Wave Groups and Small Scale Variability of Wave Heights Observed by Altimeters

Marine de Carlo, Fabrice Ardhuin, Annabelle Ollivier, Adrien Nigou

Wave groups and small scale variability of wave heights observed by altimeters

Marine De Carlo[1], Fabrice Ardhuin[1], Annabelle Ollivier[2], Adrien Nigou[2]

[1]Univ Brest, CNRS, Ifremer, IRD, Laboratoire d'Océanographie Physique et Spatiale (LOPS), IUEM, F29280, Plouzané, France.

[2]Collecte Localisation Satellite (CLS), Ramonville Saint-Agne, France

Key Points:

- Wave groups contribute to small-scale fluctuations in altimeter wave heights, explaining 25% of the variance measured by CFOSAT in 80 km.
- For the same wave height, fluctuations are larger in the presence of long and narrow-banded waves, typical of swell-dominated conditions.
- Altimeters smooth out scales shorter than the square root of half the H_s times the altitude, and distort spatial patterns at that scale.

Corresponding author: Marine De Carlo, mdecarlo@ifremer.fr

Abstract

Recent satellite altimeter retracking and filtering methods have considerably reduced the noise level in estimates of the significant wave height (H_s), allowing to study processes with smaller spatial scales. In particular, previous studies have shown that wave-current interactions may explain most of the variability of H_s at scales 20 to 100 km. As the spatial scale of the measurement is reduced, random fluctuations emerge that should be associated to wave groups. Here we quantify the magnitude of this effect, and the contribution of wave groups to the uncertainty in H_s measurements by altimeters, with a particular focus on extreme extra-tropical storms. We take advantage of the low orbit altitude of the China-France Ocean Satellite (CFOSAT), and the low noise level of the nadir beam of the SWIM instrument. Our estimate of wave group effects uses directional wave spectra measured by off-nadir beams on SWIM, and signal processing theory that gives statistical properties of the wave envelope, and thus the local wave heights, from the shape of the wave spectrum. We find that the standard deviation of H_s associated to wave groups is a function of satellite altitude, wave height and spectral bandwidth. For CFOSAT these fluctuations generally account for about 25% of the variance measured over a 80 km distance. This fraction is largest in storms and in the presence of long swells. When the estimated effect of wave groups is subtracted from the variance of H_s measurements, the remaining variability is higher in regions of strong currents.

Plain Language Summary

Satellite altimeters routinely provide measurements of the height of ocean waves, and improved instruments or processing techniques have led to more precise and detailed measurements. Here we use a combination of simulations and data from the China France Ocean satellite (CFOSAT) to interpret the small scale fluctuations in wave height measurements as the effect of wave "groups" which are fluctuations of the heights of consecutive waves associated to random waves. Due to spatial averaging within the radar footprint, we find that the fluctuations of significant wave heights (H_s) associated to wave group are a function of satellite altitude, wave height and other properties of the ocean waves. For CFOSAT, wave groups give a standard deviation of H_s that is of the order of 3 to 5% of H_s, typically half of the standard deviation in H_s measurements.

1 Introduction

Wind-waves impact all activities at sea, air-sea interactions and remote sensing, and there is a general need for obtaining more accurate and higher resolution information about the sea state. Today, satellite radar altimetry is the most extensive source of measurements with a global coverage, providing routine estimates of the significant wave height H_s (Ardhuin et al., 2019). As these data are getting used for a wide range of applications, it is important to understand what can be measured with altimeters, at what scale and with what uncertainty.

The fundamental measurement of an altimeter is the echo power as a function of the travel time of radar pulses. This function is known as a "waveform". Time is converted to "range", i.e. the distance between the radar and the ocean surface, and this waveform is discretized in range gates with a resolution δ_R, with some examples shown in Appendix A. From the shape of the waveform one can estimate H_s. The horizontal scale of the measurement was particularly discussed by Chelton et al. (1989), who introduced the concept of *oceanographic footprint*. This footprint is the ocean area that contains the sea surface points that contribute to the measurement of sea level and wave height, and it is a disc of radius

$$r_C = \sqrt{\frac{2h(H_s + \delta_R)}{1 + h/R_E}} \tag{1}$$

where h is the satellite altitude, R_E is the Earth radius, H_s is the significant wave height and the range resolution $\delta_R = c/(2B)$ is defined by the radar bandwidth B and the speed of light c. We note that all Ku-band altimeters have used $B = 320$ MHz giving $\delta_R = 0.47$ m. With $B = 500$ MHz, SARAL-AltiKa uses $\delta_R = 0.32$ m. As a result, r_C is always larger than 1 km. That size of the oceanic footprint corresponds to a single radar pulse. The sea echoes detected by the radar come from facets of the sea surface that are horizontal and reflect the signal back to the radar. These facets are randomly distributed within the oceanic footprint, with ranges that vary over many times the electromagnetic wavelength. The signal recorded in any given range gate is thus the sum of a large number of echoes with random phases, giving rise to large fluctuation in measured power, generally known as Rayleigh fading. In the context of radar remote sensing, these fluctuations are called "speckle noise" and have generally been considered to be the dominant source of altimeter measurement noise (Quartly et al., 2001; Tourain et al., 2021). This noise is reduced by averaging echoes from many pulses over a fraction of a second,

and these averaged waveforms are processed to estimate H_s. Because low Earth orbit satellites fly over the ocean at a speed around 7 km/s, averaging altimeter data over 0.05 s corresponds to a spatial average over 350 m, which is much smaller than r_C and thus does not change much the effective footprint of the measurement.

