.0448	0.9650	0.1710
Cfosat	0.5489	0.1122	0.9149	0.2425	0.3285	0.1125	0.9693	0.1508
H2b	0.5835	0.0850	0.9298	0.2429	0.3175	0.0816	0.9781	0.1348
H2c	0.4069	0.1434	0.9605	0.1821	0.4444	0.1442	0.9488	0.2028
Jason-3	0.6458	0.0757	0.8969	0.2865	0.3708	0.0834	0.9649	0.1675
S3a	0.5830	0.1039	0.9321	0.2469	0.3701	0.0998	0.9711	0.1619
S3b	0.3830	0.0928	0.9637	0.1726	0.3030	0.0938	0.9767	0.1385
S6a	0.7075	0.0974	0.8753	0.2995	0.3999	0.1016	0.9594	0.1720

Wind Speed: temporal criterions

50 km / 30 min	**Mean Averaging**				**Closest observation**			
Dataset	**RMSE** $[m/s]$	**Bias** $[m/s]$	**CC**	**SI**	**RMSE** $[m/s]$	**Bias** $[m/s]$	**CC**	**SI**
Altika	3.0741	0.2536	0.7480	0.3875	2.2615	0.2386	0.8687	0.2856
C2	2.7846	0.0918	0.8091	0.3268	2.4853	0.0924	0.8561	0.2931
Cfosat*								
H2b	3.3302	1.0285	0.7496	0.3928	2.7020	1.0288	0.8470	0.3196
H2c	2.8835	1.4300	0.7679	0.3916	2.6308	1.4285	0.8220	0.3575
Jason-3	3.9842	0.5582	0.5828	0.4741	2.6015	0.6077	0.8287	0.3143
S3a	3.2504	0.7266	0.7517	0.3854	2.6191	0.7514	0.8425	0.3128
S3b	2.2612	0.3548	0.8611	0.2603	2.3273	0.4209	0.8500	0.2706
S6a	3.4784	0.6030	0.7005	0.4050	2.6971	0.4926	0.8243	0.3146

Table 5.5: Results for different temporal criterions.

5.2 Bias correction results

After the spatio-temporal matching method with the best criterions has been done, the final satellite datasets are ready to be calibrated and corrected. It is reminded that in this work 4 bias correction techinques are used.
It is a common practice when a bias correction is used, to divide the assimilated dataset into a "calibration" and "validation" datasets. In the first phase, it is where the effective calibration takes part: for this phase, the first 10 days of every month have been selected for the calibration dataset. Whereas for the "validation" part, remaining 20 days of every month have been selected to validate the bias correction technique used in the "calibration" part.

5.2.1 Delta technique and Linear regression

Here are presented the results of 2 first bias correction techniques: Delta method and calibration via Linear Regression. These two methods are compared against each other because they are similar techniques in a certain way: they both calibrate a dataset utilizing constant factor whereas the other techniques focus on adjusting data based on the CDF.
During "calibration" phase, Δ and $[a, b]$ have been computed respectibely for both techinques, therefore these correction factors are used to validate data and see how these techinques perform. It is reminded that Δ is the correction factor, computed as the difference between mean value of assimilated dataset and observation dataset.

In the figure 5.3 it is illustrated comparison between two techniques and can be seen

Significant wave height			
Dataset	Δ $[m]$	a $[m]$	b
Altika	-0.0958	-0.0560	0.9822
C2	-0.0345	-0.0642	1.0128
Cfosat	-0.1136	-0.0722	0.9815
H2b	-0.0936	-0.0489	0.9821
H2c	-0.1207	0.0221	0.9364
Jason-3	-0.0864	0.0461	0.9427
S3a	-0.1085	0.0435	0.9422
S3b	-0.0983	-0.0948	0.9983
S6a	-0.0888	-0.0173	0.9705
Wind speed			
Dataset	Δ $[m/s]$	a $[m/s]$	b
Altika	-0.5011	-0.1652	0.9535
C2	-0.5116	-0.2763	0.9729
Cfosat*	-	-	-
H2b	-0.8430	-1.3378	1.0511
H2c	-1.3801	-1.6513	1.0311
Jason-3	-0.5652	-0.0763	0.9443
S3a	-0.2430	-0.2469	1.0003
S3b	-0.6842	-0.4366	0.9698
S6a	-0.1946	-0.2245	1.0032

Table 5.6: Calibration coefficients.

the difference: Delta method adds a constant factor to all data and basically all the points are shifted downward, standing closer to the bisector. Whereas Linear calibration adds a constant factor and then slightly "rotates" all points towards the bisector (this is the effect of $a + by_i^{assim}$). This effect can be seen even more in the figure 5.4: because wind speed data is more sparse than H_s, the difference between techniques is highlighted more.

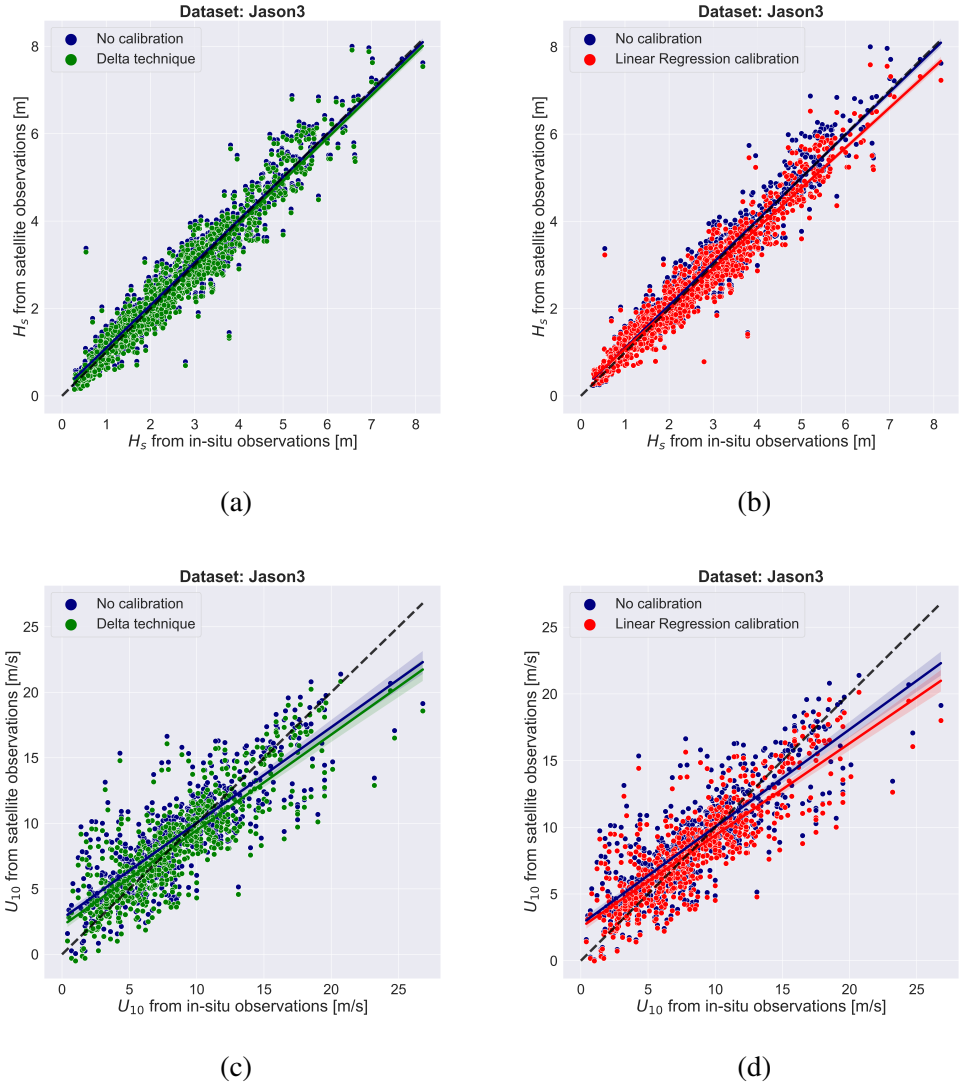


Figure 5.3: (a,b,c,d) are the Scatter Plots comparison of H_s and U_{10} for Jason-3 between Delta technique and Linear Calibration.

