Directional and frequency spread of surface ocean waves from SWIM measurements

Eva Le Merle, Danièle Hauser, Charles Peureux, Lotfi Aouf, Patricia Schippers, Christophe Dufour, Alice Dalphinet

Directional and frequency spread of surface ocean waves from SWIM measurements

E. Le Merle[1], D. Hauser[1], C. Peureux[2], L. Aouf[3], P. Schippers[4], C. Dufour[1]and A. Dalphinet[3]

[1]LATMOS (CNRS, Université Versailles Saint-Quentin, Paris Sorbonne Université)
[2]CLS (Collection Localisation Satellite)
[3]Météo-France
[4]ACRIS-ST

Key Points:

- The SWIM ocean wave spectrum is the narrowest and the most peaked for situations with waves close to the fully development stage
- For the Southern Ocean conditions, it is found that SWIM directional spread at the dominant frequency varies according to the wave development stage.
- Compared to the MFWAM model, the SWIM observations indicate narrower and more peaked omni-directional spectra but similar directional spreads of the dominant waves in well-developed sea state conditions.

Corresponding author: Danièle Hauser, daniele.hauser@latmos.ipsl.fr

Abstract

The China France Oceanography Satellite (CFOSAT) launched in 2018 now routinely provides directional ocean wave spectra at the global scale. It consists of analyzing the normalized radar cross-section measured by the near-nadir pointing Ku-Band real-aperture scanning radar SWIM (Surface Waves Investigation and Monitoring). The significant wave height, dominant wavelength and direction are provided as the main parameters, but here, we analyze additional parameters, namely the frequency width of the omni-directional spectra, the directional spread of the dominant waves, and the related Benjamin-Feir index. This latter was proposed in the literature to estimate the probability of extreme waves. We discuss the geographical distributions of these parameters, their relation with sea-state conditions, and their similarities and differences with respect to the same parameters obtained from the MFWAM numerical wave model and buoy data. We find that the SWIM omni-directional spectra are narrower and more peaked than the model spectra and that these differences are more obvious in the high sea-state conditions encountered in the Southern Ocean. We find that under the intense conditions of the Southern Ocean, the SWIM directional spread at the peak is the smallest for swell, the largest for young wind seas, and takes intermediate values for mature wind seas. The directional Benjamin-Feir index is similar for SWIM and MFWAM, but this is mainly due to compensating effects in the parameters contributing to this index. The results indicate that these shape parameters may be used in the future to better describe the wave space-time evolution.

Plain Language Summary

The France Oceanography Satellite CFOSAT was launched in 2018. It routinely provides for all ocean basins detailed information on the ocean waves, namely the distribution of wave height with direction and wavelength (or wave frequency), obtained from radar measurements. In this study, we analyze several parameters which quantify how the energy spreads around the dominant frequency and the dominant wave propagation direction of the waves. Several of these parameters are also combined to estimate an index, which characterizes the probability of occurrence of extreme waves. To our knowledge, it is the first time that such parameters are accessible from space observations at the global scale. In the paper, we discuss the geographical distributions of these parameters, their relation with sea-state conditions, and their similarities or differences with respect to the same parameters obtained from a numerical wave prediction model and from buoy observations. We find that compared to the satellite observations, the model indicates narrower and more peaked distributions of energy in frequency, and we propose some explanations on this. Overall, the results indicate that these shape parameters from satellite observations may be used in the future to further understand or validate the physical processes impacting the evolution of waves during growth order decay.

1 Introduction

Ocean surface waves are commonly described by their directional density spectrum of height, which characterizes the distribution of wave energy as a function of wave frequency (or wave number) and wave propagation direction. However, in both numerical modelling studies and observation analyses, the full information of directional spectra is seldom discussed. Apart from significant wave height, which is widely used to characterize or forecast the total energy of waves, the mean or peak period and the mean or peak directions are the parameters most often used in these studies. This reduces the information on directional spectra to only few parameters, whereas information on directional spread or frequency spread contains useful information.

Most of the numerical models developed for describing or forecasting the evolution with time and space of the waves over the globe or at the regional scale are based on a spectral approach also called phase-averaged approach (Komen et al., 1994; Tolman & Chalikov, 1996; The SWAN Team, 2010). Basically, these models solve the wave energy or the wave action conservation equation discretized in wave frequency and wave direction intervals, with source and sink terms separated in three, namely the wind input term, the dissipation term and the non-linear interaction between waves. During the last two decades, most of these models have reached an appropriate degree of maturity to provide a fairly good agreement with respect to in situ and satellite observations, for the main parameters of the sea-state, namely the significant wave height, the mean direction and the mean period (The WISE Group et al. (2007); Bidlot (2017), and ECMWF reports[1]). However, uncertainties on these parameters still remain in certain conditions, in particular those strongly forced by the wind like in the Southern Ocean (Aouf et al., 2021; Young et al., 2020) or within tropical storms (Fan et al., 2020). Furthermore, the dominant wave period from the models still shows important biases with respect to observations (see ECMWF reports). The results from Stopa et al. (2015) also show that for the model Wave Watch 3 (WW3), the directional spread parameter is in general not consistent with buoy observations regardless of the model parameterization.

The successive improvements of numerical models are most often defined and assessed through comparisons of the significant wave height. Models may be satisfying in terms of significant wave height (or integrated wave energy), but their good performance may hide some unsatisfying behavior of the energy distribution in wave frequency or wave direction (Rogers & Wang, 2007). The reason of this is probably due to the fact that, since the 90s, model assessment at the global scale is mainly based on space-borne altimeter observations. These latter provide the significant wave height but no information on wave direction nor wave frequency. In consequence, the model improvements or tunings are not strongly constrained by parameters other than significant wave height. This can induce some artificial shortcomings on the physical parameterizations of the models. It is known from theory and academic numerical tests that the shape of the wave spectra in frequency and direction is highly sensitive to the physical parameterization of the models (e.g., Zieger et al., 2015; Alves & Banner, 2003), and to numerical simplifications in the non-linear interaction term (Gagnaire-Renou, 2009; Rogers & Van Vledder, 2013; Zieger et al., 2015; Forristall & Greenwood, 1998; Rogers & Wang, 2007), but there are only very few studies which have used the spectral shape information to adjust and validate the numerical models in their operational version.

Alleviating such shortcomings with in situ measurements as reference is not trivial because appropriate observations are scarce. Indeed, although systematic in situ observations from wave buoys or wave gauges also provide wave spectra as main information, most of these systems provide only the non-directional frequency spectra and the spectral resolution is rather poor. Only a few of them provide directional information. Mean or peak period and mean or peak direction are the parameters most often accessible to compare with the model results. Also, most of the observations are located close to the coasts, so that it may be insufficient to use these in situ observations to fully tune or validate the models.

Among the satellite observation techniques, two of them are appropriate to provide detailed information on ocean wave spectra. The high resolution images provided by Synthetic Aperture Radars (such as on ERS1, ENVISAT, Radarsat, Sentinel 1A/1B) can be inverted to provide wave spectra (e.g., Kerbaol et al., 1988). But it is well-known that they have important limitations because they cannot detect waves with wavelengths

[1] see ECMWF reports at `https://www.ecmwf.int/en/forecasts/charts/catalogue/w_wave_intercomparison?facets=undefined&time=2018123000,0,2018123000&area=All%20stations%20combined&statistics=Bias`

shorter than about a cut-off limit of about 200 m particularly when waves propagate along the satellite track (Alpers & Brüning, 1986). This azimuth cut-off induces a distortion of the wave directional spectrum, mainly along the range direction. Furthermore, today, none of these satellite missions provide wave spectra over the full globe, because SAR are also used in moderate or low resolution modes for other applications close to the sea-ice and to the continents. Since 2018, the new satellite called CFOSAT with its SWIM instrument onboard dedicated to wave measurements (SWIM= Surface Waves Investigation and Monitoring) provide directional spectra of ocean waves for waves between 70 and 500 m in wavelength (Hauser et al., 2017, 2020). Hauser et al. (2020) analyzed the performance on the wave height, dominant wave direction and dominant wavelength mainly through comparisons between SWIM data and model outputs, where the model used was the MFWAM wave prediction model, i.e. the French version of the ECMWF WAM model. This study showed that except for waves which propagate in a $\pm 15°$ sector along-track, the main parameters are in good agreement with the MFWAM model.

In spite of their limitations, it was proved that wave direction and wavelength of swell partitions estimated from SAR observations (from Sentinel-1A/1B) can significantly improve the model results in terms of wave height when they are assimilated in the model (Aouf et al., 2016). The MFWAM model now operationally assimilates these SAR observations. First tests of SWIM data assimilation also show a significant impact on the model results when the assimilation process includes the wave direction and wavelength of the dominant waves (Aouf et al., 2021). This is a way to better constrain the model with directional observations in addition to the sole significant wave height.

To go a step further, and in particular to help identifying or correcting the model inaccuracies, it is necessary to use additional metrics, such as those characterizing the spectral shape (Resio et al., 2016). The most natural ones are those related to the width of the omni-directional spectrum and the directional spread of the waves around their mean direction. One aim of the present paper is therefore to present and discuss such spectral shape parameters estimated from the SWIM data.

Another motivation for studying with more details the frequency and directional spread of the wave spectra is related to the needs for better characterizing or forecasting the probability of extreme waves (or freak waves). Using statistical simulations Socquet-Juglard et al. (2005) show that in the uni-directional case, the spectral development of long waves due to the modulation instability is significantly dependent on the initial spectral width in frequency and on the significant steepness. Janssen and Bidlot (2009) and Mori et al. (2011) show analytically how the kurtosis in the wave height distribution can be related to spectral parameters. In order to characterize the excess of kurtosis, they propose two parameters, namely the Benjamin-Feir instability index (BFI) in the uni-directional case and a two-dimensional version of it, called BFI_{2D}. BFI is proportional to the ratio of significant steepness to the spectral bandwidth in frequency (see Equation 8 in Section 2.5), whereas BFI_{2D} takes into account in addition, the directional spread of the waves. Although prognostics values on maximum heights are today currently provided as products of numerical wave models (Janssen & Bidlot, 2009; ?, ?), the parameters on which they are based (spectral frequency width, BFI, directional spread, BFI_{2D}) still need to be fully assessed.

Finally, directional spreading is important not only to better understand, model and forecast the evolution of surface waves, but also because it impacts the Stokes drift, which in turn is playing an important role in the ocean mixing in the upper layers (Webb & Fox-Kemper, 2015).

Considering the issues on the knowledge of spread parameters, our goal in this paper is to present and discuss the spectral shape of the wave spectra obtained from the analysis of the CFOSAT/SWIM observations at the global scale. In order to discuss the quality or limit of these estimations, we compare them to numerical model and buoy data.

The model considered here is the MFWAM, the French version of the WAM model. The paper is organized as follows. In Section 2, we describe the SWIM, MFWAM and buoy data sets and recall the definition of the spectral parameters used in the analysis. We also present in this section, the wind and wave height conditions of the subsets of data considered in this study. In Section 3, we present the spectral shape parameters (directional and frequency spread, BFI_{2D}) at the global scale, we discuss their geographic distributions, the relations between the different parameters and the comparison between observed and modeled parameters. In Section 4, a focus is made on results obtained in the Southern Ocean, a region of very high waves, strongly forced by stormy winds. From this subset of observations, we also discuss the possible reasons of the differences between the model and the observations. Section 5 summarizes the main results and conclusions.

2 Dataset and wave spectral parameters

2.1 SWIM dataset

The SWIM instrument deployed on board CFOSAT since October 2018, is a wave scatterometer with six radar-beams rotating around the nadir axis at small incidences (0°, 2°, 4°, 6°, 8°, 10°). Using the concept of real-aperture scanning radar, SWIM provides measurements in all azimuth directions, which allows to estimate directional wave spectra. The measurement principle of the concept is explained in Jackson et al. (1985a) whereas the space-borne configuration with SWIM is detailed in Hauser et al. (2017).