Even with this averaging, H_s estimates are typically more noisy than buoy measurements, which has led users of altimetry data to take longer averages of H_s, typically 1 to 10 s, corresponding to a distance that spans 7 to 70 km (Dodet et al., 2022). While it effectively reduces noise, such averaging also blurs potentially interesting features, in particular the peaks of storms, coastal gradients (Passaro et al., 2021), and the signature of surface currents (Quilfen & Chapron, 2019). Away from surface current gradients and coastlines, sea states are uniform over scales of the order of 70 km (Tournadre, 1993). Still, within these scales, the random nature of the wave field is another source of expected geophysical variability. Theoretical analysis, in situ time series and remote sensing (Borge et al., 2004) show that small scale variations in wave height contain a signature of wave groups that can be estimated from the Power Spectral Density (PSD) of the surface elevation, hereinafter simply called "wave spectrum" (Arhan & Ezraty, 1978; Tayfun & Lo, 1989). These groups are the result of the linear superposition of many independent wave trains. Wave groups have typical time scales of a few tens of seconds to a few minutes, that translate to spatial patterns at scales of a few kilometers, hence around the possible resolution limit of altimeters, of the order of r_C. At larger scales, non-linear wave-wave interactions should contribute to fluctuations at scales 10 to 20 minutes, with spatial scales around 10 km, that should be important for wave growth (Lavrenov, 2001) and may contain information on the wave period (Badulin, 2014).

Co-location of altimeter, buoy and model data with wave heights from 1 to 8 m, has been used to estimate a typical uncertainty 7% for $H_s > 2$ m, in the case of 1 s averaged altimeter data (Dodet et al., 2022). Unfortunately, no estimate can be made for higher values due to the lack of co-located buoy and altimeter data. Understanding what makes up this uncertainty will help extrapolate uncertainties to higher values of H_s, providing a better understanding of the climatology of sea state extremes. From the principle of the estimation of wave heights from radar waveforms (Brown, 1977), satellite altimeters should be able to measure H_s values exceeding 30 m with a relative precision that gets better for higher values of H_s because the signal is spread out over a wider part of the waveform. At the same time, the higher winds that often occur with high waves

will lead to a lower signal recorded by the radar and a lower signal to noise ratio for the waveform. So far, only numerical models could be compared to the highest values of wave heights, and their random differences is usually lowest for the highest wave heights (Alday et al., 2021; Alday & Ardhuin, 2023). The highest values of H_s reported by altimeters, up to 20.1 m, are consistent with all other observations including the presence of swells with very long periods (Hanafin et al., 2012). There is thus no fundamental reason to doubt that altimeters can measure the highest possible wave heights, but there is not yet a clear understanding of biases and random errors for H_s above 8 m.

In the present paper we focus our analysis on the fluctuations of H_s associated to wave groups and its contribution to Delay-only altimeters that provide the existing reference time series for wave climate analysis (Timmermans et al., 2020; Dodet et al., 2020). The main question that we wish to answer is: how much wave groups contribute to the variability in H_s measurements? For this we take advantage of the unique opportunity provided by the SWIM instrument onboard the China-France Ocean Satellite (CFOSAT). SWIM provides both directional wave spectra from which we compute the spectrum of the wave envelope that contains wave groups, and along-track nadir altimetry. Our analysis uses SWIM data over the globe for two full years 2020 and 2021.

We start with two illustrative and contrasting examples in section 2, before providing results for the globe in section 3. Discussions and conclusion follow in section 4. A side question that we had to address is: how does an altimeter measure H_s over a realistic surface that contains local perturbations associated with wave groups? For this we used a simplified simulated altimeter with numerical results shown in section 2 and an analytical derivation in Appendix A. Those results suggest that altimeters report a particular kind of average of wave heigts over a radius that is close to $r_C/2$. When using a least-square fit to radar waveforms, estimated H_s give a spurious amplification of true H_s perturbations located at a distance around $r_C/4$ from nadir, and are blind to perturbations located right at the nadir.

2 One particular storm and a theory of wave groups

From now on, we will use the notation $\widehat{H}_s$ for estimates of the significant wave heights provided by altimeter data, to clearly differentiate these from the true significant wave height H_s. $\widehat{H}_s$ exhibits spatial variability that may be related to a true spatial variation

of H_s. As we will be considering different sizes of variability, we introduce the notation std(x, d) and var(x, d) that represent respectively the standard deviation and the variance of a variable x performed over a spatial distance d.

As described in Hauser et al. (2017, 2021), the instrument SWIM is a Ku-band wave scatterometer that illuminates successively 6 incidence angles (0°, 2°, 4°, 6°, 8° and 10°). The nadir beam (0°) works as all previous Poseidon radar altimeters and provides $\widehat{H}_s$ measurements every 0.22 s, using an average over 0.055 s. As a result, the nadir beam data is expected to be similar to data from previous Ku-band altimeters, such as Poseidon-3B on Jason-3, with the specific difference given by a lower data rate (4.5 Hz instead of 20 Hz) and a different measurement geometry associated to a rather low orbit. In principle, the low orbit altitude $h = 519$ km of CFOSAT makes it possible to resolve smaller scale variations of H_s as r_C is reduced by a factor 1.4 compared to the Jason satellites that orbit at 1340 km altitude. The low noise level of the satellite and specific processing of the SWIM instrument also contribute to the capability of SWIM to resolve smaller along-track scales in the variations of H_s compared to previous altimeter datasets (Tourain et al., 2021).

The off-nadir beams use the concept of the wave spectrometer (Jackson et al., 1985) based on a real-aperture radar and the normalized radar cross-section (NRCS) sensitivity to local surface slope at near-nadir incidences, providing estimates of the directional wave spectra (with a 180° ambiguity in direction). The CNES mission center (or CFOSAT Wind and Waves Instrument Center - CWWIC) provides Level 2 products, hereafter called L2, both for the nadir beam and the off-nadir beams 6° to 10°. The off-nadir L2 products consist of 2D wave spectra discretized into 12 directions evenly spaced from from 0 to 165°, and 32 wavenumbers forming a geometric progression from 0.0125 to 0.28 m^{-1} with a common ratio of 1.105. These spectra are constructed from overlapping of antenna scans over 180° (on each side of the track) over boxes of about 70 km by 90 km. In order to allow comparison, the nadir product is resampled by averaging values of $\widehat{H}_s$ over the box size (c. 80 km along track), its variation at this scale is quantified by taking its standard deviation over the same distance.