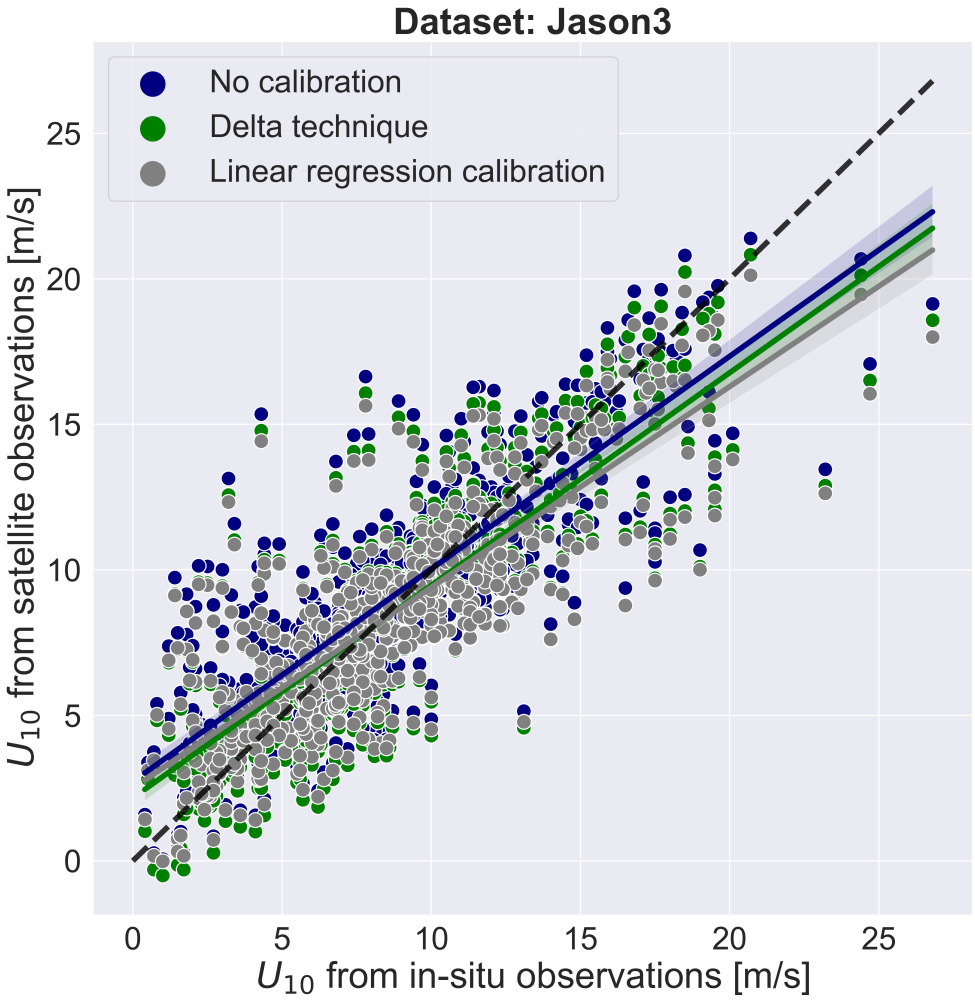


Figure 5.4: Another comparison between *Delta* and *Linear Regression calibration* techniques.

5.2.2 Quantile Mapping and Full Distribution Mapping

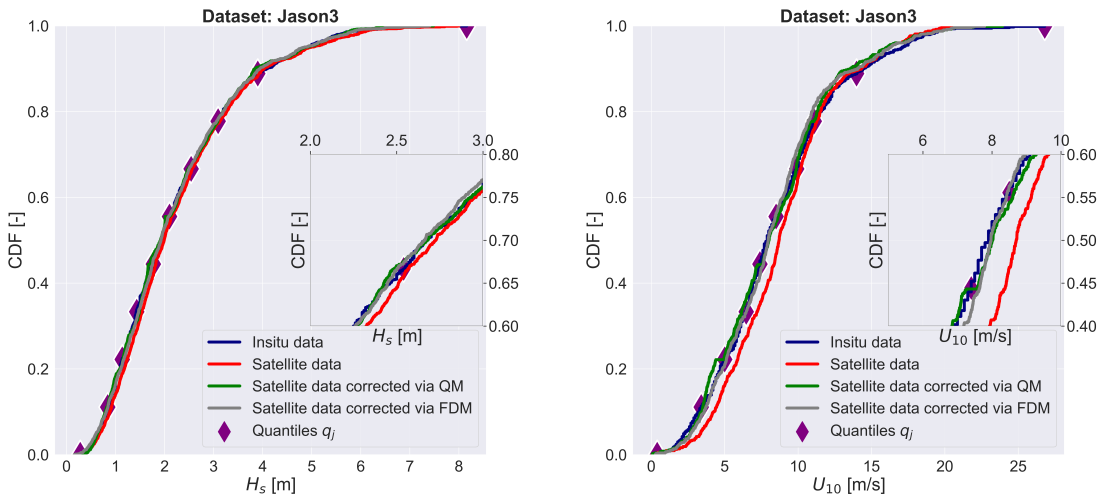
Here are presented results for the distribution mapping techniques: in the equations 4.13 and 4.14 correction factor X is interpolated with a polynomial function of order n. It is

n	QM	FDM
H_s	3	3
U_{10}	1	3

Table 5.7: n order of the interpolation polynomial.

reminded that in this work, for the QM technique only 10 linearly spaced quantiles have been used against the 50 used in this work [44]. Only in the QM technique, degree n of the polynomial function is 1 whereas the other cases is 3. The reason behind this choice stands in the lack of data. It was seen that using $n = 3$ for wind speed gives unrealistic metric results, bringing data to "explode".

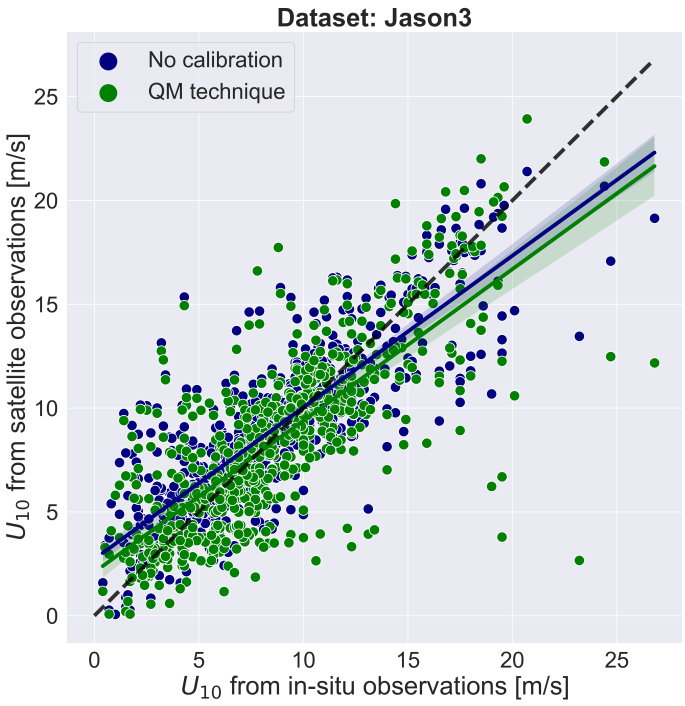
Similarly to previous techniques, here is presented in figures 5.6b and 5.6b a comparison



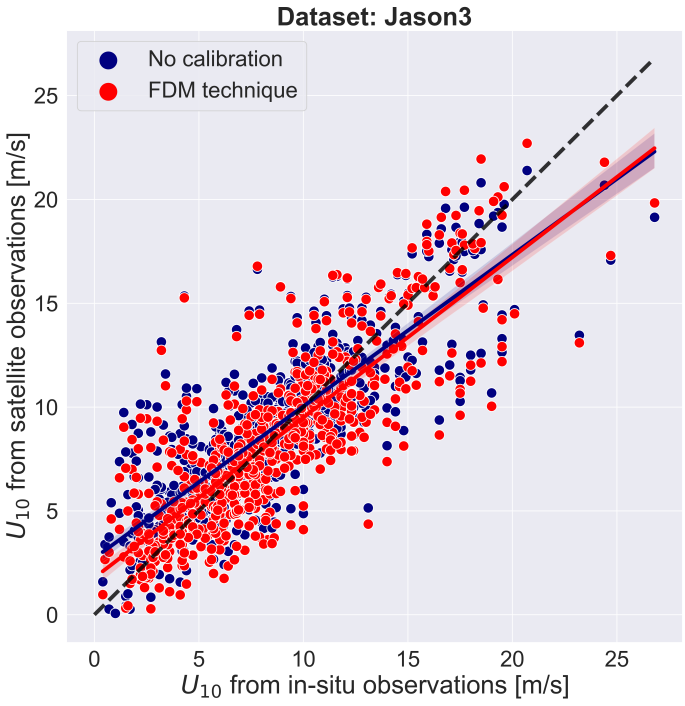
(a) (b)

Figure 5.5: CDF of both techniques compared to each other.

of the scatter plots of QM and FDM.