The wave products are distributed by the CNES (Centre National d'Etudes Spatiales) mission center; they consist in directional ocean wave spectra of wave cells of about 70 km per 90 km on each side of the satellite track (Hauser et al., 2017). The SWIM spectra are provided in the wavenumber domain from 0.013 rad/m and 0.28 rad/m which corresponds to the wavelength domain [22, 500] m. When converted in frequency by assuming the dispersion relationship in deep water, this wavenumber domain corresponds to the frequency domain from 0.056 Hz and 0.26 Hz. SWIM wave spectra are provided with a 180° ambiguity for the wave propagation direction and their angular discretisation is 15°.

Hauser et al. (2020) published the first analysis of the SWIM geophysical products.They compared to the MFWAM model wave parameters, such as significant wave height, dominant wavelength and direction. They concluded that with an updated speckle noise correction, the main parameters of the SWIM spectra are in reasonable agreement with the same parameters obtained from the MFWAM model and from in situ observations. A limitation was however found when waves propagate in a direction close to the satellite track. It is due to an inaccurate correction of speckle noise in this direction. They also concluded that the 10° beam gives the best results compared to the MFWAM data and to in situ observations. For this reason, in the present study, we only use the wave spectral information from the 10° incidence beam. Moreover, the data set used here corresponds to the updated version of the processing (version 5). This update concerns the optimized speckle noise correction as presented in Hauser et al. (2020), and the application of a normalisation on the spectral energy by using the significant wave height (SWH) from the nadir beam:

$$S(i_k, i_\phi) = \left(\frac{4}{SWH}\right)^2 \sum_{i_k=0}^{Nk-1} \sum_{i_\phi=0}^{N\phi-1} S_{pm}\left(i_k, i_\phi\right) dk\left(i_k\right) d\phi\left(i_\phi\right) / k\left(i_k\right) \quad (1)$$

With $S(i_k, i_\phi)$ the wave slope spectrum, and $S_{pm}\left(i_k, i_\phi\right)$ the modulation spectrum. This version is the one adopted to produce the operational products since October 12, 2020 and used to reprocess the SWIM data set since the beginning of the mission.

For the analysis presented below, the original SWIM spectra expressed as a function of wavenumber have been converted as frequency spectra assuming the deep water dispersion relationship.

SWIM first order parameters such as significant wave height, dominant wavelength and dominant direction are calculated by the operational processor on the 2D polar wave slope spectra.

The data set chosen for the analysis covers a 13 day period (i.e. a full orbital cycle of CFOSAT), and extends from the 10th to the 22th of September 2019. Observations over land and sea-ice have been filtered out. This global dataset represents 67 981 measurement points. All these samples are used for comparisons with model data. By comparison, a co-location with 43 buoys over 8 months of data represents only 6000 samples and much less diversity in terms of wave conditions (see Section 3).

2.2 MFWAM dataset

In this study, SWIM data are compared to MFWAM model data. The MFWAM wave model is the French version of the third generation WAM model. It is based on the ECMWF version (referred as ECW AM-IFS-38R2) with a parameterization taken from the ST4 version of the WW3 model (Ardhuin et al., 2010). The MFWAM products used here have a grid resolution of 10 km, are driven by 3-hourly winds from the IFS-ECMWF atmospheric system and are co-localised to the closest point from the SWIM measurement point. Wind information is also available in the MFWAM products. Wave/current interactions are taken into account with daily surface currents provided by the global PSY4-CMEMS ocean forecasting system. The MFWAM wave spectra are discretized in 24 directions between 0° and 360°, wich corresponds to directional bins of 15°, and 30 frequencies starting from 0.035 Hz to 0.58 Hz (1300 m to 5 m of wavelength). In its operational version as used here, the MFWAM model assimilates in real-time the significant wave height from various altimeter missions, and the directional information from Sentinel 1A and 1B SAR images. The MFWAM wave parameters are provided in output of the Météo-France operational processor. The peak periods are converted to peak wavelengths, assuming the deep water dispersion relationship.

2.3 Buoy dataset

Additional comparisons have been made with the buoys from the National Data Buoy Center (NDBC). We have selected 43 buoys moored at more than 60 km from the coasts to avoid coastal contamination in SWIM data. Buoys are located along the Pacific and Atlantic coasts of North America, around the archipelago of Hawaï, in the Gulf of Mexico, in the Caribbean Sea and along the east coasts of the archipelago of The Antilles. Significant wave heights, dominant periods and directions are provided in the data files. The dominant wavelength is then calculated from the dominant period considering the dispersion equation and taking into account the depth provided in the buoy data. Wave height omnidirectional spectra are also provided in the buoy data over the frequency range [0.02, 0.485] Hz. In our analysis of the spectral shape with frequency we have reduced this frequency range to the same interval as the one relative to SWIM. To estimate the directional spread, we have used the Fourier coefficient parameters provided in the buoy data files (Section 2.5).

The comparisons between the SWIM and the buoy data have been made over a period of 8 months between April 26 and December 31, 2019. For the comparison, the maximum distance accepted between SWIM and the buoy was choosen as 100 km and the maximum time lapse as 30 min. As the buoy are moored relatively close to the coasts, SWIM wave spectra expressed as a function of wavenumber have been converted into

frequency wave spectra using the dispersion relationship considering the depth given in the buoy data.

2.4 Overview of wind and wave height conditions for the global analysis

Figure 1 shows the significant wave height from the SWIM 10° incidence beam and the wind speed used for the forcing of the MFWAM model. The global conditions observed during this period show high sea states in the Southern Ocean with wind speed reaching 18 m/s and significant wave heights higher than 5 m. Similar conditions of high wind speed and wave height are also observed along the North American and the Greenland coasts, as well as in the Gulf of Alaska and Bering strait. Along the US coasts the high wind and wave conditions are due to the passage of the hurricane Humberto between 16 and 19 September 2019. Moreover, close by the Greenland coasts, along shore winds stronger than 20 m/s were observed between the 18th and 22th September 2019.

Figure 2 is a two-dimensional histogram representing the occurence of a given significant wave height for a given wind speed as derived from the MFWAM wind and wave parameters. The black curves reported in Figure 2 correspond to the relations derived from the Elfouhaily et al. (1997) wave spectrum under young, mature and developed conditions (wave age of 0.5, 1, 1.2 respectively). Over the 13 days period, the majority of situations correspond to significant wave heights between 1.5 m and 3 m and wind speeds projected along the wave direction between 5 m/s and 11 m/s. Based on the significant wave height to wind speed relationship for wind-sea, Figure 2 indicates that the majority of samples correspond to fully developed wind sea or swell conditions.

In order to clarify the discussions of the results, SWIM and MFWAM spectra have been classified in different categories of sea-state according to the inverse wave age Ω defined as classically as:

$$\Omega = \frac{U_{10} \cos\left(\theta_{wind} - \theta_{waves}\right)}{c_p} \tag{2}$$

Where c_p is the phase velocity, U_{10} is the wind speed at 10 m, θ_{wind} is the wind direction to the north and θ_{waves} is the wave propagation direction to the north. The inverse wave age is calculated using the MFWAM data.

According to the inverse wave age values, we established three sea state categories: the swell, the young and the mature wind sea. The swell category lists the ocean wave spectra with Ω lower than 0.84 (which corresponds to cases with wave age greater than 1.2). Ocean wave spectra with Ω higher than 1 form the young wind sea category. And cases with Ω between 0.84 and 1 are considered as mature wind sea. Over 67981 co-location points, 63300 are swell conditions, 1809 are young wind sea situations and 2872 are mature wind seas. Note however that when a sample belongs to one of this class, this does not mean that sea-state is pure swell or pure developed wind sea. Using the swell and wind sea partitioning of the MFWAM model spectra we have estimated the statistics of dominant swell, dominant wind sea and of mixed conditions, based on the ratio of swell or wind-sea energy to the total energy. If we consider that a ratio of 0.70 indicates the predominance of one particular system, we find that swell predominates in 50% of the cases, wind sea predominates in 5% of the cases, and mixed sea conditions are found as the complementary, i.e. in 45% of the cases.

2.5 Spectral shape parameters

As mentioned in the introduction, we propose to study integrated spectral parameters that are less reported in the literature than the first order parameters, but very important to assess model performances (Rogers & Van Vledder, 2013; Gagnaire-Renou, 2009). These parameters characterize the energy distribution in frequency (σ_f) and di-

rection (σ_ϕ), the "peakedness" of the omni-directional spectrum (Qp) and the Benjamin Feir Index (BFI).

In the literature, there are several definitions to characterize the distribution in frequency as reported by Saulnier et al. (2011). In this study, the spectrum bandwith is calculated by the frequency spread, noted σ_f and defined by the Blackman and Tukey (1959) formulation:

$$\sigma_f = \frac{\left[\sum_{f_{\min}}^{f_{\max}} F(f)\, df\right]^2}{\sum_{f_{\min}}^{f_{\max}} F^2(f)\, df} \tag{3}$$

where $F(f)$ is the omnidirectional wave height spectrum. $f_{\min}$ and $f_{\max}$ define the frequency domain for the frequency spread calculation: $f_{\min} = 0.056$ Hz and $f_{\max} = 0.28$ Hz. As σ_f measures the spectrum bandwith in frequency, large values of σ_f indicate broad spectra.

In order to characterize the spectral peakedness, we also estimate the Goda parameter as introduced by Goda (1976) and noted Qp. Its formulation is:

$$Qp = \frac{2\sum_{f_{\min}}^{f_{\max}} f\ F^2(f)\, df}{\left[\sum_{f_{\min}}^{f_{\max}} F(f)\, df\right]^2} \tag{4}$$

This parameter is an indicator of how sharp is a spectrum. Large values of Qp indicate sharp spectra around the dominant energy peak. For this study, the frequency spread and Qp parameter are calculated over the same frequency domain [$f_{\min}$, $f_{\max}$] for SWIM and MFWAM. Note that Qp and σ_f vary in the opposite way: large values of σ_f which indicate broad spectra correspond to low values of Qp, and the other way around for narrow spectra. Although initially defined for narrow-band single mode spectra (as several other width definitions based on spectral moments), it was shown by Rao (1988) that the Goda "peakedness" parameter can be considered as an appropriate parameter to characterize the spectral width because, due to its dependence on the square of the frequency spectrum, it is mostly sensitive on the spectral shape near the dominant peak. To further asses this statement, we have carried out additional tests on synthetic spectra, by simulating bimodal spectra with two gaussian shapes of same variance separated in frequency. The main result is that when increasing from 0.5 to 1 the energy ratio of the second to first peak, it changes σ_f by more than 25% whereas it only marginally affects Qp (change of 5%). Actually, in case of bimodal spectra, Qp takes a value close to that of the most energetic peak (and not much affected if both peaks are of similar amplitude), whereas σ_f increases rapidly as the second peak is increased and reaches an asymptotic value.