The Ifremer Waves and Wind Operational Center (IWWOC) is in charge of developing and testing different processing and provides an alternative Level 2 product for off-nadir beams. This product is referred to as L2S product and consists of 1D wave mod-

ulation spectra, one for each measurement azimuth. Whereas the L2 product uses the nadir H_s to rescale the spectrum, the L2S product is based on a theoretical modulation transfer function to transform the NRCS spectra into surface elevation spectra (Jackson et al., 1985). Also the L2 product uses a maximum wavelength of 500 m in order to avoid amplifying noise, where the L2S product does not use such a fixed value for the maximum wavelength. As a result, L2 spectra are often badly distorted in the most severe storms where the spectrum is dominated by waves longer than 500 m. Because we focused on extreme storms, we have used the L2S product to generate spectra similar to those in the L2 product, using the same "boxes" as the L2 product and rescaling the spectrum energy to correspond to the root mean square average of the significant wave height in that box. This rescaling is particularly motivated by our investigation of wave properties along the nadir track, and it corrects an average bias of the order or 14% for the total energy of the spectrum. Alternative data processings are discussed in section 4. In practice we also used the homogeneously reprocessed Level 2+ (L2P) product provided by the Copernicus Marine Environment Monitoring System (CMEMS), in particular we used the quality flags specific to that product.

2.1 Significant wave height variability in Storm Dennis

On February 14^{th} 2020, the European windstorm Dennis, which became one of the most intense extratropical cyclones ever recorded, underwent through its explosive intensification in the middle of the North Atlantic. Around 9:10 UTC that same day, Dennis was sampled by CFOSAT, with $\widehat{H}_s$ values up to 19.7 m for the native (4.5 Hz) sampling, and 17.9 m for the 1 Hz sampling (averaging over 1 second). Fig. 1.a shows a model snapshot of H_s in the north Atlantic and the corresponding descending track of CFOSAT, while Fig. 1.b shows the altimeter-estimated significant wave heights $\widehat{H}_s$ for the three different samplings: native (4.5 Hz), 1 Hz and 80 km box averaged.

On the periphery of the storm, where the average H_s is around 10 m, we were struck by the factor two difference in std($\widehat{H}_s$)that spans more than 420 km (1 minute of data). Our working hypothesis is that this variability of $\widehat{H}_s$ may be dominated by fluctuations associated to wave groups. These fluctuations have different magnitudes and spatial scales which can be estimated from the directional wave spectrum (Longuet-Higgins, 1984). Hence CFOSAT is a unique instrument for studying this effect as we have both $\widehat{H}_s$ variability along the satellite track and directional wave spectra. In the following, we illustrate the

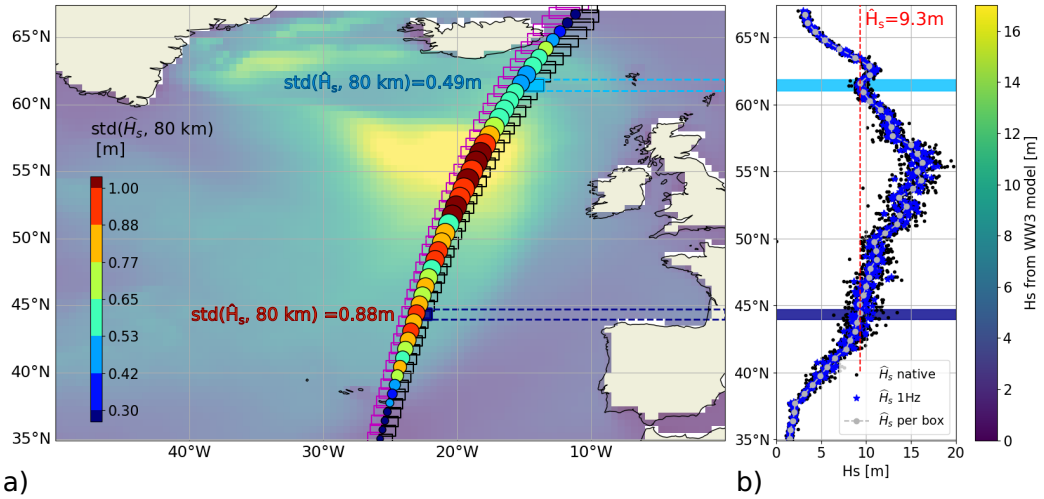

Figure 1. a) Map of significant wave heights in the North Atlantic at 09:00 on 14 February 2020, as provided by the model hindcast of (Alday et al., 2021), overlaid with circles located at the center of SWIM box estimates for the L2 wave spectra. Circles are sized by the L2 H_s estimate and color corresponds to std($\widehat{H}_s$); b) corresponding $\widehat{H}_s$ values as a function of latitude (y-axis) : black small dots represent native measurements at 4.5 Hz, blue stars represent the 1 Hz averaged and grey circles represent the $\widehat{H}_s$ averaged over a box. Two boxes are selected for the case study: box A — highlighted in light blue — is at 62°N, and box B — in dark blue — is at 44°N.

expected signature of wave groups for the two sea state conditions that correspond to the particular SWIM boxes highlighted in light and dark blue. It is worth noting that in these two examples, the $\widehat{H}_s$ values obtained from the sum of the L2S spectrum prior rescaling are around 7.5 m, which is lower than the 9 m given by the nadir beam and used in the L2 product and in this study to rescale the spectrum energy.

2.2 Variability of $\widehat{H_s}$ and envelope spectrum

The patterns of individual waves vary with the shape of the wave spectrum, as illustrated in Fig. 2. A key difference between the north-side (left column) and south-side (right column) of storm Dennis is that the south-side has a longer peak wavelength around 600 m, and a narrower spectrum, in particular in directions. The smaller width in directions gives longer wave crests while the smaller width in wavenumber magnitude gives