(a)



(b)

Figure 5.6: Scatter plots of both techniques compared to each other.

5.2.3 Final comparison

Now that all results are presented, a final comparison between all techniques is done to see which one performs better and suits for this specific applications. In table 5.8 are presented metrics results comparison for Jason-3, Altika and Sentinel-6A datasets In this case the logic behind colors of cells is different: for every metric coefficient, colored cells are the ones that represents the best result in that metric, among all BC techniques. Following this logic, for H_s of Jason-3, among all the correction, FDM produces the lowest bias among all techniques. This logic is applied to all other metrics, whereas the different color cell is to differentiate visually all coefficients.
Across all the bias correction techniques, FDM seems to perform better across all satellites datasets. Delta and Linear techniques do not change the distribution of data, whereas QM and FDM do therefore for this reason CC is always improved with last 2 methods.
For the wind part, Delta and Linear techniques perform almost in all the datasets, whereas

Significant wave height: *Jason-3*				
Technique	Bias $[m]$	RMSE $[m]$	CC	SI
No correction	0.070548	0.336538	0.968605	0.151888
Delta	-0.015863	0.329443	0.968605	0.148686
Linear	-0.014140	0.325947	0.968605	0.147108
FDM	-0.008884	0.323765	0.969035	0.146123
QM	-0.019095	0.322534	0.969289	0.145568

Significant wave height: *Altika*				
Technique	Bias $[m]$	RMSE $[m]$	CC	SI
No correction	0.079929	0.302390	0.978148	0.137455
Delta	-0.015947	0.292071	0.978148	0.132764
Linear	-0.016594	0.290112	0.978148	0.131874
FDM	-0.018939	0.289528	0.978437	0.131608
QM	-0.011540	0.306533	0.975682	0.139338

Wind speed: *Jason-3*				
Technique	Bias $[m/s]$	RMSE $[m/s]$	CC	SI
No correction	0.447781	2.483290	0.824621	0.293477
Delta	-0.117452	2.445407	0.824621	0.289000
Linear	-0.124243	2.432158	0.824621	0.287434
FDM	-0.165597	2.476804	0.825914	0.292711
QM	-0.195044	2.484800	0.827212	0.293656

Wind speed: *Altika*				
Technique	Bias $[m/s]$	RMSE $[m/s]$	CC	SI
No correction	0.253137	2.220850	0.864463	0.264350
Delta	-0.247991	2.220270	0.864463	0.264281
Linear	-0.314455	2.238011	0.864463	0.266393
FDM	-0.220076	2.341728	0.858806	0.278739
QM	-0.242094	2.409095	0.849513	0.286757

Significant wave height: *Sentinel-6A*				
Technique	Bias $[m]$	RMSE $[m]$	CC	SI
No correction	0.089827	0.368868	0.964629	0.157819
Delta	0.000961	0.357765	0.964629	0.153068
Linear	0.001054	0.352677	0.964629	0.150891
FDM	0.002438	0.356834	0.964882	0.152670
QM	0.001529	0.360068	0.964375	0.154053

Significant wave height: *H2b*				
Technique	Bias $[m]$	RMSE $[m]$	CC	SI
No correction	0.075522	0.319583	0.976623	0.137378
Delta	-0.018148	0.311061	0.976623	0.133714
Linear	-0.016268	0.307558	0.976623	0.132209
FDM	-0.019236	0.310692	0.976713	0.133556
QM	-0.031102	0.331640	0.974399	0.142561

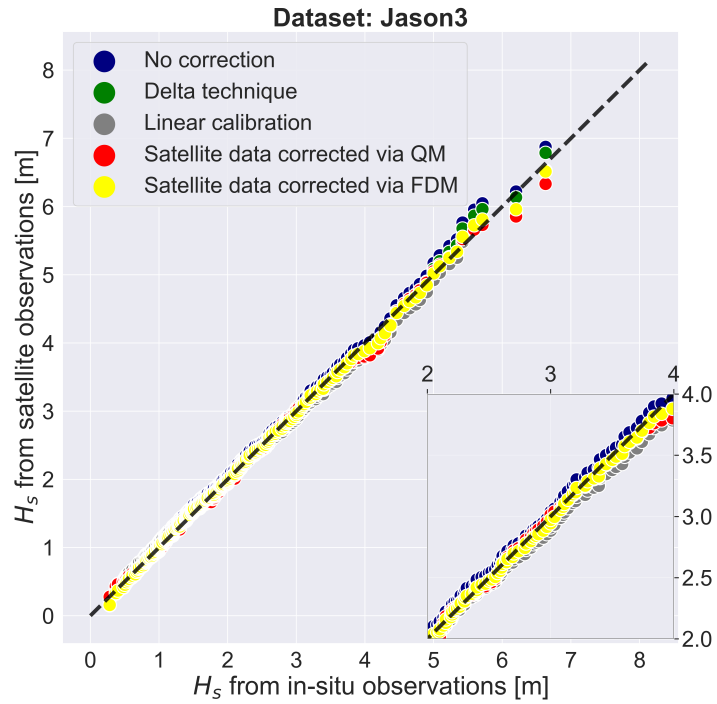
Wind speed: *Sentinel-6A*				
Technique	Bias $[m/s]$	RMSE $[m/s]$	CC	SI
No correction	0.267033	3.021507	0.794062	0.334869
Delta	0.072422	3.010555	0.794062	0.333656
Linear	0.072702	3.010254	0.794062	0.333622
FDM	0.019737	3.050159	0.794126	0.3380459
QM	-0.027888	3.075201	0.792315	0.340820

Wind speed: *H2b*				
Technique	Bias $[m/s]$	RMSE $[m/s]$	CC	SI
No correction	0.730285	2.522038	0.847983	0.291604
Delta	-0.112729	2.416623	0.847983	0.279416
Linear	-0.127950	2.445458	0.847983	0.282750
FDM	-0.133617	2.528572	0.852644	0.292359
QM	-0.177816	2.575535	0.846221	0.297789

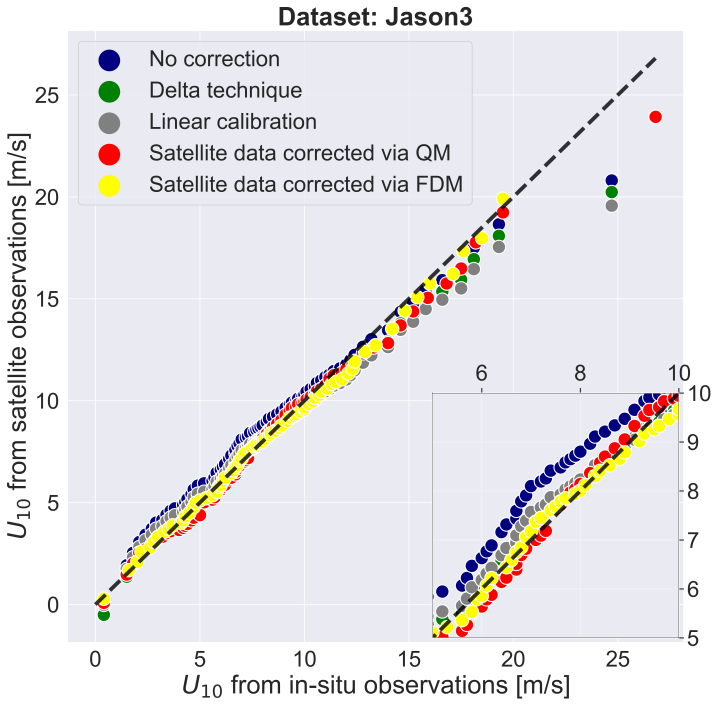
Table 5.8: Final metrics results and comparison between all techinques.

FDM and QM do not reach decent levels of improvement, also because to the lack of data. A final comparison between all techniques is presented with Quantile-Quantile plots (see figures 5.7a and 5.7b). Thanks to these graphs we can see how strong is relantionship between satellite and in-situ measurements, after calibration process. Due to the strong variability of wind speed, it is hard to see all the points aligning along bisector, whereas

for wave height strong relationship can be observed until 5-5.5 m.