The directional spread σ_ϕ and the mean direction ϕ_{mean} are calculated using the first pair of the Fourier coefficients $a_1(f)$ and $b_1(f)$:

$$\sigma_\phi(f) = \sqrt{2 \times \left(1 - \sqrt{a_1(f)^2 + b_1(f)^2}\right)} \tag{5}$$

$$\phi_{\text{mean}}(f) = \arctan\left(\frac{b_1(f)}{a_1(f)}\right) \tag{6}$$

The expressions for $a_1(f)$ and $b_1(f)$ are those proposed by Longuet-Higgins et al. (1963) and recalled by Kuik et al. (1988) or Pettersson et al. (2003) among others. For SWIM and MFWAM we have calculated $a_1(f)$ and $b_1(f)$ from the directional spectra, whereas for the buoy data we have directly used the Fourier coefficients provided in the NDBC buoy files. The directional spread at the dominant wavelength of the spectrum is analyzed in Section (3.2) whereas in Section (4.3) we present results on the directional spread as a function of the normalized frequency f/f_p, where f_p is the frequency at the energy

peak of the spectrum. Note that if one considers that the directional distribution can be represented by a shape in $\cos^{2s}$ as proposed by e.g Mitsuyasu et al. (1975) then s and σ_ϕ are related by:

$$\sigma_\phi = \sqrt{\frac{2}{s+1}} \tag{7}$$

The last parameter investigated in this study is the directional version of the Benjamin Feir Index (BFI) introduced by Mori et al. (2011). BFI has been proposed in the literature as an appropriate indicator of non-linearities of wave interactions and probability of occurrence of extreme waves in the case of unidirectional seas (Janssen & Bidlot, 2009):

$$BFI = k_0\sqrt{m_0}Qp\sqrt{2\pi} \tag{8}$$

With k_0 the mean wavenumber and m_0 the 0th order moment of the energy of the density spectrum. BFI is proportional to Qp and to the significant wave slope : $k_0\sqrt{m_0}$. Its values span between 0 and 1, and the highest it is, the highest is the kurtosis of the probability density function (pdf), i.e the highest is the probability of occurrence of extreme waves. This parameter does not include the directional effects whereas it was shown by Mori et al. (2011) that the directional spread impacts the excess of kurtosis of the pdf wave heights. For a given value of BFI, an increase of the directional spread reduces the excess of kurtosis. Therefore, Mori et al. (2011) introduced an extension of BFI, named BFI_{2D}, in order to include the directional effects. Its formulation is:

$$BFI_{2D} = \frac{BFI}{\sqrt{1+\alpha_2 R}} \tag{9}$$

Where the constant α_2 is equal to 7.10 (fitted from numerical simulations by (Mori et al., 2011)) and R measures the importance of directional width compared to the frequency width:

$$R = \frac{1}{2}\,\sigma_\phi^2 \pi Q p^2 \tag{10}$$

3 Results on the shape parameters at the global scale

In this Section we discuss the shape spectrum parameters and the Benjamin-Feir Index at the global scale. Because the wave conditions corresponding to the co-located SWIM/buoy data set are rather limited (small significant wave heights and short wavelengths), the comparison of SWIM to buoy parameters is mainly used in the following to explain some of the differences between SWIM and MFWAM in these conditions.

3.1 Frequency shape parameters

Figure (3a) shows the geographic distribution of the SWIM frequency spread calculated with Equation 3. When comparing to sea state conditions illustrated in Figure 1, one can see that the smallest values of σ_f (typically less than 0,10 Hz) are encountered mainly in regions of high winds and high significant wave heights such as in Southern Ocean or along the coasts of Alaska or Greenland. This is rather counter intuitive because these regions are characterized by wave actively forced by the wind and we may expect broader spectra in these conditions than in those dominated by swell. But actually, these regions where the smallest values of σ_f are encountered, correspond to wind-waves close to the fully developed stage. Hence, the frequency spread is relatively small compared to young seas, and these latter are relatively rare at the global scale (see Figure 2).

Large values of σ_f (typically between 0.17 and 0.18 Hz) are found either along the Equatorial band, and in some areas of the central and eastern parts of the North Pacific as well as in enclosed oceanic basins (e.g. Gulf of Mexico, Mediterranean Sea, South China Sea). In most of these regions, the significant wave height is low (less than 1.8 m, see

Figure (3a)). Especially for situations in enclosed oceanic basins, these large values of σ_f are due to the presence of spurious energy peaks at the lowest frequency in the SWIM wave height spectra (Tourain et al., 2020). These spurious peaks appear at the lowest frequency in the spectra, due to the amplification of the remaining noise floor at low frequency (i.e. low wavenumber k), even if the noise floor is close to zero in the wave slope spectra. To convert wave slope to wave height spectra we apply a division by k^2 which dramatically amplifies small remaining noise when it is present at low frequencies.

The results obtained for σ_f with the MFWAM model are illustrated in Figure 3b as a geographical map. As for the SWIM results, σ_f values are the smallest in the regions of fully developed wind-waves (Southern Ocean, Greenland and Alaska coastal regions), whereas large values of σ_f are found along the Equatorial band, in the central and eastern part of the North Pacific. In these regions, MFWAM gives larger values of σ_f than SWIM. These relatively large values are partly due to the presence of mixed sea conditions. Indeed, the study of the ratio of swell and wind sea energy to the total energy indicates the presence of mixed sea conditions in these regions which impact the σ_f value. In opposite, regions of enclosed oceanic basins (e.g. Mediterranean Sea, Gulf of Mexico), are characterized by intermediate values of σ_f for MFWAM, smaller than in the case of SWIM. As mentioned above, the presence of a spurious energy peak impacts the σ_f value in these low sea state conditions.

Table 1 presents the mean and median values of SWIM σ_f by category of sea-state conditions (swell, young wind sea or mature wind sea, see Section 2.4). Overall, the swell conditions correspond to the smallest σ_f whereas, young wind sea correspond to the broader spectra, but note that in each category, mixed sea cases may also exist. The statistics for MFWAM results are also presented in Table 1. They indicate that for both mature wind sea and swell conditions, MFWAM spectra are broader than the SWIM spectra in average. In opposite in the young wind sea cases, MFWAM spectra are narrower than the SWIM spectra probably because of the spurious peaks on SWIM spectra as mentioned previously. This is confirmed by a complementary analysis carried out by comparing SWIM data to buoy data (see below discussion on Figure 4).

The statistics on the differences between SWIM and MFWAM σ_f values are illustrated in Figure 4a with the histogram of σ_f differences (SWIM-MFWAM) as a function of the significant wave height. Figure 4 shows the statistics of differences between SWIM and buoy data. In the mean, the difference between SWIM and MFWAM is small but systematically negative, with a mean value of -0.010 Hz, and a root mean square difference (RMSD) of 0.023 Hz. For values of Hs greater than about 2 m, the mean difference is clearly negative (SWIM values smaller than MFWAM values). For the smallest Hs values (less than about 2 m), although the largest occurrence is for negative differences, the number of points with positive differences is also significant. At these small Hs, positive differences are also found with respect to the buoy values (Figure 4b). Therefore, it is likely that the positive differences between SWIM and MFWAM at low Hs indicate non reliable values from SWIM, probably due to the spurious peaks in the SWIM spectra in these conditions. In opposite, for the negative differences, SWIM σ_f smaller than MFWAM values, there is no evidence that SWIM results should be questioned. The same trend is also found when we analyze the histogram of differences as a function of the dominant wavelengths (not shown): there is a clear negative mean difference between SWIM and MFWAM σ_f values for all peak wavelengths above 100 m. The histograms of differences have also been analyzed for mixed sea conditions only (not shown) and the negative mean difference is also observed. Hence, for all conditions except for the cases of small Hs and/or short dominant wavelengths, we can conclude that there is a systematic positive bias of MFWAM σ_f with respect to σ_f from SWIM, i.e. the model spectra are systematically broader than the SWIM spectra. This conclusion is further assessed in Section 4 with the analysis of the Southern Ocean data set.

Figure 5a shows the Goda peakedness factor Qp from SWIM. It shows that maximum values (around 2.4) are encountered in the high sea-state regions (the same as those with low σ_f values). In opposite, in the areas of low significant wave height (enclosed ocean basins and some areas in the open ocean), Qp shows the lowest values (around 1.0). These are the regions where we mentioned here above that SWIM spectra may be polluted by spurious peaks at low frequency. So, results from SWIM must be considered with caution in these regions. In almost all other regions, Qp takes intermediate values. The comparison with the results from MFWAM (Figure 5b) shows that Qp values from SWIM are larger than those from MFWAM in average. In the areas of strong winds and high significant wave heights, such as in the Southern Ocean or along the Greenland or Alaska coasts, this overestimation by SWIM is particularly obvious.

Table 2 shows the statistical results by category of sea-state conditions (young wind sea, mature wind sea and swell), for both SWIM and MFWAM results. For young and mature wind sea cases, MFWAM values are larger than SWIM ones, whereas the opposite is true for swell conditions. For MFWAM, the mean and median values of Qp values decrease from young wind sea to mature wind-sea and from mature wind-sea to swell. For SWIM, the young wind-sea Qp is in-between values for swell and mature wind sea. It is likely that this is due to the limitations of SWIM already mentioned for the young wind-sea cases which predominantly correspond to low significant wave height and short wavelength conditions. For SWIM, similarly to MFWAM, swell cases show the smallest peakedness parameter among all the cases. Again, it must be noted that this analysis does not distinguish mixed sea cases, but as explained in Section 2.5, Qp is expected to be less sensitive to bimodal spectra than σ_f.

Figure 6a shows the histogram of SWIM-MFWAM differences for Qp as a function of the significant wave height. For Hs lower than about 1.8 m, we observe a large scatter, and a systematic negative difference between SWIM and MFWAM values (bias of -0.20). As mentioned above, these points correspond to situations of low sea-state in enclosed seas where the SWIM data are probably affected by spurious energy peaks which appear at low frequency (long wavelength). The same type of difference is also found from the comparison of SWIM with the NDBC buoy data set (Figure 6b). This seems to indicate that SWIM spectral shape must be analyzed with caution in these situations. On the other hand, for wave heights larger than about 1.8 m, there is a systematic positive bias of 0.14 (0.29) for SWIM with respect to MFWAM. As evidenced with the results presented in Table 2, these larger values of Qp for SWIM with respect to MFWAM are associated to swell cases and mature wind sea in areas strongly forced by the wind (e.g. Southern Ocean).

3.2 Directional spread

Figure 7a shows the geographical distribution of the SWIM directional spread σ_ϕ estimated at the frequency peak of each spectrum. According to this map the directional spread at the peak of the spectrum is minimum (15°-25°) in the regions of high sea states like in the Southern Ocean, along the coasts of Alaska, Greenland and North America. Elsewhere, the directional spread takes values from about 25° to 60° without any clear relation with the Hs map (Figure 1a) nor with the map of dominant wavelengths (not shown).

σ_ϕ estimated from MFWAM is shown in Figure 7b. As for SWIM, in regions of active generation, σ_ϕ also indicates relatively narrow spectra. This is again counter intuitive but is confirmed by the mean and median values estimated when sorting the data set by sea-state categories (Table 3). At the global scale the narrower spectra are observed for the mature wind sea conditions for both SWM and MFWAM whereas young wind sea conditions correspond to the broader spectra.

Moreover, MFWAM data shows narrow spectra not only in the active generation regions but also in other areas like west of Chile, west of Africa and in the Gulf of Bengal whereas, SWIM values are large. After analyzing the maps of SWIM and MFWAM dominant wavelengths (not shown), we could identify that these differences correspond to locations where SWIM dominant wavelengths are significantly smaller than those from MFWAM. Comparing the ratio of the swell and the wind sea energy to the total energy of the spectrum, we find that mixed sea conditions are mostly present in these areas, according to the MFWAM data. The presence of mixed sea in these regions can explain the different identification of the dominant wave system between SWIM and MFWAM and induce differences in the directional spread. Indeed, smaller dominant wavelengths and broader directional distributions at the peak of the spectrum are observed for SWIM compared to MFWAM. Significant differences are also observed in enclosed seas and near the coasts. In these areas it seems to be a limitation of SWIM because of the presence of a spurious peak as mentioned above. The comparison of Figures 7a and 7b also shows some large negative differences in the directional spread between SWIM and MFWAM in the Mid-Pacific Ocean. According to MFWAM one can find mixed sea conditions in this region but these cases are not associated to any significative differences in dominant wavelengths between the two sets of data.

Figures 8a-b highlight the problem of mis-association between SWIM and MFWAM. Indeed, the MFWAM histogram indicates a larger density of points for cases with dominant wavelength around 250-300 m than SWIM. However, the mean directional spreads at these long wavelengths are of the same order of magnitude for SWIM and MFWAM (25°). At shorter wavelengths (lower than 200 m), the average value of the SWIM σ_ϕ (45°) is higher compared to MFWAM (30°). Which means that in average, situations with relatively short dominant wavelengths (like young wind sea conditions) are more spread in direction than MFWAM.