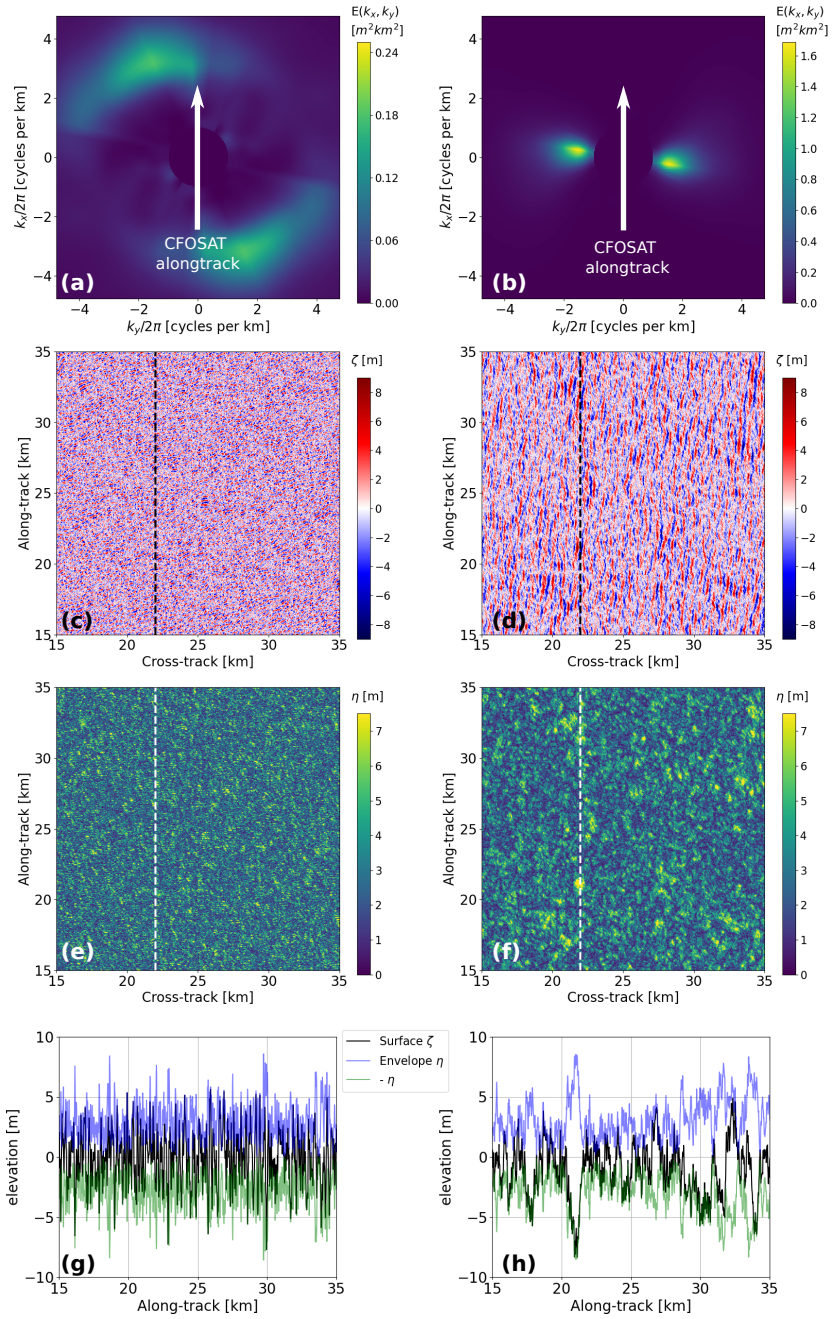

Figure 2. Left column corresponds to our chosen northern CFOSAT box, and right column to the southern box. From top to bottom, (a, b): L2S wave spectra $E(k_x, k_y)$ as provided by IWWOC. (c, d): simulated surface elevation maps generated from the wave spectra using random phases, (e, f): envelope η of the surface elevation, (g, h) along-track slices of elevation ζ and $\pm\eta$ at $x = 22$ km.

longer wave groups (see Longuet-Higgins, 1984, for a detailed definition of the length of wave groups).

Although the surface elevations in Fig. 2.c or 2.d are realizations of a uniform sea state each given by a well defined wave height H_s and spectrum shape, any measurement that is sensitive to the surface elevation over a finite area will provide an estimate $\widehat{H}_s$ that differs from H_s due to spatial fluctuations. We will now link this estimate $\widehat{H}_s$ in the case of a radar altimeter, to the properties of the envelope. For this purpose we need to define a local wave height.

2.2.1 Definition of a local wave height

Let ζ_c be the complex surface such that $\zeta = \mathrm{Re}(\zeta_c)$ is the free surface. The envelope η of the signal is defined by $\eta = |\zeta_c|$. This defines a local amplitude of the signal, that does not contain the small scale crest-to-trough (positive to negative) variations of the original surface. From this envelope η we define the wave height H_r as a spatial average over a disc of radius r

$$H_r(x,y) = 4\sqrt{\frac{2}{\pi}}(\eta \otimes g_r)(x,y) \tag{2}$$

where $\otimes$ is the convolution operator and g_r is a filtering kernel of radius r. Under the Gaussian approximation of the distribution of sea surface elevations this spatial average actually converges to the usual significant wave height H_s.

The envelope of a signal is known to act as a low-pass filter and its fluctuations, larger than those of an individual wave, can be related to wave groups, both in size and amplitude (Arhan & Ezraty, 1978; Longuet-Higgins, 1984; Masson & Chandler, 1993).

Hence wave groups may contribute to the fluctuations of the estimated $\widehat{H}_s$ provided by the nadir beam of SWIM, as indicated on Fig. 1.b. We will now attempt to quantify that contribution. In order to understand how much wave groups may contribute to $\widehat{H}_s$ fluctuations in satellite data, we have to address two questions: First, what are the scales affected by wave groups?

And second, what are the scales of the H_r variation that are resolved by satellite altimeters?

2.2.2 Surface elevation envelope and H_r spectrum

One simple way to quantify the different scales present in the envelope is to compute its spectrum. The most simple theoretical result comes directly from the theory of Fourier transforms that gives the spectrum of a product of functions as the convolution of the Fourier transforms. In our case, the envelope squared is the product of the elevation by its complex conjugate, and this is true for spectra in one or two dimensions. For waves in one dimension with wavenumber k, the spectrum of the envelope squared $\Psi_2(k)$ is the convolution of the spectrum of the single-sided surface elevation spectrum $E(k)$ by itself,

$$\Psi_2(k) = 8 \int_0^\infty E(u)E(u+k)\mathrm{d}u, \tag{3}$$

and we note that $\Psi_2(k)$ is also single-sided.