(a)



(b)

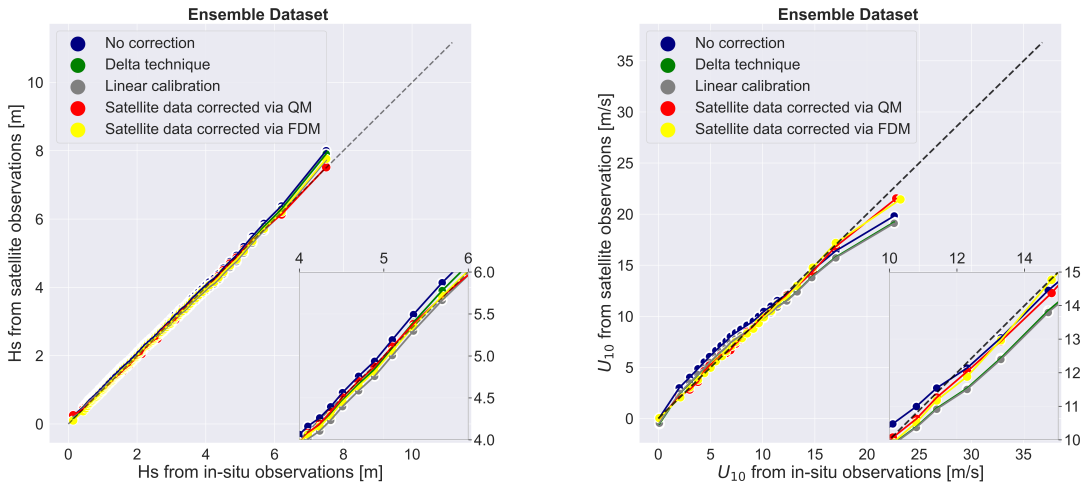
Figure 5.7: Quantile-Quantile comparison along all techniques (wave and wind).

5.2.4 Final ensemble

Now that all bias correction results have been presented, an ensemble is done, where all satellite datasets are assimilated into a final big dataset to see how effective all these techniques are.
Final performance results in table 5.9 show that there is not a predominant technique in all metric coefficients, but rather some patterns that occur: FDM is the best method where CC is improved almost all the times (for wave), whereas Linear calibration and Delta seem to be the best options to use to correct RMSE and Bias in wave and wind data. Same color logic, on what has been done in table 5.8, is applied here.

Significant wave height: *Ensemble*				
Technique	Bias $[m]$	RMSE $[m]$	CC	SI
No correction	0.08375	0.3336	0.9712	0.1482
Delta	-0.009901	0.3230	0.9712	0.1435
Linear	-0.009523	0.3197	0.9712	0.1421
FDM	-0.008956	0.3211	0.9713	0.1427
QM	-0.01418	0.3215	0.9714	0.1429
Wind speed: *Ensemble*				
Technique	Bias $[m/s]$	RMSE $[m/s]$	CC	SI
No correction	0.518596	2.556269	0.828595	0.300283
Delta	-0.084112	2.504525	0.828595	0.294204
Linear	-0.085499	2.502566	0.828595	0.293974
FDM	-0.105112	2.568031	0.830136	0.301664
QM	-0.155444	2.576730	0.830724	0.302686

Table 5.9: Metrics results for final ensemble.



(a) (b)

Figure 5.8: Quantile-Quantile comparison along all techniques (wave and wind).

Chapter 6

Discussions and conclusions

6.1 Spatio-temporal matching

The aim of this thesis was to analyze and valuate satellite altimetry data performance for offshore wave and wind applications, by following a spatio-temporal matching between satellite and in-situ instruments. In particular, all the mooring buoys are located in the North Sea, since the last 20 years, number of wave related platforms have grown greatly. Additionally, growth of ORE sector has brought a need of more fast and accurate data: since in-situ platforms give accurate data but are not so reliable, indeed satellite altimetry data plays an important role in this phase.
This analysis has brought interesting results in relation to cross-radius and time-frame and how they affect accuracy of altimetry data. Other than that, spatial and temporal criterions (beyond cross-radii and time-frame) such as Minimum Distance or IDW can affect final results depending also on which variable is analysed.
Three different cross-radii ($30, 50$ and $70\ km$) were used to investigate how this variable could affect accuracy and assimilation process, togheter with 2 spatial criterions were used to determine from the satellite track points which value had to be picked in the spatial dimension. All the mooring paltforms were selected based on the off-shore distance, to be at least $100\ km$ from the coast, in order to have accurate satellite data and interfered by coasts. For this reason, spatial matching analysis was used to investigate performance close to the coast.
Given the lower number of cross-over points at $30\ km$ compared to $50\ km$ cross radii (for *Cfosat*, 836 points vs 1605 points, almost the double) and the metric performance (RMSE comparison, $0.3337\ m$ vs $0.3285\ m$), $50\ km$ is the best cross-radius for altimetry data collocation analysis.
It is interesting to see that for $30\ km$ and $70\ km$ IDW criterion performs well for H_s and U_{10} (also at $50\ km$ for wind speed), whereas for $50\ km$ the Minimum Distance criterion seems to be the best option, considering metrics performance and number of collocation points. This analysis demonstrated that the standard cross-radius used for altimeter collocation method is still $50\ km$.
Along with spatial analysis, temporal influence to the collocation method is investigated. Similarly on what was done for spatial part, three different time-frames ($15, 30$ and 60 min) were used and two temporal criterions (Closest Observation and Mean Averaging) on which point to select were presented. For this work 18 mooring buoys have been selected and not all have same characteristics, such as different times averages: some of them have 10-min, while others have 1-hour. This temporal lag is linked to the difference between

different type of instrumentation used to record metocean data. In addition to this aspect, it is reminded that all 18 mooring buoys deliver continous data of H_s, while only 10 of them deliver measurements of the wind speed U_{10}.
The results of temporal matching analysis showed that 30 min time-frame and Closest Observation are the best choice for altimeter collocation methods, whereas Mean Averaging tends to produce less Bias at the cost of higher RMSE, SI and lower CC.
For the wave part, almost all datasets have comparable performance results, except for

	Significant wave height				**Wind speed**			
Dataset	RMSE $[m]$	Bias $[m]$	CC	SI	RMSE $[m/s]$	Bias $[m/s]$	CC	SI
Altika	0.3279	0.0925	0.9755	0.1509	2.2556	0.2396	0.8693	0.2843
C2	0.3838	0.0448	0.9650	0.1710	2.4700	0.1119	0.8576	0.2910
Cfosat	0.3285	0.1125	0.9693	0.1508	-	-	-	
H2b	0.3175	0.0816	0.9781	0.1348	2.6478	0.9841	0.8516	0.3126
H2c	0.4444	0.1442	0.9488	0.2028	2.5647	1.4109	0.8316	0.3469
Jason-3	0.3708	0.0834	0.9649	0.1675	2.5444	0.6262	0.8342	0.3076
S3a	0.3701	0.0998	0.9711	0.1619	2.6429	0.6394	0.8374	0.3162
S3b	0.3030	0.0938	0.9767	0.1385	2.3205	0.3427	0.8492	0.2708
S6a	0.3999	0.1016	0.9594	0.1720	2.6555	0.5300	0.8332	0.3091

Table 6.1: Final metrics results for all datasets.

H2c that produces a $CC = 0.9488$ and $RMSE = 0.444$. For the wind part, we have more sparse data with higher biases, where SI are much higher than the wave part and CC are not comparable with those for wave part as well. Lastly, in figures 6.1b, 6.1a, 6.1d and 6.1c scatter plots of both variables are compared, where *H2c* is less accurate.