Figure 9a shows the histogram of the differences between SWIM and MFWAM for the directional spread, as a function of the SWIM significant wave height. In spite of the visual aspect of Figure 9, the SWIM to MFWAM differences are almost similar for small and large significant wave heights (less or greater than 1.8m): bias of about 4° and RMSD of 15° to 18°. At small Hs, the upward part of the scatter plot (positive bias) is partly due to the mis-association of the dominant wavelengths of SWIM and MFWAM spectra as mentioned above and partly due to the limitations of SWIM partition at small Hs as evidenced in Figure 9b obtained from the SWIM to buoy comparison. Even if the number of co-located points is less in Figure 9b than in Figure 9b, one can see that there there is also a positive bias of SWIM with respect to buoy data (mean of about 3°), with however an important scatter (RMSD of about 21°). Even if small, this positive bias for SWIM directional spread compared to both MFWAM and buoy data, and the relatively large value of RMSD indicate that one must remain cautious on the SWIM results in these regions of enclosed seas and coastal zones.

Table 3 shows that for all categories, MFWAM values are slightly smaller than those from SWIM, meaning a larger directional spread from SWIM than from MFWAM in the mean. The maximum difference ($\sim 8°$) is for the young wind sea. As mentioned above, it is clear that at the global scale, the mature wind sea conditions correspond to the narrower spectra for SWIM and MFWAM. This is not really expected from the literature (the swell systems are always mentioned as corresponding to the narrowest angular spread conditions), but we will see in Section 4.3 that when we limit our analysis to the Southern Ocean area, the results are slightly different.

3.3 Benjamin Feir Index

As mentioned in Section 2.5, the BFI_{2D} is an appropriate indicator of non-linearities of wave interactions and probability of occurrence of extreme waves. Figure (10a) shows

the spatial distribution of BFI_{2D} calculated with the SWIM data. As could be expected, the Southern Ocean, the region along the Greenland coasts and the Gulf of Alaska are the areas where BFI_{2D} values are the highest. This is also the case for BFI (not shown here).

The comparison of BFI_{2D} calculated with SWIM and MFWAM data shows a very good agreement (Figure 10b) with a mean bias equal to -7 10^{-3} and a RMSD equal to 2.1 10^{-2}. In opposite to the BFI comparison (not shown here), there is no overestimation of the SWIM BFI_{2D} compared to the MFWAM BFI_{2D}. This can be explained by the fact that BFI_{2D} is expressed as a function of BFI and of the R coefficient (Equations 9 to 10), both dependent on the Qp parameter. Comparison of the R coefficient between SWIM and MFWAM indicates an overestimation of SWIM (not shown here). This is due to larger values of Qp and, to a lesser extent, to the largest values of σ_ϕ. Hence, it turns out that for BFI_{2D} from SWIM compared to MFWAM, the overestimation of BFI which appears at the numerator is compensated by the overestimation of R which is at the denominator. Although the mean difference between SWIM and MFWAM for BFI_{2D} is very small, it does not result from a good agreement between SWIM and MFWAM but from a compensation effect between BFI and the R coefficients. We will see below that the same result is found for the case of the Southern Ocean where σ_ϕ values are similar for SWIM and MFWAM. This means that this compensating effects is mainly supported by Qp.

4 Focus on the Southern Ocean

4.1 Comparisons of SWIM and MFWAM shape parameters

The Southern Ocean is the only ocean connected to the three main oceans in the world: Pacific, Atlantic and Indian and it plays a non-negligible role on the Earth's climate. This region is dominated by strong westerly winds, with almost unlimited fetches (Young et al., 2020). In winter, the wind speed reaches high values of the order of 20 m/s or locally more (Young et al., 2020). These intense winds and unlimited fetch conditions generate extreme sea states. This is one of the reason why we focus here our analysis on SWIM and MFWAM data in the Southern Ocean (latitudes between -40° and -70°). The other reason is that in this region, the limitations of SWIM are the smallest because in these high significant wave height situations, the SWIM measurements are of the highest quality (negligible effect of a potential spurious peak, minimum impact of the speckle noise). Moreover for this data set, there is a good agreement between SWIM and MFWAM for the dominant wavelengths (mean values of 219 m for SWIM and 236 m for MFWAM), with limited number of cases where there is a mis-association of the dominant wave systems between SWIM and MFWAM. For the period of investigation, the mean significant wave height in the considered area is 4.3 m which is twice the mean Hs at the global scale. Hs values range between 1 m and 11 m.

Figure (11) shows the comparisons of the frequency and directional spread, Qp and BFI_{2D} between SWIM and MFWAM. The frequency spread (Figure 11a) is clearly larger for MFWAM than for SWIM data, with a mean difference of 0.015 Hz, larger than at the global scale (0.010 Hz). Moreover, in contrary to the comparison at the global scale, in the Southern Ocean the scatter remains limited because most of the situations encountered in this area have high signifiant wave heights. Correlatively, the comparison of Qp in Figure (11b) shows that SWIM values are larger than those from MFWAM, with a bias between SWIM and MFWAM of 0.26. It indicates that SWIM spectra are more peaked than the MFWAM spectra. So, SWIM and MFWAM wave spectra significantly differ in their omni-directional shape with broader and less peaked spectra for MFWAM as compared to SWIM.

An overestimation of the frequency spread of model spectra with respect to observations has already been mentioned in several studies. From comparison between MFWAM results with buoy and KuROS airborne radar data (KuROS is a concept similar to SWIM), Le Merle et al. (2019) concluded on the same kind of overestimation of the model frequency spread compared to buoy and radar observations in fetch limited conditions in the Mediterranean sea. Several authors have also discussed in the past, that the frequency width of the modelled wave spectrum is highly sensitive to the way the non-linear interactions are taken into account (Rogers & Van Vledder, 2013; Gagnaire-Renou, 2009; S. Hasselmann & Hasselmann, 1985a, 1985b). Here, we find that spectra from the MFWAM model are systematically broader than the SWIM spectra. As MFWAM uses the Discrete Interaction Approximation -DIA- to represent the non-linear interactions, it is possible that this is due to this approximation. Such bias in frequency spread was also evidenced in S. Hasselmann and Hasselmann (1985b) on few test cases. More recently, Annenkov et al. (2021) shows with academic tests, that in fully developed conditions they obtain closer values of Qp with observations using the Zakharov-Kolmorov approch instead of the classical representation of non linear interactions by the DIA.

The comparison on the directional spread calculated at the peak frequency (Figure 11c) shows that for this parameter, there is a good agreement between SWIM and MFWAM with a mean bias of 0.4° and a RMSD of 12° which is less than at the global scale. So, although the wave evolution modeled by MFWAM seems to produce wave spectra which are too broad in frequency compared with the observations, this does not seem to affect the directional spread in the conditions encountered in the Southern Ocean. This is due to the fact that, even if mixed sea conditions are present, wind sea and swell systems propagate in the same directions in average. This has been checked when analyzing the directional wave spectra in this region.

The BFI_{2D} comparison (Figure 11d) indicates a good correspondance between SWIM and MFWAM (mean difference of 0.0031). As mentioned in Section 3.3, the rather small difference between SWIM and MFWAM BFI_{2D} is in fact due to a compensating effect of larger values of BFI and R (Equation 10) parameters from SWIM compared to MFWAM. Moreover, in the case of the Southern Ocean data set, the larger values of R for SWIM compared to MFWAM are mainly due to the Qp factor and not to the directional spread.

4.2 Frequency spread according to sea state categories

In this section, the distributions of the frequency spread and the Qp parameter are estimated and compared for each sea state category. Similarly to the analysis at the global scale, sorting the data according to the 3 classes of inverse wave age does not exclude that there are mixed sea situations in certain areas or periods. For the Southern Ocean, we have verified that over our data set (16667 samples) swell remains dominant (i.e. contributes to more than 70% of the total energy in the spectrum) in 50% of the cases, whereas wind waves are dominant (i.e. their energy contributes to more than 70% of the total energy in the spectrum) in 8% of the cases. The remaining (42%) may be mixed seas, according to MFWAM. By analyzing the directional wave spectra in this region, we could check that these situations of mixed seas correspond to wind sea and swell propagating in the same directions. Figure (12) shows the histograms of the frequency spread for SWIM and for MFWAM data. The frequency spread is calculated for each individual omni-directional spectrum over the dimensional frequency vector. The mean and median values of the frequency spread distributions are reported in Table 1. In the case of SWIM, the mean and median frequency spread are the smallest for swell conditions (mean and median values of 0.097 and 0.093 Hz, respectively) and the largest in young wind sea conditions (mean

and median values of 0.110 Hz and 0.108 Hz, respectively), with intermediate values in mature wind sea conditions (mean and median values of 0.098 Hz and 0.092, respectively). The frequency spread from MFWAM is significantly broader than that from SWIM in all conditions. The largest differences are for the swell conditions (mean value of 0.109 Hz). This confirms the results discussed in Section 3 from the global data set, but extends the results to the young wind sea cases, which are here not affected by uncertainties at low signifiant wave heights of SWIM spectra.

Figure (13) shows the distributions of the Goda peakedness factor for SWIM (a) and for MFWAM (b) data. The mean and median values of the Qp parameter distributions are reported in Table 2. For SWIM, the smallest mean values are found for the swell category ($Qp = 2.00$), the largest for the mature wind seas category ($Qp = 2.11$) and intermediate values are found for young wind sea category ($Qp = 2.07$). When considering the median instead of the mean values, the ranking is the same: Qp values are the smallest for swell, and the largest for the young wind sea, whereas intermediate values are found for young wind sea (Table 2). Hence, with the SWIM data, we find that the mature wind sea cases have sharper omni-directional spectra than both the swell and young wind sea cases. The conclusion is slightly different with the MFWAM distributions, which indicate that the largest values are found for the young wind sea, the smallest for swell and intermediate values for mature wind seas. In addition, the most significative difference is that swell spectra from MFWAM are much less sharp than SWIM spectra (mean value of Qp is 1.76 for MFWAM instead of 2.00 for SWIM). For young and mature wind sea cases, MFWAM mean values of Qp are also significantly smaller for MFWAM than for SWIM. As mentioned above, one explanation could be the use of the DIA to represent the wave-wave non linear interactions in the MFWAM model. Indeed, our results can be compared to those of Rogers and Van Vledder (2013) who showed by comparing simulated spectra to wave buoy spectra, that the DIA tends to produce underestimated values of Qp compared to those obtained from the buoy measurements. This underestimation does not exist any more when Rogers and Van Vledder (2013) use the exact formulation of the non-linear interactions in their model. Note however that in our case, the underestimation of Qp from model with respect to observations is evidenced for all conditions of young wind sea, mature wind sea and swell, although this difference is less obvious for the wind sea cases.

4.3 Directional spread of mean ocean wave spectra

In this section the directional spread is investigated as a function of the adimensional frequency (f/f_p) for mean spectra estimated for both SWIM (Figure 14a) and MFWAM (Figure 14b). Mean spectra are calculated for each sea state category. In order to estimate these mean spectra, individual wave spectra are transformed before averaging to express them as a function of the adimensional frequency vector: f/f_p, where f_p stands for the peak frequency. Moreover, the individual directional spectra are rotated in direction in order to set the mean wave propagation direction along the north direction. These spectral transformations allow to force all the spectra used in the averaging procedure, to have their maximum matching each other in direction and adimensional frequency.