In practice people have rather studied the variations of H_s and not that of H_s^2. Although the details of the theory are more complex, the important result is that, for low frequencies, the spectrum of the envelope $\Psi(k)$ has the same shape as the spectrum of the envelope squared $\Psi_2(k)$ (Rice, 1944). More specifically, Tayfun and Lo (1989) have showed that a good approximation for the spectrum of the envelope is given by

$$\Psi(k) = \frac{8-2\pi}{H_s^2}\Psi_2(k) \tag{4}$$

This same result is valid for spectra in two dimensions. We now consider the double-sided wave spectrum $E(k_x, k_y)$, defined for (k_x, k_y) in the entire wavenumber plane and centrally symmetric, the region of the envelope spectrum for $k \ll k_p$, with k_p the wavenumber peak, is proportional to

$$\Psi_2(k_x, k_y) = 8 \int_{-\infty}^\infty \int_{-\infty}^\infty E(u,v)E(u+k_x, v+k_y)\mathrm{d}u\mathrm{d}v, \tag{5}$$

in which Ψ_2 is also double-sided.

From eq. (2), the spectrum of H_r is

$$\Psi_{H_r}(k_x, k_y) = \frac{32}{\pi}\Psi(k)G_r(k_x, k_y) \tag{6}$$

$$= \underbrace{\frac{32}{\pi}\frac{8-2\pi}{H_s^2}\Psi_2(k_x, k_y)}_{\Psi_{H_0}(k_x,k_y)} G_r(k_x, k_y) \tag{7}$$

with H_s the usual significant wave height and G_r the square of the Fourier transform of the filtering kernel g_r. We call $\Psi_{H_0}(k_x, k_y)$ the spectrum obtained before applying the filter $G_r(k_x, k_y)$.

Integrating this spectrum for $k_x > k_1$, with the x-axis taken in the along-track direction, amounts to integrating the expected variance up to the cut-off distance, $d_1 = 2\pi/k_1$, giving $\mathrm{var}(H_r, d_1)$. The group-induced variation of $\widehat{H}_s$ is thus equal to

$$\mathrm{var}(H_r, d_1) = 2\int_{-\infty}^{\infty}\int_{k_1}^{\infty} \Psi_{H_0}(k_x, k_y) G_r(k_x, k_y) \mathrm{d}k_x \mathrm{d}k_y \tag{8}$$

We now need to estimate the filter G_r associated with the SWIM altimeter.

2.3 Altimeter measurements over varying wave heights H_r

Going back to the fundamental altimeter measurement that is the waveform, Brown (1977) assumed a uniform ocean reflectivity and showed that the waveform is an area-weighted histogram of the radar echoes as a function of range R. Over a flat sea surface, this histogram is a Heaviside function because the part of the ocean surface with ranges between R and $R + \delta_R$ is an annulus of radius $r = \sqrt{R^2 - h^2}$ centered on the nadir point, with an area $2\pi R\delta_R$ that is almost constant as long as $R \approx h$. In the presence of waves, echoes from the sea surface at the elevation $z = \zeta$ and at the nadir (vertically under the satellite), will have a higher range when $\zeta < 0$. These echoes will have the same range as other echoes from $z = 0$ and horizontal locations away from nadir. Given the very small incidence angles, the change in range caused by waves is $\Delta_R = -\zeta$. For a Gaussian distribution of ζ with standard deviation $\sigma_H = H_s/4$, the presence of waves gives a smoothing of the histogram. Here we use the most simple theoretical waveform shape that is obtained in the limit of a very broad radar antenna pattern (Brown, 1977),

$$w_B(R, \sigma_H) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{R-h}{\sqrt{2}\sigma_H}\right)\right]. \tag{9}$$

When "retracking" altimeter data, eq. (9) is inverted, giving the estimate $\widehat{H}_s$ equal to 4 times the σ_H of the theoretical waveform that best fits the data. In practice the theoretical waveform may also include effects of the antenna pattern and mispointing. Different fitting methods have been developed to reduce the effect of noise or spurious echoes in the measured waveform (Rodriguez, 1988; Passaro et al., 2014; Tourain et al., 2021). Another important assumption needed to obtain the Brown waveform is that the sea state is homogeneous within the footprint. We thus have to discuss what sets the scale of the footprint, or more precisely where are the points on the sea surface that give the distinctive shape of the waveform and allow the fit to distinguish different values of H_s.

Compared to a flat sea surface, the elevation ζ at a distance r from the nadir point will change the range R of the surface point and make it look as if it was located at a different distance $r+\delta_r$, so that points from different locations on the sea surface will map to the same range R. This is the same "range bunching" or "overlay" or "surfboard effect" that is common to all radar systems (Peral et al., 2015). Following Chelton et al. (1989) we can estimate the apparent horizontal displacement. For a satellite altitude h and using $\zeta \ll R$, the calculation for a flat mean sea surface gives

$$\delta_r \simeq \sqrt{r^2 - 2h\zeta} - r. \tag{10}$$

For a spherical Earth of radius R_E, ζ should be replaced by $\zeta/(1+h/R_E)$. In the particular case where $\zeta = -H_s$ and $r = 0$, δ_r is the radius r_C of the oceanographic footprint defined by Chelton et al. (1989), and given by eq. (1), when the range resolution δ_R is neglected compared to H_s. For $H_s = 10$ m, and $h = 519$ km, this gives $r_C = 3.3$ km. Based on the Gaussian distribution of the sea surface elevation, there is only a 0.003% probability that $\zeta > H_s$. Hence, we have the same negligible probability that points located at r_C from nadir (i. e. at the edge of the "Chelton footprint") contribute to the waveform at ranges $R < h$, i.e. in the first half of the rising part of the waveform. We may thus give the following interpretation of r_C: points located at $r > r_C$ are outside of the footprint and have a very limited impact on the estimated value $\widehat{H}_s$.

If wave heights H_r vary as a function of distance to nadir, then the waveform does not follow exactly the Brown form, as detailed in Appendix A. As different regions of the waveform contain different regions of the sea surface, one could imagine fitting different parts of the waveform to measure variations in H_r as a function of distance from nadir. The theoretical limit to this capability is the blurring due to range bunching over a distance of the order $r_C/2$. Speckle noise is another limiting factor, and probably the main one in practice for small values of H_s.