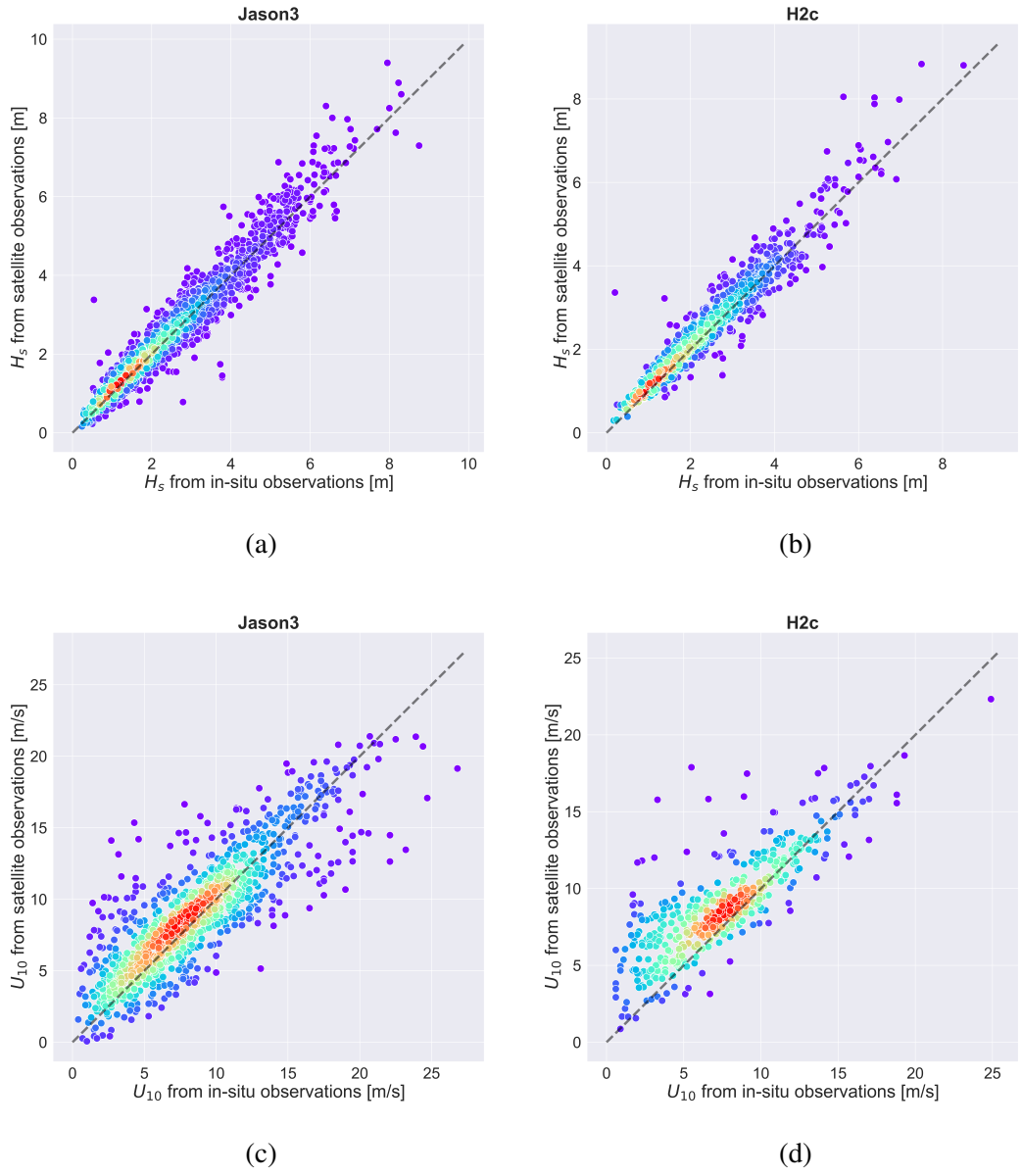


Figure 6.1: Scatter plots of *H2c* and *Jason-3* of H_s and U_{10} at comparison: hot colors in the graph show areas of high density points.

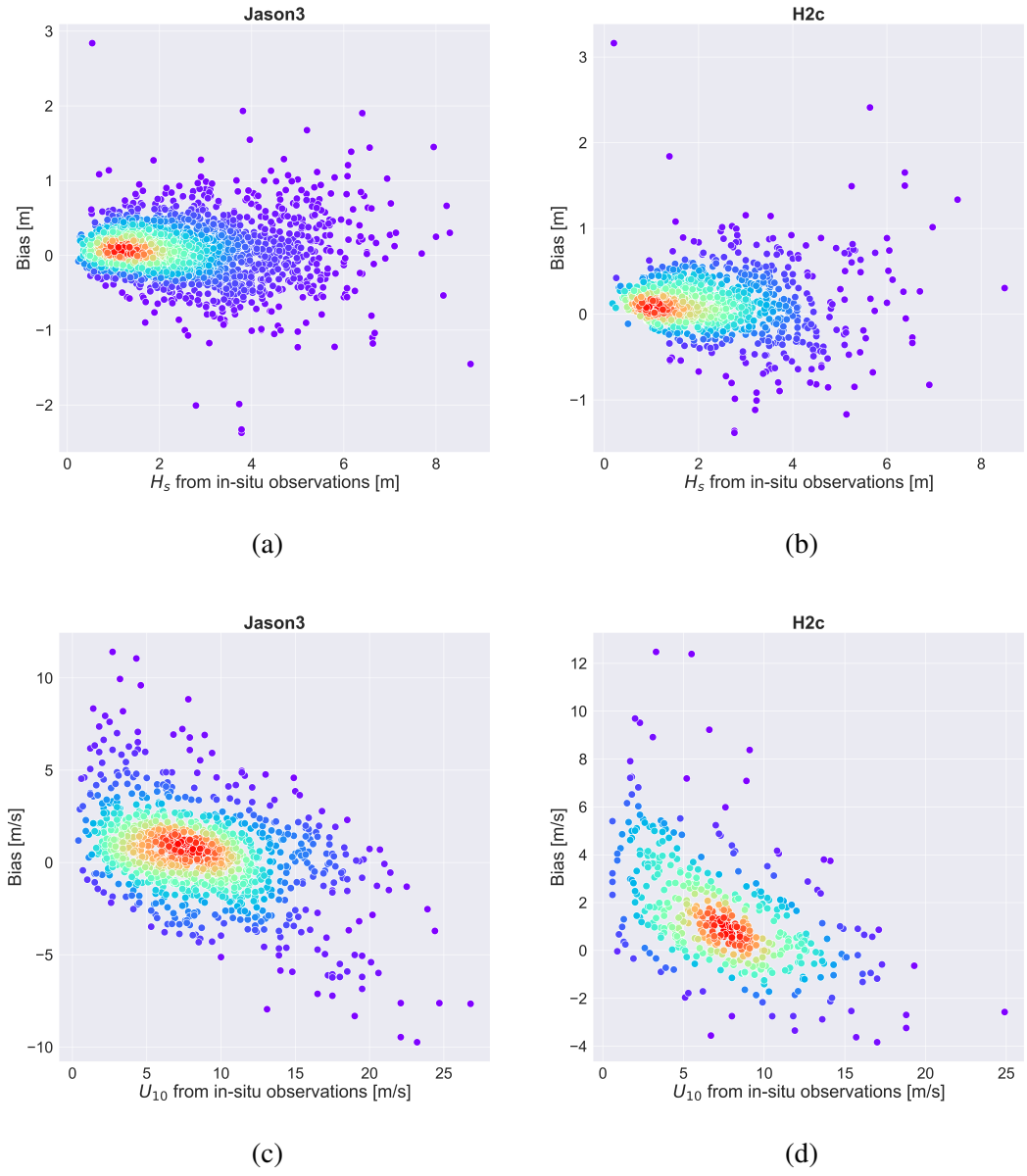


Figure 6.2: Comparison of Bias plots for *H2c* and *Jason-3*: hot colors in the graph show areas of high density points..