The results from the SWIM spectra (Figure 14a) show that the minimum of directional spread is found at f/f_p=1. This is in agreement with many other experimental results obtained in wind wave conditions as mentioned by Forristall and Ewans (1998) in his review and also described by Mitsuyasu et al. (1975); Ewans (1998); Babanin and Soloviev (1998); Hwang et al. (2000); Pettersson et al. (2003); Romero and Melville (2010). D. E. Hasselmann et al. (1980) and Donelan et al. (1985) found slightly different position of the minimum of the directional spread (1.05 f_p in the former case, and 0.95 in the latter case). With our data set the resolution in frequency is not sufficient to conclude more precisely about this position. From the SWIM spectra we find that the min-

imum value of the directional spread is weakly dependent on sea-state conditions with broader distributions for young wind sea (around 33°) and narrower distributions for mature wind sea (29°) and swell (27°) conditions. Mitsuyasu et al. (1975); Donelan et al. (1985) and Romero and Melville (2010) also found such a small sensitivity with wave age whereas almost all the other authors mentioned here-above could not evidence such a trend with wave age. Figure (14b) shows that for MFWAM, the direction spread slightly decreases with wave development (smaller for swell and mature wind sea than for yound wind sea) however it has to be noted that for MFWAM, the position of the minimum of the directional spread varies significantly with the sea-state category: it occurs at f/f_p = 0.8 to f/f_p = 0.9 for swell cases but at f/f_p >1 for wind waves and intermediate values for mature wind waves. MFWAM differs on this point from SWIM and from other experimental results of the literature, but on the other hand MFWAM and SWIM minimum angular spread are very similar. This shows that MFWAM is able to reproduce the observed angular spread near the peak. This is consistent with the results of Rogers and Wang (2007) who showed that in spite of the numerical simplification of the representation of the non-linear interactions in the SWAN model, the directional spreading near the peak frequency is quite close to that of buoy estimates.

For both SWIM and MFWAM, there is a marked broadening of the angular distribution on each side of its minimum position. At frequencies higher than about 1.2 f_p, SWIM spectra do not show any dependence with sea-state categories (swell or wind sea), in opposite to the case of the MFWAM mean spectra. On this point, SWIM spectra are in good agreement with the results from the literature (Mitsuyasu et al., 1975; D. E. Hasselmann et al., 1980; Ewans, 1998; Babanin & Soloviev, 1998), whereas MFWAM spectra are not. At frequencies smaller than f_p, both SWIM and MFWAM indicate a broadening towards the smallest frequencies and wider spectra for young wind sea than for mature wind sea and swell. The difference between the three categories is however more pronounced on MFWAM than on SWIM spectra. At these normalized frequencies less than 1, this dependence of angular spread with the wave development was also found by Mitsuyasu et al. (1975); D. E. Hasselmann et al. (1980); Babanin and Soloviev (1998), but not observed by Ewans (1998).

Finally, it is interesting to note that the trend of the angular spread variation with f/f_p is significantly different between the SWIM and the MFWAM spectra. Whereas for SWIM, the trend is steeper towards the low frequencies than towards the high frequencies, in agreement with most of the above-mentioned results from the literature, the results obtained from MFWAM spectra show a rather symmetric trend towards the high and low frequencies.

In summary, we find that for the SWIM data the behavior of the angular spread with the normalized frequency and with the wave development are in good agreement with the literature whereas the spectra from the MFWAM model show some differences on several points (position in frequency of the minimum, trend with the normalized frequency and trend with wave development at high frequencies). On the other hand, the minimum values of the directional spread close to the frequency peak of the spectrum are very similar for SWIM and MFWAM. Considering every spectra in each categories of wave development, both SWIM and MFWAM mean spectra are slightly broader than the typical values found for the literature. For swell conditions our results indicate a mean angular spread of 32° and 33° for respectively SWIM and MFWAM (Table 3), whereas typical values as reported by Mitsuyasu et al. (1975) correspond to 18.7° (using Equation 7 to convert the s exponent into σ_ϕ values). Similarly, wind sea conditions for our data set correspond to mean directional spread between 31° and 37° (larger values for young wind sea conditions, and SWIM values larger than MFWAM - see Table 3) whereas values reported by Mitsuyasu et al. (1975) correspond to about 23°. However, because of the lack of independent co-located in-situ observations in the situations encountered

in the Southern Ocean, it is difficult to conclude whether this difference is due to the specificity of the data set or to another reason.

5 Discussion and conclusion

We have presented above an analysis of shape parameters of ocean wave spectra obtained from the observations of the SWIM instrument on-board CFOSAT. To our knowledge it is the first time that global statistics and maps of spectral shape parameters and Benjamin-Feir Index are provided from satellite observations. For most of the open ocean conditions (significant wave heights larger than 1.8 m and wavelength larger than 70 m, which corresponds to 73% of the dataset), we can conclude that SWIM provides consistent values of these parameters with sensitivity of these parameters with the sea state condition, compatible with what we know from the literature.

The analysis with co-located observations of SWIM and buoys (8 months of co-located points for 43 NDBC buoys) allows to conclude that SWIM parameters are biased in conditions of low significant wave height ($Hs < 1.8$ m) and short dominant wavelength ($\lambda_p <$ 70 m). We attributed this to the frequent occurence of a spurious peak at the very lowest frequencies of the SWIM spectra, which modifies the shape of the spectra and may dominate in these conditions.

For the other conditions, the comparison of the parameters derived from SWIM to those derived from the MFWAM model, using exactly the same frequency interval and the same expressions for both sources of information, indicates systematic positive differences for the peakedness parameter Qp (sharpest spectra for SWIM), and negative differences for the frequency spread (narrowest spectra from SWIM). These differences are more marked in extreme conditions like those encountered in the Southern Ocean. By analyzing the data sorted according to the sea-state development stage, we could conclude that these differences are the most significant in swell and mature conditions.

Concerning the angular spread of the dominant waves, the main conclusion comes from the Southern Ocean data set because in this data set, there is less uncertainty on the determination of the dominant energy peak and on its association between SWIM and MFWAM data sets. For this subset of data, SWIM spectra show clearly that the narrowest spectra are obtained for swell components, and the broader for young sea components. Compared to MFWAM, the results are not very different. However, SWIM spectra are slightly broader in cases of young wind sea (about 37°) compared to MFWAM (about 33°). When converted into a $2s$ exponent in the $\cos^{2s}$ expression, these values lead to an exponent $2s$ between 6 and 8 for the MFWAM values and between 8 and 10 for the SWIM values. In both cases, this is larger than the classical values of 2 or 4 which are often found in the literature.

Using global data sets and a single version of the model it is difficult to formally conclude on the causes of the differences on frequency distribution, especially for the negative bias at low and high Hs. The complementary comparison of SWIM and buoy data over 8 months of NDBC buoy observations do not help to conclude either, because the large majority of co-located data between SWIM and buoy corresponds to low sea-state and short wavelengths of the dominant waves. However, provided that we exclude these situations, we think that a possible reason may come from the model which uses, as all similar operational 3rd generation wave models, the Discrete Interaction Approximation -DIA- (S. Hasselmann & Hasselmann, 1985a) to represent the energy transfer by non-linear interactions between waves. According to K. Hasselmann (1962), Babanin and Soloviev (1998), Romero and Melville (2010), the non-linear interactions between quadruplets of ocean waves, with different wavenumbers and direction is a dominant mechanism which must be invoked to explain the transfer of energy between directions and between wavenumbers. Although its representation by the DIA gives satisfactorily numerical results in terms

of total energy (or significant wave height), it was previously shown (e.g., Rogers and Van Vledder (2013) with his study with the SWAN model) that the frequency width of the omni-directional spectrum is broader and in less agreement with buoy observations, when this approximation is used than when the exact solution (exact NL) is implemented. However, for practical reasons, the exact solution cannot yet be implemented in operational global numerical models. Therefore, some authors are looking for alternative physical solutions to represent the wave-wave interactions (e.g., Annenkov & Shrira, 2001, 2006, 2018; Annenkov et al., 2021) which would overcome this practical impossibility while providing wave spectra shapes in agreement with observations. Although promising, the approach needs to be validated with observations. It is the aim of future studies to contribute to this kind of study by analyzing SWIM data sets in specific situations.

Finally we have shown in this paper that SWIM observations can be used to provide an index to characterize the deviation from a Gaussian shape of the wave height distribution, and hence the probability of occurrence of extreme waves. The BFI and BFI_{2D} indexes can both be estimated by SWIM and used either for prediction purposes or for climatological surveys.

Acknowledgments

The SWIM data set used here corresponds to the files reprocessed in version 5 and made available by CNES on the ftp serveur of AVISO+ (ftp-access.aviso.altimetry.fr, directory cfosat/swim_l2_op05), accessible to anyone after registration. The authors would like to thank all the CFOSAT teams from CNES, CNSA, NSOAS for their involvement at the different stages in the mission. The buoy data from NDBC (National Data Buoy Center) have been downloaded from the NOAA site https://www.ndbc.noaa.gov/ by Wang Jiuke and kindly provided to us after selecting those at distances larger than 60 km from the coasts. This research was funded by the Centre National d'Étude Spatiale under the TOSCA program.

References

Alpers, W. R., & Brüning, C. (1986). On the relative importance of motion-related contributions to the sar imaging mechanism of ocean surface waves. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-24*(6), 873-885. doi: 10.1109/TGRS.1986.289702

Alves, J. H. G. M., & Banner, M. L. (2003). Performance of a saturation-based dissipation-rate source term in modeling the fetch-limited evolution of wind waves. *J. Phys. Oceanogr.*, *33*(6), 1274-1298. doi: 10.1175/1520-0485(2003)033⟨1274:poasds⟩2.0.co;2

Annenkov, S. Y., Shrira, V., Romero, L., Melville, K., Le Merle, E., & Hauser, D. (2021). Wave development and transformation under strong offshore winds: modelling by dns and kinetic equations and comparison with airborne measurements. In *Egu general assembly 2021.* doi: 10.5194/egusphere-egu21-10437

Annenkov, S. Y., & Shrira, V. I. (2001). Numerical modelling of water-wave evolution based on the zakharov equation. *Journal of Fluid Mechanics*, *449*, 341371. doi: 10.1017/S0022112001006139

Annenkov, S. Y., & Shrira, V. I. (2006). Role of non-resonant interactions in the evolution of nonlinear random water wave fields. *Journal of Fluid Mechanics*, *561*, 181207. doi: 10.1017/S0022112006000632

Annenkov, S. Y., & Shrira, V. I. (2018). Spectral evolution of weakly nonlinear random waves: kinetic description versus direct numerical simulations. *Journal of Fluid Mechanics*, *844*, 766795. doi: 10.1017/jfm.2018.185

Aouf, L., Hauser, D., Chapron, B., Toffoli, A., Tourrain, C., & Peureux, C. (2021). New directional wave satellite observations : Towards improved wave forecasts

and climate description in Southern Ocean. *Geophysical Research Letters*, (in press).

Aouf, L., Hauser, D., Tison, C., & Mouche, A. (2016). Perspectives for directional spectra assimilation: Results from a study based on joint assimilation of cfosat synthetic wave spectra and observed sar spectra from sentinel-1a. In *in proc. ieee int. geosci. remote sens. symp. (igarss)* (p. 58205822). Beijing, China.

Ardhuin, F., Rogers, E., Babanin, A. V., Filipot, J.-F., Magne, R., Roland, A., ... Collard, F. (2010). Semiempirical dissipation source functions for ocean waves. Part I: Definition, calibration, and validation. *Journal of Physical Oceanography*, *40*(9), 1917–1941. doi: 10.1175/2010JPO4324.1

Babanin, A., & Soloviev, Y. (1998). Variability of directional spectra of wind-generated waves, studied by means of wave staff arrays. *Marine and Freshwater Research*, *49*, 89-101. doi: 10.1071/MF96126

Bidlot, J.-R. (2017). *Wave forecast system inter-comparison against in-situ data for september to november 2017* (WMO report). ECMWF.

Blackman, R. B., & Tukey, J. W. (1959). *The measurement of power spectra*. Dover Publications, Inc.