Based on the analysis in Appendix A we expect that variations of H_r at scales much smaller than $r_C/4$ will be smoothed out in altimeter data, whereas variations at scales much larger than $r_C/2$ have no effect on the waveform that will follow the Brown shape for the local wave height. For our analysis of CFOSAT data we will define an "effective altimeter radius" r_a such that the variance associated to the random fluctuations of H_{r_a} — the envelope filtered with g_{r_a}, a Gaussian filter of standard deviation r_a — is the same

as that produced by an altimeter. The actual shape of the "altimeter filter" is discussed in Appendix A

2.4 Estimation of the equivalent r_a scale for an idealized altimeter

We have simulated the sampling of our simulated sea surface by a highly simplified altimeter. We thus neglect radar noise, speckle and variations in ocean backscatter, and compute simulated waveforms as histograms of the number of discrete pixels as a function of range R discretized with the same resolution $\delta_R = 0.47$ m used in actual SWIM data. The histogram is computed for a finite region of size r_C by r_C centered at the nadir point. The value of $\widehat{H}_s$ for each simulated histogram is given by the least square fit to the theoretical waveform in by eq. (9) for R varying from $h - H_s$ to $h + H_s$. As detailed in Appendix A, even this simplified altimeter makes a much more complex measurement than a simple Gaussian smoothing of the H_r field. We briefly tested that more realistic waveforms and different fitting procedures yield some quantitative differences.

Taking the simulated sea surface from Fig. 2, we compare a map of simulated altimeter data $\widehat{H}_s$ in Fig. 3 with maps of local wave heights H_r, smoothed on different scales. As expected, the patterns of the envelope with radiuses larger than 2 km, those that persist in Fig. 3.f., match the large scales of the simulated altimeter data. From a quantitative point of view, the standard deviation of the simulated altimeter data, here 0.705 m, is of the same order as the standard deviation of actual SWIM measurements over the same SWIM box (0.88 m). We also note that this value is very close to that obtained for a filtering of the envelope between a scale $r = r_C/5$ and $r = r_C/4.5$. As $r = r_C/4.5$ gives the closest value, we define the effective altimeter radius r_a as $r_a = r_C/4.5$.

Looking at Fig. 3.a. it is clear that the map of $\widehat{H}_s$ contains much smaller features than the envelope smoothed with $r_a = r_C/4.5$. All of these are spurious amplification of envelope perturbations that happen to be at the right distance from nadir, around $r_C/2$, as explained in Appendix A. As a result, maxima of $\widehat{H}_s$ given by the altimeter are not located at the true wave height maxima but slightly displaced by a distance of the order of $r_C/2$. A striking example in Fig. 3 is the region of waves higher than 11 m around $x = 20$ km, $y = 19$ km. The altimeter gives a local minimum where the true wave height is maximum, and a round halo of maxima surrounding that point. Conversely a ring-shaped maximum in the envelope, such as around $x = 35$ km, $y = 45$ km, gives a lo-

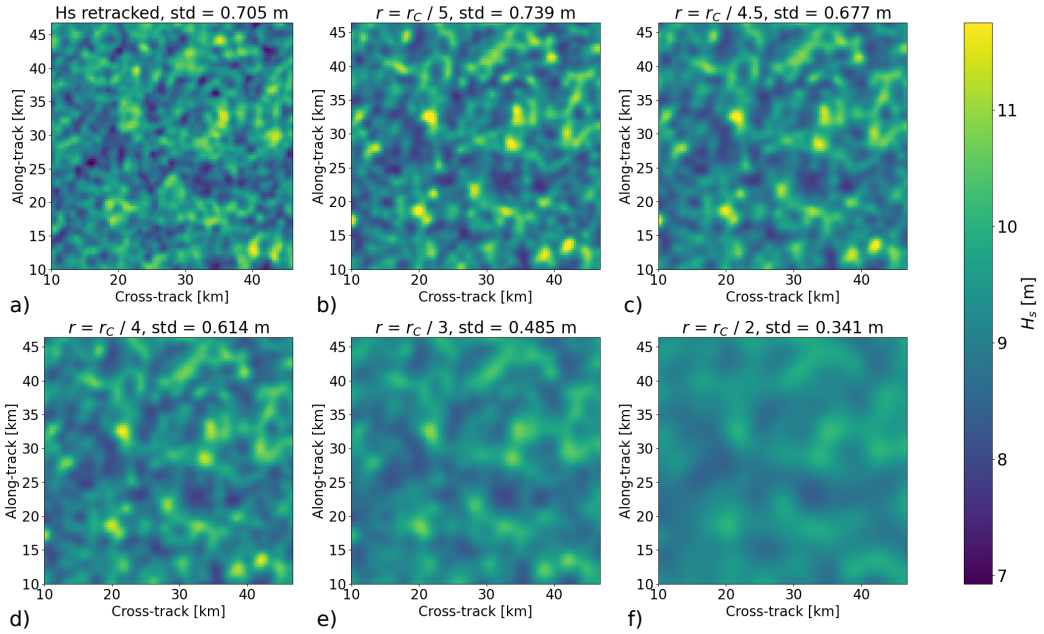

Figure 3. Maps of different estimates of the wave heights obtained by either (a) simulating altimeter processing or (b — f) computing the local average H_r using eq. (2) where the g_r kernel is a Gaussian filter of standard deviation r varied from $r_C/5$ to $r_C/2$. In this example, $r_C = 3063$ m thus r values are respectively 613, 681, 766, 1021, and 1531 m. In practice the smoothing is applied in a finite box of size $4r_C$ by $4r_C$.

cal maximum in the simulated altimeter data. We have found that different retracking methods produce the same large scale patterns but may differ in small scale details. These differences are beyond the scope of the present paper. Even though the patterns do not exactly coincide, we will now assume that the sampling of the sea surface by the altimeter is equivalent, in terms of variability of wave height to filtering the envelope with a Gaussian of standard deviation $r_a = r_C/4.5$.

2.5 Predicting H_r variability and its contribution of $\widehat{H_s}$ variability

Based on our analysis, we expect that SWIM measurements are contaminated by wave group structures at scales of a few kilometers, following the variation of r_C with wave height and satellite altitude. As illustrated by the two examples with different spectral widths, we note that for the same wave height, a wider spectrum leads to smaller scales of wave groups, part of which scales are smoothed away by the altimeter footprint and therefore not resolved. For a narrower spectrum, wave groups have larger scales and

amplitudes and a larger contribution to the variability of wave heights estimated by an altimeter.