6.2 Bias correction

After the spatio-temporal matching analysis has been done, every dataset has been calibrated by several bias correction techniques. Throughout metrics coefficients, performance and effectiveness have been evaluated for every technique.
It is remineded that whenever a dataset is bias corrected, it is a common procedure to divide data into 2 parts: *calibration* dataset is used to determine the correction factors (based on the technique) and then *validation* dataset where those factors are used to validate the technique. For this work, all data have filtered from 1st January 2021 to 31th December 2023: more specifically, first 10-day of every month have been selected for the calibration phase, whereas remaining 20-days for validating the technique.
Tables 6.2 and 6.3 tell us that Delta Technique and Linear calibration work well for both

	Significant wave height				**Wind speed**			
Jason-3	RMSE $[m]$	Bias $[m]$	CC	SI	RMSE $[m/s]$	Bias $[m/s]$	CC	SI
No correction	0.3365	0.0705	0.9686	0.1518	2.4832	0.4477	0.8246	0.2934
Delta	0.3294	-0.0158	0.9686	0.1486	2.4454	-0.1174	0.8246	0.2890
Linear reg.	0.3259	-0.0141	0.9686	0.1471	2.4321	-0.1242	0.8246	0.2874
FDM	0.3237	-0.0088	0.9690	0.1461	2.4768	-0.1655	0.8259	0.2927
QM	0.3225	-0.0190	0.9692	0.1455	2.4848	-0.1950	0.8272	0.2936

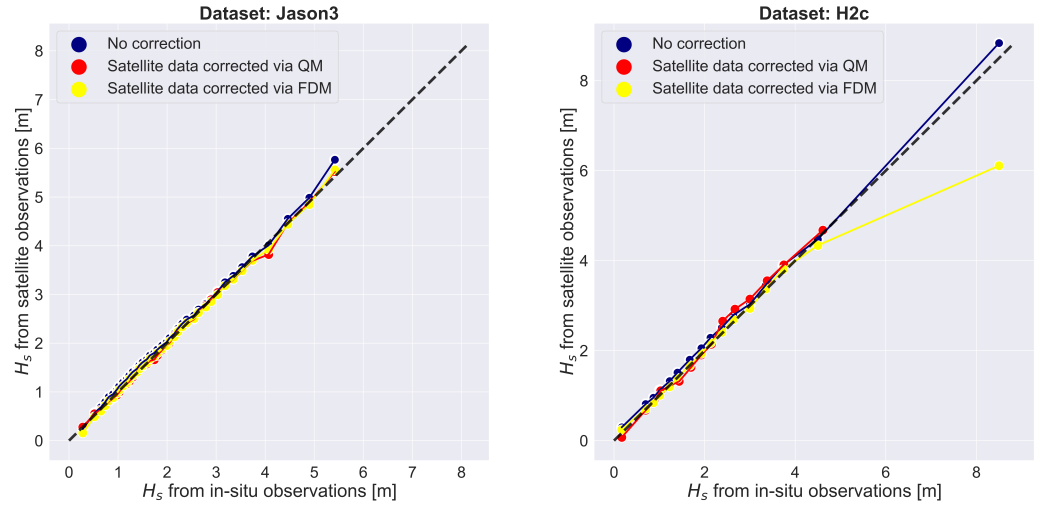
Table 6.2: Final metrics results for *Jason-3*.

	Significant wave height				**Wind speed**			
H2c	RMSE $[m]$	Bias $[m]$	CC	SI	RMSE $[m/s]$	Bias $[m/s]$	CC	SI
No correction	0.3691	0.1193	0.9679	0.1627	2.8061	1.5644	0.8040	0.3702
Delta	0.3493	-0.0014	0.9679	0.1540	2.3368	0.1843	0.8040	0.3083
Linear reg.	0.3402	-0.0102	0.9679	0.1500	2.3509	0.1981	0.8040	0.3102
FDM	0.3488	-0.0208	0.9664	0.1538	2.5741	0.1336	0.7936	0.3396
QM	0.5390	0.0399	0.9273	0.2376	2.5614	0.1431	0.8078	0.3379

Table 6.3: Final metrics results for *H2c*.

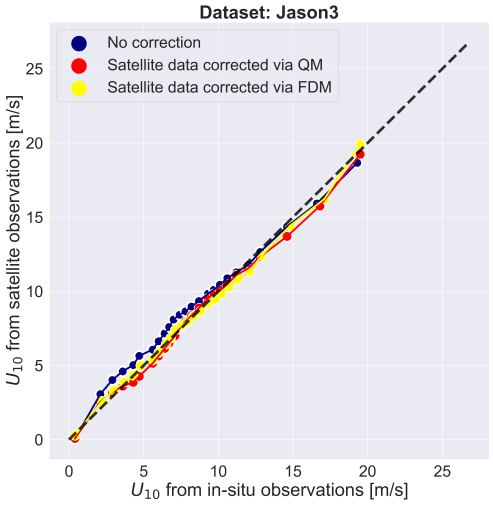
variables, improving in all metric coefficients (except CC). These results demonstrate that for simple bias correction, these techniques are suitable even for sparser data like wind speed. In these table *Jason-3* and *H2c* datasets have been compared because, first of all these two have a different number of collocation points (2919 vs 1149 for wave, whereas 2006 vs 782 for wind), and secondly *H2c* was the dataset that presented poor results in the spatio-temporal matching analysis. The number of data have definitely influenced distribution mapping techniques, indeed, FDM and QM improved performance of *Jason-3* even more than the other 2 techniques. Whereas in *H2c* metric coefficients got worse regard to Delta and Linear calibration.
Figures 6.3a and 6.3b show us how different is data distribution for H_s, confirming the lower accuracy for *H2c*, thanks also to a lower number of data. Even with distribution

mapping techniques, from 4.5 *m* to 9 *m*, the data deviates even more from the bisector, having some outliers whereas *Jason-3* data follows more precisely the bisector.
In figures 6.3c and 6.3d the effect of FDM and QM correction is more observable: after adjusting satellite data to in-situ observations, distribution of points is alligned more closely to the bisector.

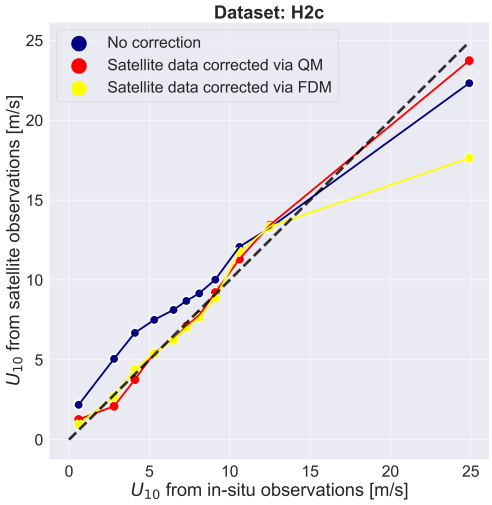


(a)

(b)



(c)



(d)

Figure 6.3: Q-Q plots for *H2c* and *Jason-3*.

At the end, all datasets have been grouped into a final ensemble. This procedure was done to valuate even more effectiveness of bias correction techniques, by having a bigger pool of data. For this purpose, in figure 6.4 Q-Q plot of wind speed is presented because, compared to H_s, it has several differences between the distributions of data.

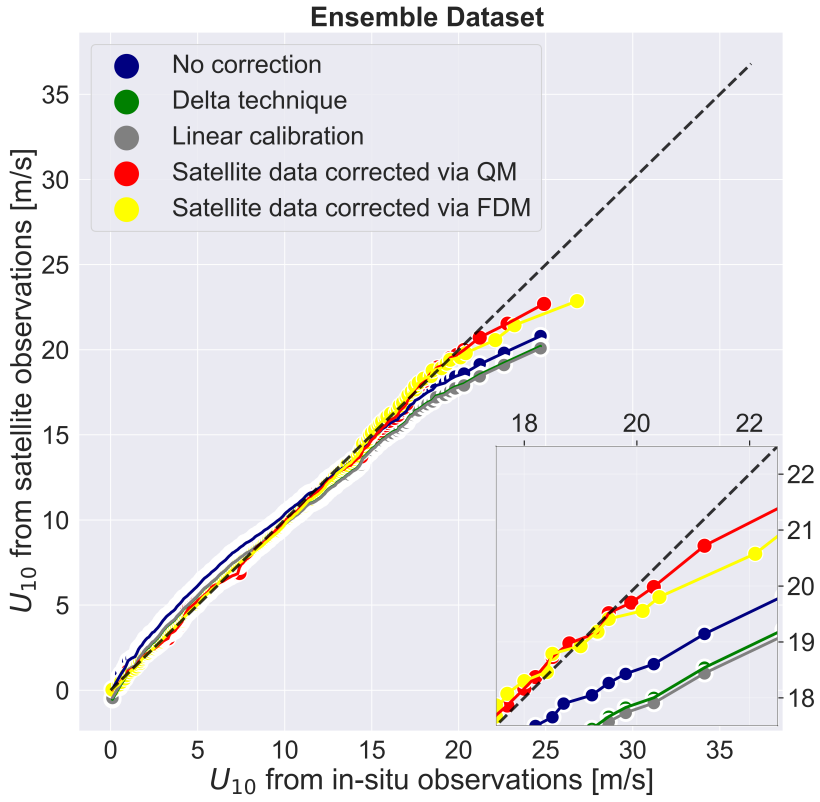


Figure 6.4: Final ensemble of all datasets: comparison of distribution of data, after correction techniques are applied.