Donelan, M. A., Hamilton, J., Hui, W. H., & Stewart, R. W. (1985). Directional spectra of wind-generated ocean waves. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *315*. doi: 10.1098/rsta.1985.0054

Elfouhaily, T., Chapron, B., Katsaros, K., & Vandemark, D. (1997). A unified directional spectrum for long and short wind-driven waves. *Journal Of Geophysical Research-oceans*, *102*(C7), 15781-15796. Retrieved from `https://archimer.ifremer.fr/doc/00091/20226/` doi: 10.1029/97JC00467

Ewans, K. (1998). Observations of the directional spectrum of fetch-limited waves. *Journal of Physical Oceanography*, *28*, 495-512. doi: 10.1175/1520-0485(1998)028⟨0495:OOTDSO⟩2.0.CO;2

Fan, Y., Hwang, P., & Yu, J. (2020). Surface gravity wave modeling in tropical cyclones. In *Geophysics and ocean waves studies.* IntechOpen. doi: 10.5772/intechopen.93275

Forristall, G., & Ewans, K. (1998). Worldwide measurements of directional wave spreading. *J. of Atmospheric and Oceanographic Technology*, *15*(2), 440–469. doi: 10.1175/1520-0426(1998)015⟨0440:WMODWS⟩2.0.CO;2

Forristall, G., & Greenwood, J. (1998). Directional spreading of measured and hindcasted wave spectra. *Proc. 5th International Workshop on Wave Hindcasting and Forecasting*.

Gagnaire-Renou, E. (2009). *Progress in spectral wave modeling using a quasi-exact method for nonlinear wave-wave interactions* (Theses, Université du Sud Toulon Var). Retrieved from `https://tel.archives-ouvertes.fr/tel-00595353`

Goda, Y. (1976). On wave groups. *Proceedings of the First Behaviour of Offshore Structure Conference76*.

Hasselmann, D. E., Dunckel, M., & Ewing, J. A. (1980). Directional wave spectra observed during JONSWAP 1973. *Journal of Physical Oceanography*, *10*(8). doi: 10.1175/1520-0485(1980)010⟨1264:DWSODJ⟩2.0.CO;2

Hasselmann, K. (1962). On the non-linear energy transfer in a gravity-wave spectrum part 1. general theory. *Journal of Fluid Mechanics*, *12*(4), 481500. doi: 10.1017/S0022112062000373

Hasselmann, S., & Hasselmann, K. (1985a). Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum. Part I: A new method for efficient computations of the exact nonlinear transfer integral. *Journal of Physical Oceanography*, *15*(11), 1369-1377. doi: 10.1175/1520-0485(1985)015⟨1369:CAPOTN⟩2.0.CO;2

Hasselmann, S., & Hasselmann, K. (1985b). Computations and parameteri-

zations of the nonlinear energy transfer in a gravity-wave spectrum. Part II: Parametrizations of the nonlinear energy transfer for application in wave models. *Journal of Physical Oceanography*, *15*(11), 1378-1391. doi: 10.1175/1520-0485(1985)015⟨1378:CAPOTN⟩2.0.CO;2

Hauser, D., Tison, C., Amiot, T., Delaye, L., Corcoral, N., & Castillan, P. (2017, May). Swim: The first spaceborne wave scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(5), 3000-3014. doi: 10.1109/TGRS.2017.2658672

Hauser, D., Tourain, C., Hermozo, L., Alraddawi, D., Aouf, L., Chapron, B., ... Tran, N. (2020). New observations from the swim radar on-board cfosat: Instrument validation and ocean wave measurement assessment. *IEEE Transactions on Geoscience and Remote Sensing*, 1-22. doi: 10.1109/TGRS.2020.2994372

Hwang, P., Wang, D., Walsh, E., Krabill, W., & Swift, R. (2000). Airborne measurements of the wavenumber spectra of ocean surface waves. Part II: Directional distribution. *Journal of Physical Oceanography*, *30*. doi: 10.1175/1520-0485(2001)031⟨2768:AMOTWS⟩2.0.CO;2

Jackson, F. C., Walton, W. T., & Baker, P. L. (1985a). Aircraft and satellite measurement of ocean wave directional spectra using scannig-beam microwave radars. *Journal of Geophysical Research*, *90*(C1), 987–1004. doi: 0148-0227/85/004C-1190$05.00

Janssen, P. A. E. M., & Bidlot, J.-R. (2009). *On the extension of the freak wave warning system and its verification* (Technical Memorandum). ECMWF. doi: 10.21957/uf1sybog

Kerbaol, V., Chapron, B., & W., V. P. (1988). Analysis of ers-1/2 synthetic aperture radar wave mode imagettes. *Journal of Geophysical Research*, *103*(C4), 78337846.

Komen, G. J., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S., & Janssen, P. A. E. M. (1994). *Dynamics and modelling of ocean waves*. Cambridge University Press. doi: 10.1017/CBO9780511628955

Kuik, A., Van Vledder, G., & Holthuijsen, L. (1988). A method for the routine analysis of pitch-and-roll buoy wave data. *Journal of Physical Oceanography*, *18*, 1020–1034. doi: 10.1175/1520-0485(1988)018⟨1020:AMFTRA⟩2.0.CO;2

Le Merle, E., Hauser, D., & Tison, C. (2019). Directional wave spectra at the regional scale with the kuros airborne radar: comparisons with models. *Ocean Dynamics*, *69*(6). doi: 10.1007/s10236-019-01271-5

Longuet-Higgins, M. S., Cartwright, D. E., & D., S. N. (1963). Observations of the directional spectrum of sea waves using the motions of a floating buoy. *Ocean Wave Spectra: Proceedings of a Conference*.

Mitsuyasu, H., Tasai, F., Suhara, T., Mizuno, S., Ohkusu, M., Honda, T., & Rikiishi, K. (1975). Observation of the directional spectrum of ocean waves using a clover-leaf buoy. *Journal of Physical Oceanography*, *5*, 750-760. doi: 10.1175/1520-0485(1975)005⟨0750:OOTDSO⟩2.0.CO;2

Mori, N., Onorato, M., & Janssen, P. A. E. M. (2011). On the estimation of the kurtosis in directional sea states for freak wave forecasting. *Journal of Physical Oceanography*, *41*, 1484-1497. doi: 10.1175/2011JPO4542.1

Pettersson, H., Graber, H., Hauser, D., Quentin, C., Kahma, K., Drennan, W., & Donelan, M. (2003). Directional wave measurements from three wave sensors during the fetch experiment. *Journal of Geophysical Research*, *108*. doi: 10.1029/2001JC001164

Rao, C. (1988). Spectral width parameter for wind-generated ocean waves. *Indian Academy of Sciences Proceedings: Earth and Planetary Sciences*, *97*(2), 173-181. doi: 10.1007/BF02861852

Resio, D. T., Vincent, L., & Argan, D. (2016). Characteristics of directional wave spectra and implications for detailed-balance wave modeling. *Ocean Modelling*,

103, 38 - 52. (Waves and coastal, regional and global processes) doi: 10.1016/j.ocemod.2015.09.009

Rogers, W. E., & Van Vledder, G. P. (2013). Frequency width in predictions of windsea spectra and the role of the nonlinear solver. *Ocean Modelling*, *70*, 52 - 61. (Ocean Surface Waves) doi: 10.1016/j.ocemod.2012.11.010

Rogers, W. E., & Wang, D. (2007). Directional validation of wave predictions. *J. of Atmospheric and Oceanographic Technology*, *24*(3), 504–520. doi: 10.1175/JTECH1990.1

Romero, L., & Melville, W. (2010). Numerical modeling of fetch-limited waves in the gulf of tehuantepec. *Journal of Physical Oceanography*, *40*, 466-486. doi: 10.1175/2009JPO4128.1

Saulnier, J.-B., Clement, A., Falcao, A., Pontes, M., Prevosto, M., & Ricci, P. (2011). Wave Groupiness and Spectral Bandwidth as Relevant Parameters for the Performance Assessment of Wave Energy Converters. *Ocean Engineering*, *38*(1), 130-147. doi: 10.1016/j.oceaneng.2010.10.002

Socquet-Juglard, H., Dysthe, K., Trulsen, K., Krogstad, H., & Jingdong, L. (2005). Distribution of surface gravity waves during spectral changes. *Journal of Fluid Mechanics*, *542*, 195 - 216. doi: 10.1017/S0022112005006312

Stopa, J., Ardhuin, F., Babanin, A., & Zieger, S. (2015). Comparison and validation of physical wave parameterizations in spectral wave models. *Ocean Modelling*, *103*. doi: 10.1016/j.ocemod.2015.09.003

The SWAN Team. (2010). *SWAN Scientific and Technical Documentation, SWAN Cycle III version 40.81* (Tech. Rep.). Delft University of Technology. Retrieved from `http://www.swan.tudelft.nl`

The WISE Group, L., Cavaleri, Alves, J.-H., Ardhuin, F., Babanin, A., Banner, M., Belibassakis, K., ... Young, I. (2007). Wave modelling - the state of the art. *Progress in Oceanography*. doi: 10.1016/j.pocean.2007.05.005

Tolman, H. L., & Chalikov, D. (1996). Source terms in a third-generation wind wave model. *J. Phys. Oceanogr.*, *26*(11), 2497–2518. doi: 10.1175/1520-0485(1996)026⟨2497:stiatg⟩2.0.co;2

Tourain, C., Hauser, D., Hermozo, L., Rodriguez-Suquet, R., Schippers, P., Aouf, L., ... Tison, C. (2020). Cal/val phase for the swim instrument onboard cfosat. In *Ieee international geoscience and remote sensing symposium (igarss)*.

Webb, A., & Fox-Kemper, B. (2015). Impacts of wave spreading and multidirectional waves on estimating stokes drift. *Ocean Modelling*, *96*, 49-64. doi: 10.1016/j.ocemod.2014.12.007

Young, I., Fontaine, E., Liu, Q., & Babanin, A. (2020, 05). The wave climate of the southern ocean. *Journal of Physical Oceanography*. doi: 10.1175/JPO-D-20-0031.1

Zieger, S., Babanin, A. V., Erick, R. W., & Young, I. R. (2015). Observation-based source terms in the third-generation wave model wavewatch. *Ocean Modelling*, *96*, 225. doi: 10.1016/j.ocemod.2015.07.014

Tables

Table 1: Mean and median values of frequency spread distributions for SWIM and MFWAM

	Global scale				Southern Ocean			
	SWIM		MFWAM		SWIM		MFWAM	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Young Wind sea	0.137	0.143	0.133	0.129	0.110	0.108	0.119	0.111
Mature wind sea	0.126	0.129	0.128	0.127	0.098	0.092	0.104	0.098
Swell	0.122	0.126	0.133	0.131	0.097	0.093	0.109	0.105

Table 2: Mean and median values of Qp parameter distributions for SWIM and MFWAM

	Global scale				Southern Ocean			
	SWIM		MFWAM		SWIM		MFWAM	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Young Wind sea	1.82	1.73	2.35	2.24	2.07	2.04	2.05	2.09
Mature wind sea	1.87	1.80	2.06	2.04	2.11	2.09	2.00	2.04
Swell	1.74	1.64	1.66	1.61	2.00	1.96	1.76	1.76

Table 3: Mean and median values of directional spread distributions for SWIM and MFWAM

	Global scale				Southern Ocean			
	SWIM		MFWAM		SWIM		MFWAM	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Young Wind sea	40.0	39.8	31.8	31.0	36.6	35.4	33.4	31.8
Mature wind sea	36.5	35.0	31.1	30.2	32.5	31.2	30.6	29.2
Swell	39.0	38.7	34.7	31.5	32.2	30.0	33.3	30.4

Figures

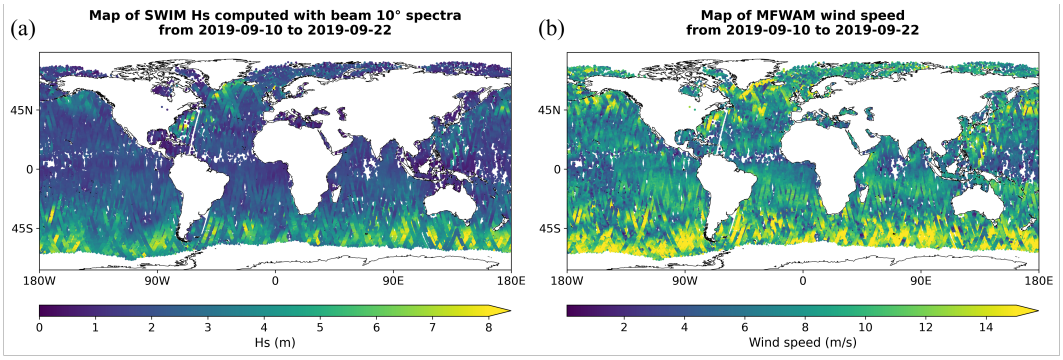

Figure 1: Maps of *(a)* the significant wave height from SWIM, and *(b)* the wind speed from the MFWAM model during the period 10-22 September, 2019.