For a quantitative analysis we first consider the simpler case of waves propagating in only one direction, with a sea surface ζ distributed with the normal law $\mathcal{N}(0, \sigma_H = H_s/4)$ with a single-sided (defined for $k > 0$) Gaussian spectrum with mean value k_p and standard deviation σ_k

$$E(k) = \frac{H_s^2}{16\sigma_k\sqrt{2\pi}} \mathrm{e}^{-(k-k_p)^2/(2\sigma_k^2)}. \tag{11}$$

The spectrum of the envelope is also Gaussian and writes, using eq. (3) and (4), and after computing the correlation (for a single-sided spectrum),

$$\begin{aligned} \Psi(k) &= \frac{8-2\pi}{H_s^2}\Psi_2(k), \\ &= \frac{8-2\pi}{H_s^2} 8 \int_0^\infty E(u)E(u+k)\mathrm{d}u, \\ &= \frac{16(4-\pi)}{H_s^2} \frac{H_s^4}{512\sqrt{\pi}\sigma_k} \mathrm{e}^{-k^2/(4\sigma_k^2)}, \\ &= \frac{(4-\pi)H_s^2}{32\sqrt{\pi}\sigma_k} \mathrm{e}^{-k^2/(4\sigma_k^2)}. \end{aligned} \tag{12}$$

Using eq. (6), with a Gaussian filter of standard deviation r_a, the spectrum of H_{r_a} writes,

$$\Psi_{H_{r_a}}(k) = \underbrace{\frac{(4-\pi)H_s^2}{\pi\sqrt{\pi}\sigma_k} \mathrm{e}^{-k^2/(4\sigma_k^2)}}_{\Psi_{H_0}(k)} G_{r_a}(k) \tag{13}$$

Wave groups contain wavelengths larger than π/σ_k, with a constant spectral density near $k = 0$. Around $k = 0$, the value of the H_{r_a} spectrum is $0.15H_s^2/\sigma_k$ m^2/(rad/m).

Fig. 4 presents one dimensional wave spectra — in solid lines — of two typical sea states with the same $H_s = 3.1$ m and their associated $\Psi_{H_0}(k)$ spectra — in dashed lines. The light blue spectrum is a JONSWAP spectrum (Hasselmann et al., 1973) with a peak period of 8 s and a peak enhancement factor $\gamma = 3.3$ that represents a moderate wind-sea. The dark blue spectrum is a narrow Gaussian spectrum with a peak period of 14 s and $\sigma_k = 0.002$ rad/m, typical of swell conditions in the open ocean. The altimeter smoothing function $G_{r_a} = \exp(-k^2 r_a^2)$ allows to define a cut-off wave number $k_a = \sqrt{\pi}/(2r_a)$. As shown in Fig. 4 the wavelengths in altimeter-filtered envelopes, larger than the associated wavelength cut-off $2\pi/k_a$ (in black dash-dotted line), are large compared to the typical wavelengths contained in the wave groups (of order $\pi\sqrt{2}/\sigma_k$ and represented by the dark and light blue vertical dash-dotted lines).

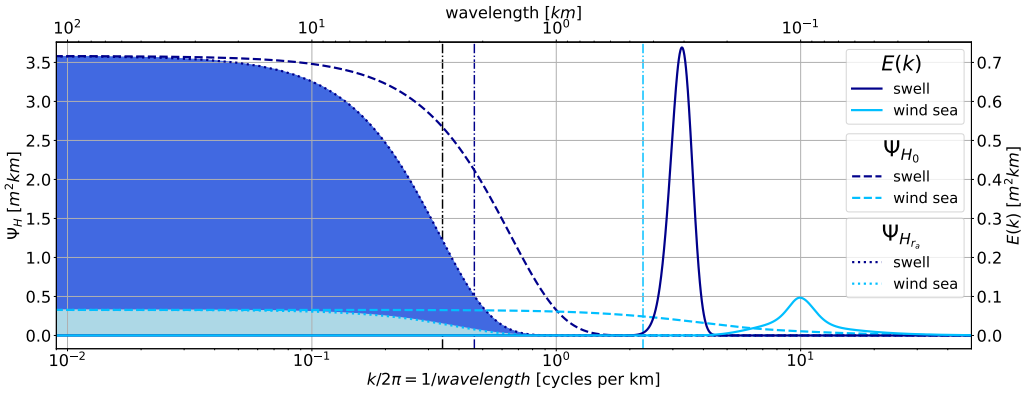

Figure 4. Example of two wave spectra — solid lines — in one dimension and the corresponding spectra of H_s — dashed lines —, for typical swell conditions in the open ocean in dark blue, and typical windsea in moderate wind conditions in light blue. Because the fluctuations of H_s are filtered by the altimeter with the function $G_{r_a}(k)$ — dotted lines —, the actual measured variance of H_s is the shaded area, in light blue for the windsea and dark blue for the swell. The vertical black line is the equivalent altimeter cut-off wavenumber at $k = k_a$, whereas the vertical dark and light blue lines represents the width of the H_s spectra.

Applying the one dimension version of eq. (8) gives the variance of altimeter-estimated $\widehat{H}_s$ as the shaded areas in Fig. 4. For a Gaussian spectrum, in cases where the altimeter filter scale is large enough not to be concerned about the shortest scales, this area is approximately k_a times $\Psi_{H_0}(k = 0)$ the PSD level at $k = 0$. This gives a standard deviation of $\widehat{H}_s$ of the order of $0.39\sqrt{k_a/\sigma_k}H_s$, which is 0.40 H_s for the one-dimensional swell example of Fig. 4.

For a generic one-dimensional wave spectrum $E(k)$, the reciprocal width $1/\sigma_k$ should be replaced by $(2\sqrt{\pi})Q_k$, with the peakedness parameter Q_k defined similarly to the reciprocal of the the usual frequency bandwidth (Saulnier et al., 2011),

$$Q_k = \frac{\left(\int_0^\infty E(k)\mathrm{d}k\right)^2}{\int_0^\infty E^2(k)\mathrm{d}k}. \tag{14}$$

For a JONSWAP spectrum with a peak enhancement factor $\gamma = 3$, this definition gives $B_k = 1.3k_p$ and the standard deviation of H_r for the wind sea case above is $0.1H_s$.