6.3 Conclusions

This tesis presented a methodology by acquiring satellite altimetry data and collocating it to fixed-point positions, being able to compare altimeter data with in-situ data.
This procedure followed a spatio-temporal matching between satellite and mooring buoys, and through this methodology several criterions were proposed. The results showed us that the best spatial and temporal criterions for collocating altimetry data respect to in-situ data are 50 km cross-radius and a time-frame of 30 min, which brought the best results in terms of number of data and metric performance evaluation. Furthermore, spatial criterion IDW resulted being the best option for analysing wind speed data through this methodology, whereas Closest Distance was the best option for significant wave height.
After that, 4 bias correction techniques have been used to correct biases in the satellite dataset and through metric coefficients, performance and effectiveness of those techniques have been valuated. Results showed us that for datasets with a larger number of data, Full Distribution Mapping and Quantile Mapping techniques seem to work better for H_s, whereas simplier techniques such as Delta Method and Linear calibration work well for U_{10}.
Finally all satellite datasets have been grouped into a ensemble dataset: this procedure was done to evaluate even more effectiveness of those techniques and to analyze performance overall. From metric results, Linear calibration and Full Distribution Mapping provide the best performance across all coefficients. Final ensemble dataset has a total of 14,700 data points (for H_s part), while a total of 8,563 data points (for U_{10}). FDM and QM techniques improved slightly CC from 0.9712 to 0.9713 and 0.9714 respectively for significant wave height, while for wind speed CC went from 0.8285 to 0.8301 and 0.8307. For the wave part CC was already high and indeed the improvement is not enormous, whereas for wind speed it was quite low but still bringing not a big improvement. The calibration done using coefficient from Linear Regression is consistent for both wave and wind: RMSE went from 0.3336 m to 0.3197 m, similarly for wind RMSE went from 2.5562 m/s to 2.5025 m/s. A final consideration on FDM and QM techniques can be done: wind data wasa much more sparser than wave, therefore can be seen that FDM and QM produced a higher RMSE (2.5680 m/s and 2.5767 m/s compared to 2.5562 m/s). Another reason behind this results are the lower number of data points compared to the wave data.
For this work, all the analysis and data gathered from 1st January 2021 to 31th December 2023, bringing a limited number of data compared to other types of researches. Future calibrations and improvemenets can be done by collecting even more data, in order to see the full effectiveness of distribution mapping techniques to correct biases in models. Extending the number of data, more restrictive spatio-temporal criterions can be applied in order to get accurate data. More studies can done similarly in the North Sea, given the number of in-situ platforms increasing, whereas most of the analysis were done on the coasts in North America.

Additional informations

This study has been conducted using E.U. Copernicus Marine Service Information:

- **Product ID**: WAVE_GLO_PHY_SWH_L3_NRT_014_001, **DOI**: 10.48670/moi-00179

- **Product ID**: INSITU_GLO_PHYBGCWAV_DISCRETE_MYNRT_013_030, **DOI**: 10.48670/moi-00036

Appendix A

Appendix

A.1 Statistics

A.1.1 Probability

Here are introduced little concepts of probability theory. A random variable $\underline{x}$ is described by the probability density function $p(x)$ or *PDF*, which is defined as the probability of $\underline{x}$ obtaining a value between x and $x + dx$ is:

$$\Pr\{x < \underline{x} \leq x + dx\} = \int_{x}^{x+dx} p(x)\, dx \tag{A.1}$$

It follows that the probability of $\underline{x}$ being less than or equal to x is:

$$\Pr\{\underline{x} \leq x\} = \int_{\infty}^{x} p(x)\, dx = P(x) \tag{A.2}$$

which is the cumulative distribution function (CDF) of $\underline{x}$ (see figure A.1).

Then we introduce the average or the mean value of $\underline{x}$, which is called *expected value*

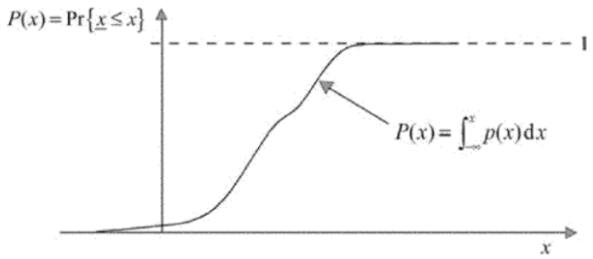


Figure A.1: CDF of a random variable [10].

denoted as μ_x or $E\{\underline{x}\}$:

$$E\{\underline{x}\} = m_1/m_0 = \int_{-\infty}^{+\infty} x\, p(x)\, dx \Big/ \int_{-\infty}^{+\infty} p(x)\, dx \tag{A.3}$$

which is defined in terms of first-order moment divided by the zeroth-order moment [10]. This average may be interpreted as the location of the PDF on the x-axis. Most of the

times, the average is not calculated from the PDF but calculated from a set of samples of $\underline{x}$:

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N}\underline{x}_i \tag{A.4}$$

where N is the number of samples. Note that these are only estimates, which will always differ from the expected values.These differences are called (statistical) sampling errors. A well-known probability function is the *Gaussian* probability density function.
While moments of spectrum are defined as:

$$m_n = \int_0^\infty f^n E(f)df \qquad \text{for } n = ..,-3,-2,-1,0,1,.. \tag{A.5}$$

This concept is quite important in the definition of H_s, for example variance of η^2 can be defined as:

$$\text{variance} = E\{\underline{\eta}^2\} = \int_0^\infty E(f)df = m_0 \qquad \text{for } \mu_\eta = E\{\eta\} = 0 \tag{A.6}$$

A.1.2 Statistical metrics

To evaluate accuracy or reliability between two datasets, there are several statistical metrics and coefficients commonly used. Here are presented 4 coefficients that have been used in this thesis [25]: Bias, RMSE (root-mean squared error), CC (correlation coefficient) and SI (scatter index), where y refers to satellite data and x refers to in-situ data:

$$Bias = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i) \tag{A.7}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2} \tag{A.8}$$

$$SI = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i - Bias)^2}}{\frac{1}{N}\sum_{i=1}^{N}x_i} \tag{A.9}$$

$$CC = \frac{\text{cov}(y,x)}{\sqrt{\text{cov}(y)\ \text{cov}(x)}} \tag{A.10}$$

(cov = covariance).
Bias is a coefficient that refers to the difference between measured values (or predicted values, in this case are satellite observations) and true values (or ground truth, that in this case are in-situ observations). Having a positive bias indicates that measured values are most of times greater than true values (overpredicted).
Similarly with bias definition, *RMSE* metric gives an indication of average error present in the predicted values. When there are big errors in the dataset, those are more sensitive

to this metric because there is the square-root. Clearly this coefficient has to be smaller as possible.
CC as it suggests the name, it is a indicator of the linear relationship between two variables: ideally if it closer to 1, it means that there is a strong correlation between those 2 variables. Lastly *SI* coefficient quantifies the RMSE referred to the mean of measured values: usually can be seen as a percentage and has to be smaller as possible. Having higher SI means that the measured values deviates strongly from ground truth.

Acknowledgements

I would like to express my deepest gratitude to all the people who contributed to the realization of this thesis. First and foremost, I sincerely thank my supervisors, Giuseppe and Giulia, for their exceptional guidance, valuable advice, and constant support throughout the project. This work has allowed me to expand my knowledge, improve my skills, and grow as a person.
I would also like to thank my classmate Riccardo, who supported me throughout this long journey and kept me company even in the most challenging moments.
My heartfelt thanks go to my entire family, who have always supported my choices and cared for me. In particular, an infinite thanks to my mother, who made all of this possible, helping me pursue my dreams despite the challenges and giving me the opportunity to build my future.
Finally, I want to express my gratitude to my girlfriend Francesca, who has supported, encouraged, and motivated me throughout this year. She has given me immense strength to overcome the obstacles I faced in this final stretch.

Bibliography

[1] DE Cartwright. "Buoy techniques for obtaining directional wave spectra". In: *Buoy Technology Symposium*. 1964, pp. 112–121.

[2] Matteo Postacchini et al. "Lagrangian observations of waves and currents from the river drifter". In: *IEEE Journal of Oceanic Engineering* 41.1 (2015), pp. 94–104.