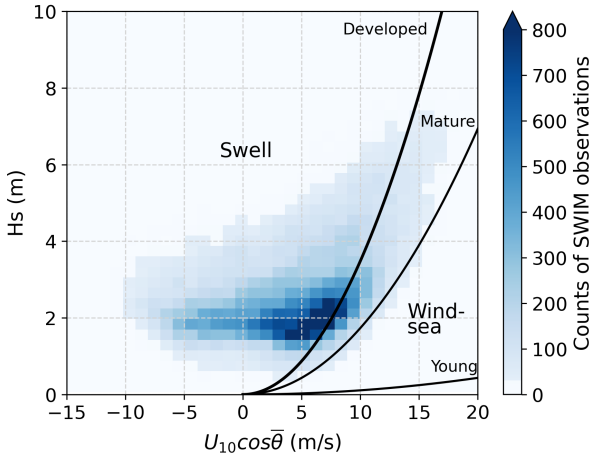

Figure 2: Two-dimensional histogram of the significant wave height Hs as a function of the wind speed projected along the wave direction from MFWAM. Solid lines describe the variation of Hs with wind speed at different wave evolution stages according to the Elfouhaily et al. (1997) spectrum. "Young" corresponds to inverse wave age $\Omega = 2$, "Mature" corresponds to $\Omega = 1$, "Developed" corresponds to $\Omega = 0.84$.

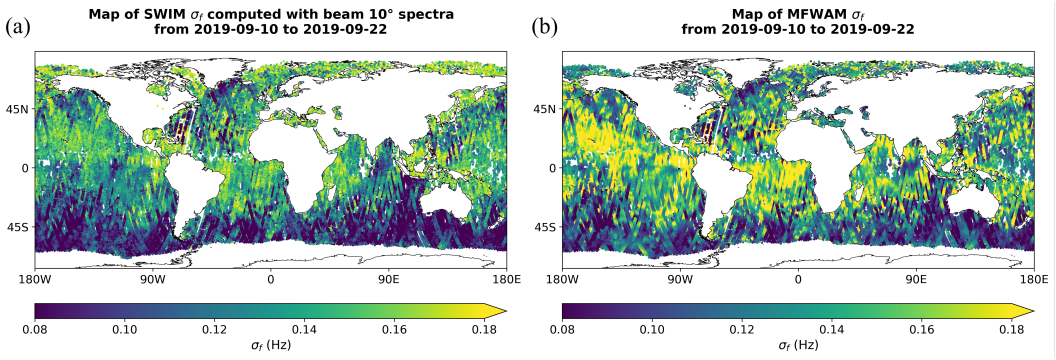

Figure 3: Maps of the frequency spread σ_f calculated with *(a)* SWIM data and *(b)* MFWAM data during the period 10-22 September, 2019.

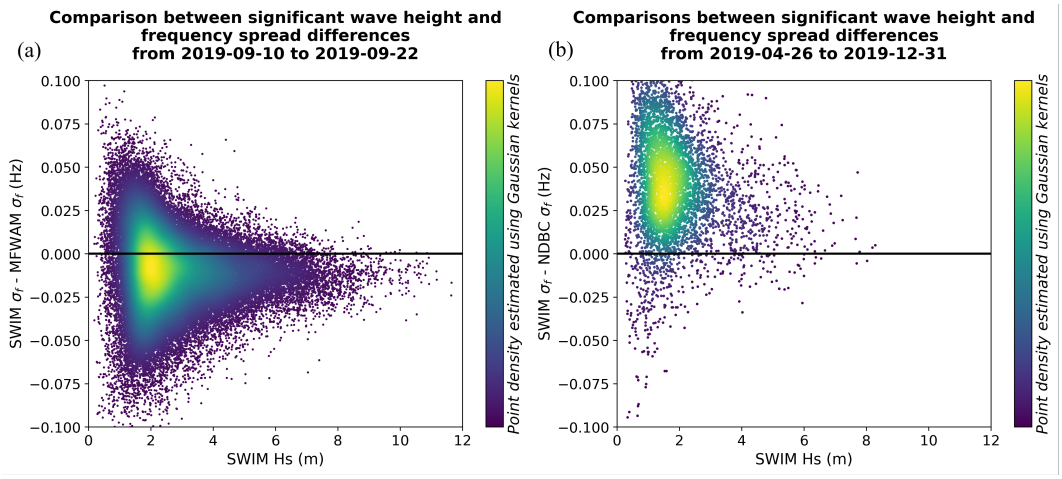

Figure 4: Histogram of the differences of the frequency spread between *(a)* SWIM and MFWAM and *(b)* SWIM and NDBC buoys as a function of significant wave height from SWIM.

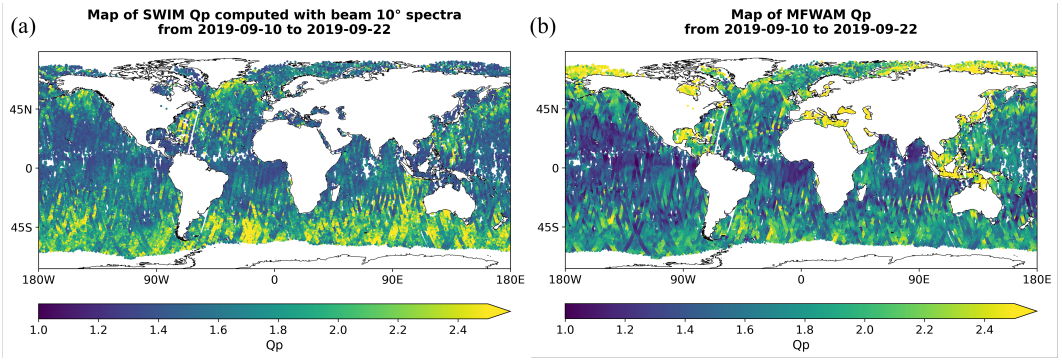

Figure 5: Maps of the Goda peakedness factor *Qp* calculated with *(a)* SWIM data and *(b)* MFWAM data during the period 10-22 September, 2019.

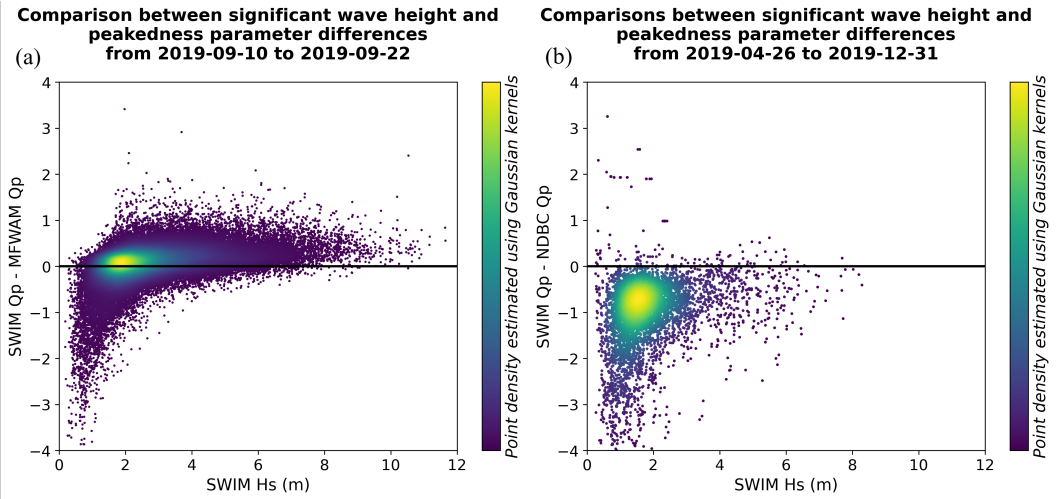

Figure 6: Histogram of the differences of the Goda peakedness factor Qp between *(a)* SWIM and MFWAM and *(b)* SWIM and NDBC buoys as a function of significant wave height from SWIM.

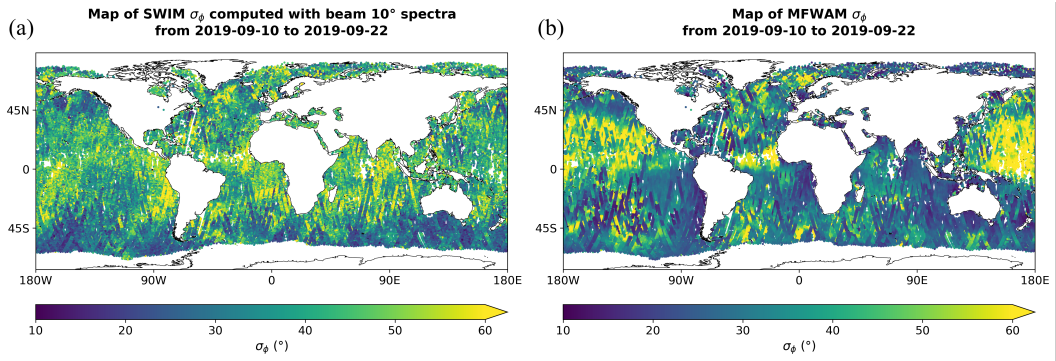

Figure 7: Maps of the directional spread σ_ϕ calculated with *(a)* SWIM data and *(b)* MFWAM data during the period 10-22 September, 2019.

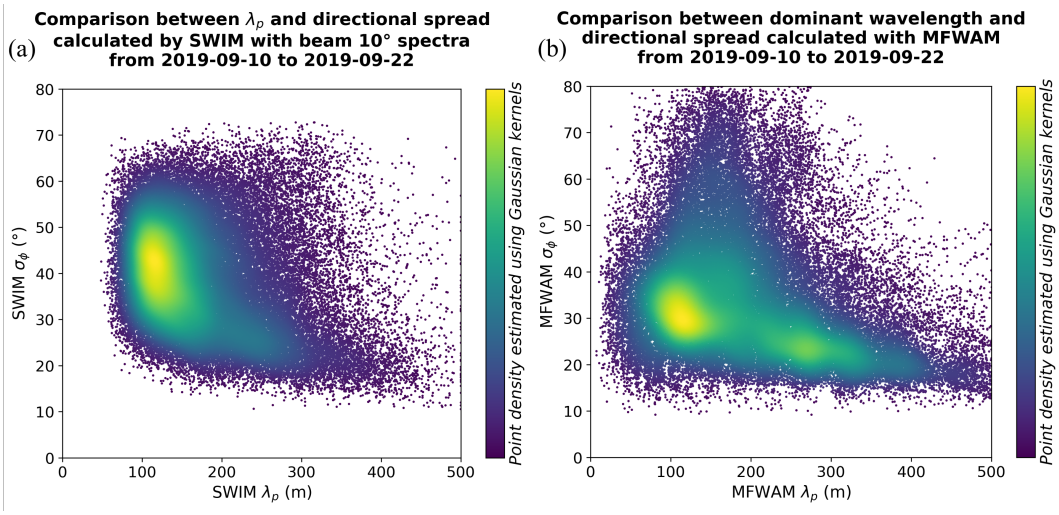

Figure 8: Histograms of the directional spread at the peak of the wave spectrum σ_ϕ as a function of the dominant wave length. *(a)* for SWIM, *(b)* for MFWAM.