[3] E Medina-Lopez et al. "Satellite data for the offshore renewable energy sector: synergies and innovation opportunities". In: *Remote Sensing of Environment* 264 (2021), p. 112588.

[4] Harald E Krogstad and Stephen F Barstow. "Satellite wave measurements for coastal engineering applications". In: *Coastal Engineering* 37.3-4 (1999), pp. 283–307.

[5] Pierre Yves Le Traon et al. "From observation to information and users: The Copernicus Marine Service perspective". In: *Frontiers in Marine Science* 6 (2019), p. 234.

[6] *Copernicus Marine Environment Monitoring Service*. URL: `\url{https://marine.copernicus.eu/}`.

[7] Rosemary Morrow and Pierre-Yves Le Traon. "Recent advances in observing mesoscale ocean dynamics with satellite altimetry". In: *Advances in Space Research* 50.8 (2012), pp. 1062–1076.

[8] IEC Marine Energy—Wave. "Tidal and Other Water Current Converters—Part 101: Wave Energy Resource Assessment and Characterization". In: *IEC TS* (2015), pp. 62600–101.

[9] Sten Tronæs Frandsen. *Turbulence and turbulence-generated structural loading in wind turbine clusters*. 2007.

[10] Leo H Holthuijsen. *Waves in oceanic and coastal waters*. Cambridge university press, 2010.

[11] Vallam Sundar. *Ocean wave mechanics: Applications in marine structures*. John Wiley & Sons, 2017.

[12] Michel K Ochi. *Ocean waves*. 1998.

[13] Posudin Yuriy. "Methods of Measuring Environmental Parameters". In: (2014), pp. 34–36.

[14] Alois Schaffarczyk. *Understanding wind power technology: theory, deployment and optimisation*. John Wiley & Sons, 2014.

[15] Michael Brower. *Wind resource assessment: a practical guide to developing a wind project*. John Wiley & Sons, 2012, pp. 37–42.

[16] NG Mortensen. "Wind measurements for wind energy applications. A review". In: *16th British Wind Energy Association Conference*. Mechanical Engineering Publications Limited. 1994, pp. 353–360.

[17] William Emery and Adriano Camps. *Introduction to satellite remote sensing: atmosphere, ocean, land and cryosphere applications*. Elsevier, 2017.

[18] Jean-Noël Thépaut et al. "The Copernicus programme and its climate change service". In: *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE. 2018, pp. 1591–1593.

[19] Craig J Donlon et al. "The Copernicus Sentinel-6 mission: Enhanced continuity of satellite sea level measurements from space". In: *Remote Sensing of Environment* 258 (2021), p. 112395.

[20] AVISO. *Satellite altimetry data*. URL: https://www.aviso.altimetry.fr/en/missions/past-missions/topex-poseidon.html.

[21] C.H. Davis. "Satellite radar altimetry". In: *IEEE Transactions on Microwave Theory and Techniques* 40.6 (1992), pp. 1070–1076. DOI: 10.1109/22.141337.

[22] Igor V. Florinsky. "Chapter 3 - Digital Elevation Models". In: *Digital Terrain Analysis in Soil Science and Geology (Second Edition)*. Ed. by Igor V. Florinsky. Second Edition. Academic Press, 2016, pp. 77–108. ISBN: 978-0-12-804632-6. DOI: https://doi.org/10.1016/B978-0-12-804632-6.00003-1. URL: https://www.sciencedirect.com/science/article/pii/B9780128046326000031.

[23] Wen Huang et al. "Integrated processing of ground-and space-based GPS observations: improving GPS satellite orbits observed with sparse ground networks". In: *Journal of Geodesy* 94.10 (2020), p. 96.

[24] EUMETSAT. *Sentinel-6 altimetry L2 and L3 data guide*. URL: https://user.eumetsat.int/resources/user-guides/sentinel-6-altimetry-level-2-and-3-data-guide.

[25] Agustinus Ribal and Ian R. Young. "33 years of globally calibrated wave height and wind speed data based on altimeter observations". In: *Scientific Data* 6 (2019). URL: https://api.semanticscholar.org/CorpusID:169035813.

[26] European Space Agency. *ESA Applications: Observing the Earth*. URL: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-6/New_sea-level_monitoring_satellite_goes_live.

[27] European Space Agency. *Copernicus Programme: Sentinel Wiki*. URL: https://sentiwiki.copernicus.eu/web/?l=en.

[28] François Boy et al. "CryoSat-2 SAR-mode over oceans: Processing methods, global assessment, and benefits". In: *IEEE Transactions on Geoscience and Remote Sensing* 55.1 (2016), pp. 148–158.

[29] Yves Le Roy et al. "SRAL SAR radar altimeter for sentinel-3 mission". In: *2007 IEEE international geoscience and remote sensing symposium*. IEEE. 2007, pp. 219–222.

[30] Liguang Jiang et al. "Evaluation of Sentinel-3 SRAL SAR altimetry over Chinese rivers". In: *Remote Sensing of Environment* 237 (2020), p. 111546.

[31] Ferran Gibert et al. "Water Elevation and Water Extent Measurements With Sentinel-6A Radar Altimeter Fully-Focussed SAR Data". In: *2022 Ocean Surface Topography Science Team Meeting*. 2022, p. 115.

[32] M Raynal et al. "Sentinel-6A Delay Doppler Processing: Assessment over the Global Ocean". In: *Proceedings of the SAR Workshop*. 2016.

[33] NASA. *Earth Observation Data Portal*. Data Processing Levels, Accessed 26-11-2024. URL: https://www.earthdata.nasa.gov/engage/open-data-services-and-software/data-and-information-policy/data-levels.

[34] Charles R McClain. "A decade of satellite ocean color observations". In: *Annual Review of Marine Science* 1.1 (2009), pp. 19–42.

[35] Jing Ye, Yong Wan, and Yongshou Dai. "Quality evaluation and calibration of the SWIM significant wave height product with buoy data". In: *Acta Oceanologica Sinica* 40 (2021), pp. 187–196.

[36] Candice Hall et al. "Emerging national data buoy center (NDBC) wave systems". In: *Conference: JCOMM DBCP-34 Meeting, Cape Town, South Africa. https://www. researchgate. net/publication/328748840_Emerging_National_Data_Buoy_Center_NDBC_Wave_S ystems*. 2018.

[37] Glenn D Hamilton. "National data buoy center programs". In: *Bulletin of the American Meteorological Society* 67.4 (1986), pp. 411–415.

[38] Copernicus Marine Service. *In-situ TAC Dashboard*. Accessed: 26-11-2024. URL: \url{https://marineinsitu.eu/}.

[39] Guillaume Dodet et al. "Error characterization of significant wave heights in multidecadal satellite altimeter product, model hindcast, and in situ measurements using the triple collocation technique". In: *Journal of Atmospheric and Oceanic Technology* 39.7 (2022), pp. 887–901.

[40] E.U. Copernicus Marine Service Information (CMEMS). *Global Ocean in-situ Nrt observations. Marine Data Store (MDS)*. DOI: 10.48670/moi-00036 (Accessed on 26-11-2024).

[41] IR Young, E Sanina, and AV Babanin. "Calibration and cross validation of a global wind and wave database of altimeter, radiometer, and scatterometer measurements". In: *Journal of Atmospheric and Oceanic Technology* 34.6 (2017), pp. 1285–1306.

[42] E.U. Copernicus Marine Service Information (CMEMS). *Global Ocean L3 Significant Wave Height From Nrt Satellite Measurements. Marine Data Store (MDS)*. DOI: 10.48670/moi-00179 (Accessed on 26-11-2024).

[43] Ricardo M Campos. "Analysis of Spatial and Temporal Criteria for Altimeter Collocation of Significant Wave Height and Wind Speed Data in Deep Waters". In: *Remote Sensing* 15.8 (2023), p. 2203.

[44] Markel Penalba et al. "Bias correction techniques for uncertainty reduction of long-term metocean data for ocean renewable energy systems". In: *Renewable Energy* 219 (2023), p. 119404.

[45] Gil Lemos et al. "On the need of bias correction methods for wave climate projections". In: *Global and Planetary Change* 186 (2020), p. 103109.

[46] Claudia Teutschbein and Jan Seibert. “Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods”. In: *Journal of hydrology* 456 (2012), pp. 12–29.

[47] Douglas Maraun. “Bias correcting climate change simulations-a critical review”. In: *Current Climate Change Reports* 2.4 (2016), pp. 211–220.