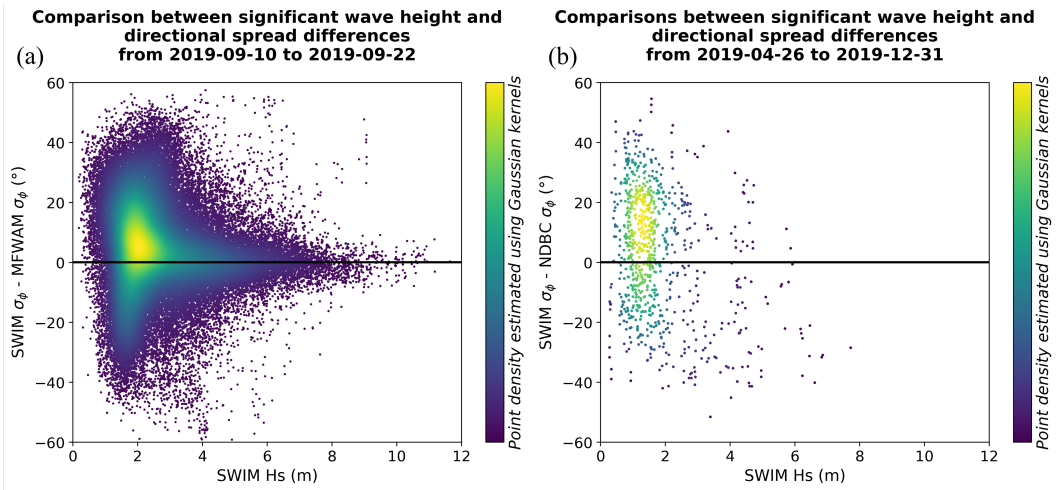

Figure 9: Histogram of the differences of the directional spread σ_ϕ between *(a)* SWIM and MFWAM and *(b)* SWIM and NDBC buoys as a function of significant wave height from SWIM.

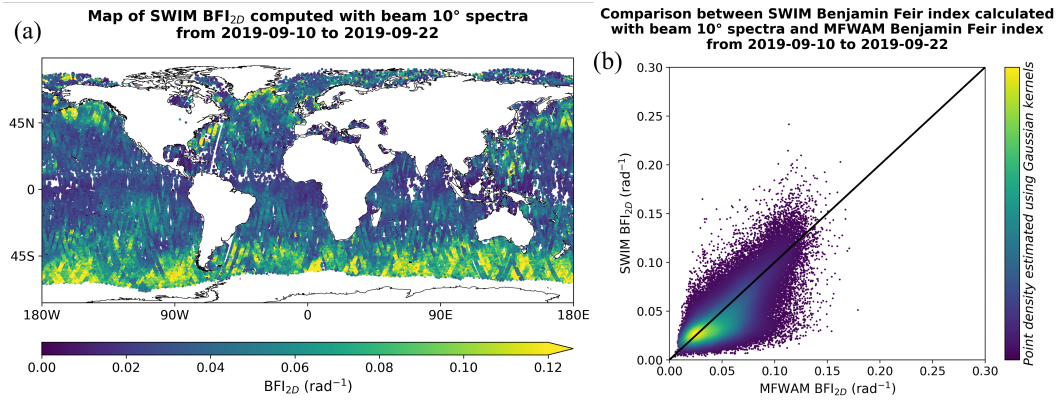

Figure 10: *(a)*: Map of the BFI_{2D} parameter for SWIM. *(b)* BFI_{2D} from SWIM (vertical) compared to BFI_{2D} from MFWAM (horizontal).

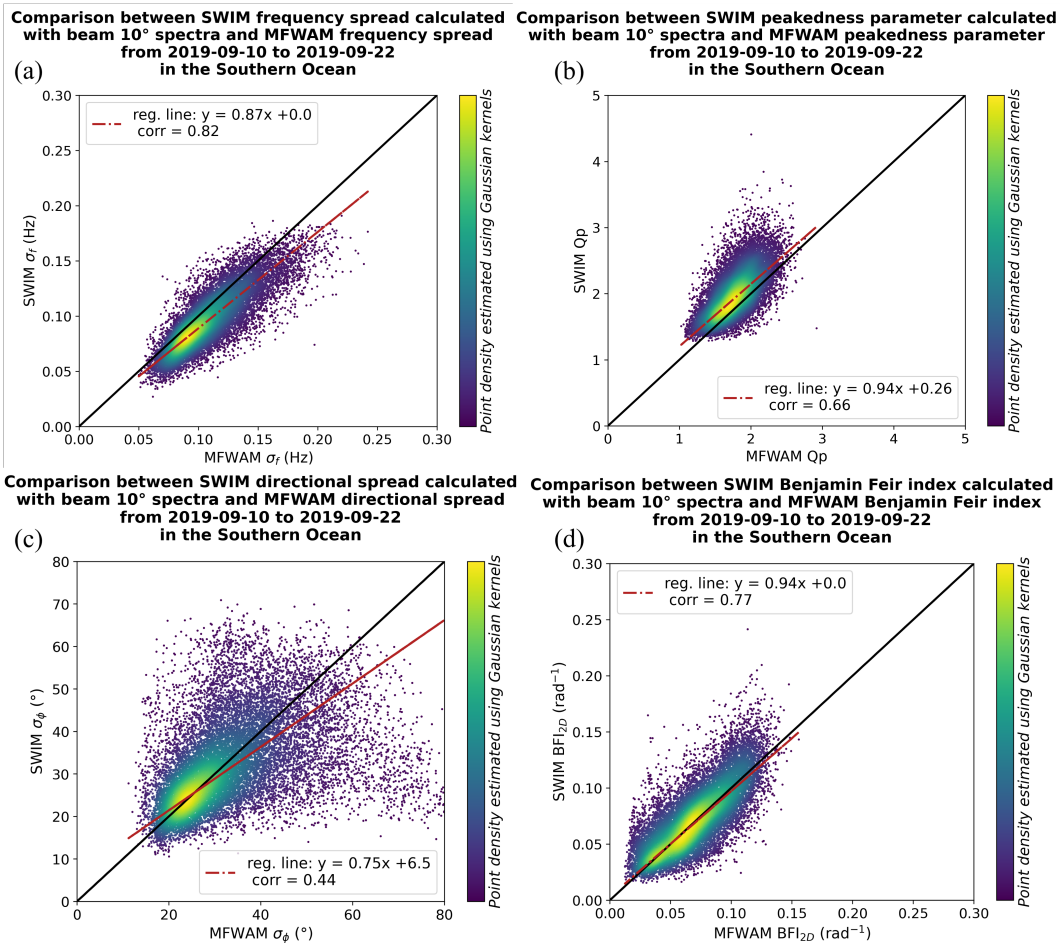

Figure 11: Comparison between SWIM and MFWAM data in the Southern Ocean. *(a)*: frequency spread, *(b)*: *Qp*, *(c)*: directional spread, *(d)*: BFI_{2D}.

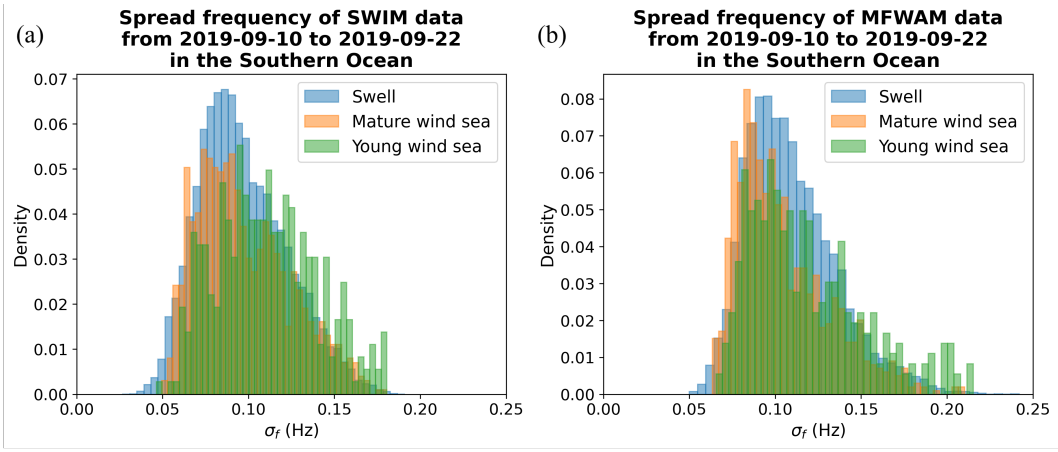

Figure 12: Distributions of frequency spread for individual spectra in the Southern Ocean. *(a)*: for SWIM data, *(b)*: for MFWAM data. The different colors refer to different wave developments: cyan for swell, orange for mature wind sea, green for young wind sea (see text).

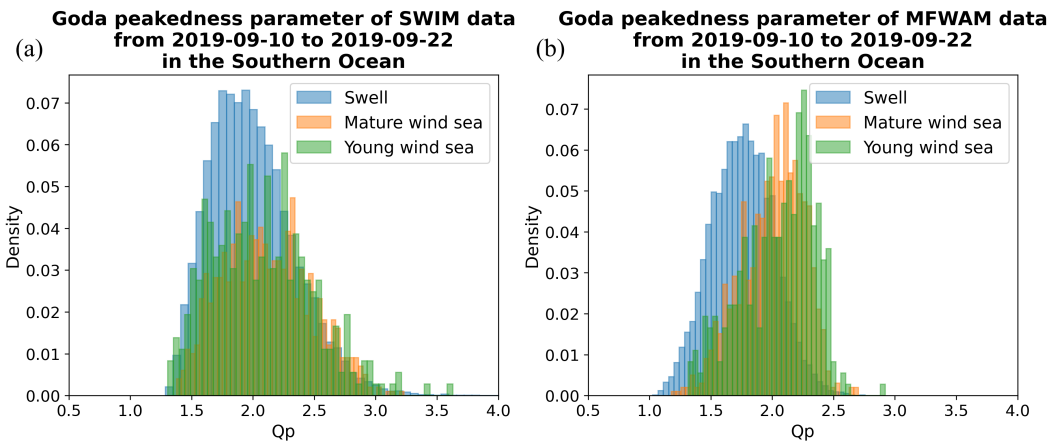

Figure 13: Distributions of the peakdness parameter Qp for individual spectra in the Southern Ocean. *(a)*: for SWIM data, *(b)*: for MFWAM data. The different colors refer to different sea-state categories: cyan for swell, orange for mature wind sea, green for young wind sea (see text).

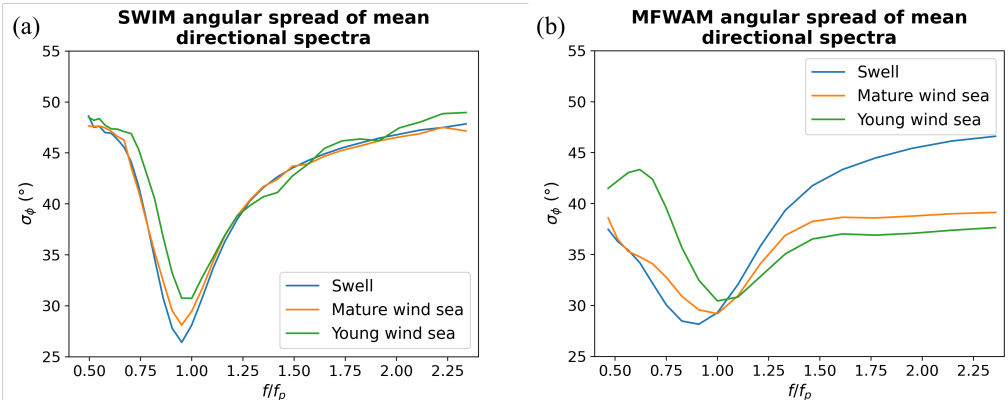

Figure 14: Directional spread as a function of the adimensional frequency for each mean directional spectrum for *(a)*: SWIM , *(b)*: MFWAM. The color code refer to mean spectra estimated for different sea-state category: cyan for swell, orange for mature wind sea and green for young wind sea.