For a realistic altimeter, we must consider waves in two dimensions, and the standard deviation of H_{r_a} is the square root the variance as given by eq. (8). This variance is the integral of the H_{r_a} spectrum in the wavenumber plane. For large enough r_a, the

integral can approximated as the value $\Psi_{H_{r_a}}$ of the spectrum at $(k_x = 0, k_y = 0)$ times an effective area in the wavenumber plane,

$$\begin{aligned}
\mathrm{var}(H_{r_a}, d_1) &\simeq \Psi_{H_0}(k_x = 0, k_y = 0) \times \int_{k_y \in \mathbb{R}} \int_{k_x \in \mathbb{R}_{\setminus [-k_1, k_1]}} G_{r_a}(k_x, k_y) \mathrm{d}k_x \mathrm{d}k_y \\
&\simeq \Psi_{H_0}(k_x = 0, k_y = 0) \times (\pi/2) \left(2/{r_a}^2 - 4k_1/(\sqrt{\pi} r_a)\right) \\
&\simeq \frac{32}{\pi} \frac{8 - 2\pi}{H_s^2} \Psi_2(k_x = 0, k_y = 0) \times (\pi/2) \left(2/{r_a}^2 - 4k_1/(\sqrt{\pi} r_a)\right) \\
&\simeq Q_{kk}^2 H_s^2 (4 - \pi) \left(2/{r_a}^2 - 4k_1/(\sqrt{\pi} r_a)\right), \qquad (15)
\end{aligned}$$

where we have defined a two-dimensional spectral peakedness Q_{kk} which is measured in meters,

$$Q_{kk}^2 = \frac{\iint_{\mathbb{R}^2} E^2(k_x, k_y) \mathrm{d}k_x \mathrm{d}k_y}{\left(\iint_{\mathbb{R}^2} E(k_x, k_y) \mathrm{d}k_x \mathrm{d}k_y\right)^2} = \frac{32 \Psi_2(k_x = 0, k_y = 0)}{H_s^4}. \qquad (16)$$

This expression gives the approximate value for the standard deviation,

$$\mathrm{std}(H_{r_a}, d_1) \simeq H_s Q_{kk} \sqrt{(4 - \pi) \left[2/r_a^2 - 4k_1/(\sqrt{\pi} r_a)\right]}. \qquad (17)$$

This variability of H_{r_a}, which we have defined as the contribution of wave groups to the variability of measured wave heights $\widehat{H}_s$ is thus the product of three factors: the significant wave height H_s, the shape of the wave spectrum as quantified by Q_{kk}, and the effective range of spatial scales over which the variance is integrated. That last factor is a function of the smoothing effect of the altimeter, represented by the scale r_a, and the distance $d_1 = 2\pi/k_1$ over which we consider the variability.

2.6 Practical implementation

In the previous subsections, various considerations have been made. Here, we summarize them to give a flow chart for estimating the part of the variability due to wave groups $\mathrm{std}(H_{r_a}, d_1)$ from CFOSAT products.

1. Assemble the 1-dimensional L2S spectra for each azimuth to obtain an equivalent L2 spectrum, make it double-sided $E_{2S}(k, \theta)$ and rescale it with $\widehat{H}_s$ from nadir measurement,
2. Interpolate $E_{2S}(k, \theta)$ over a regular (k_x, k_y) grid, with k_x the along track direction, to obtain $E(k_x, k_y)$,

3. Compute the spectrum of the envelope squared $\Psi_2(k_x, k_y)$ from the discrete correlation

$$\Psi_2(k_x, k_y) = 8 \sum_{k'_x} \sum_{k'_y} \left[E(k'_x, k'_y), E(k'_x + k_x, k'_y + k_y) \right] \mathrm{d}k'_x \mathrm{d}k'_y \qquad \text{(Step 3)}$$

4. Transform to a PSD of H_0,

$$\Psi_{H_0}(k_x, k_y) = \frac{32}{\pi} \frac{8 - 2\pi}{H_s^2} \Psi_2(k_x, k_y) \qquad \text{(Step 4)}$$

where H_s is computed as $4\sqrt{E}$, with $E = \iint_{\mathbb{R}^2} E(k_x, k_y) \mathrm{d}k_x \mathrm{d}k_y$.

5. Compute the altimeter smoothing filter G_{r_a} as

$$G_{r_a}(k_x, k_y) = |\mathcal{F}(g_{r_a})|^2 = \mathrm{e}^{-(k_x^2 + k_y^2) r_a^2} \qquad \text{(Step 5)}$$

with $\mathcal{F}(\cdot)$ the Fourier transform and $r_a = r_C / 4.5$.

6. Apply the filter to Ψ_{H_0} to obtain the PSD of altimeter H_{r_a} estimate,

$$\Psi_{H_{r_a}}(k_x, k_y) = \Psi_{H_0}(k_x, k_y) G_{r_a}(k_x, k_y) \qquad \text{(Step 6)}$$

7. Integrate $\Psi_{H_{r_a}}(k_x, k_y)$, the spectrum of H_{r_a}, over all cross-track wavenumbers k_y and over along-track wavenumbers k_x with a magnitude larger than $2\pi/d_1$, as illustrated on Fig. 5.e. Because we compare our estimate to $\widehat{H}_s$ variations within a SWIM L2 box size that is 80 km along-track, d_1 is taken as 80 km,

$$\mathrm{var}(H_{r_a}, d_1) = \int_{k_y \in \mathbb{R}} \int_{k_x \in \mathbb{R}_{\setminus [-k1, k1]}} \Psi_{H_{r_a}}(k_x, k_y) \mathrm{d}k_x \mathrm{d}k_y \qquad \text{(Step 7)}$$

Fig. 5 shows the results of steps 3, 6 and 7 for the two selected boxes of our case study. The top line corresponds to the Ψ_2 spectra obtained from correlation, the middle line shows $\Psi_{H_{r_a}}$, the PSD of our local H_{r_a} estimate. Note the different colour scales between the northern — A — and southern — B — boxes. The bottom line shows the one sided along-track k_x-spectra. The vertical line shows the lower integration limit over k_x that is used to obtain the variance of H_{r_a}.

Fig. 6 shows wave height and corresponding standard deviation, both observed and estimated from L2S spectrum, as a function of the sampling time (UTC) over storm Dennis. For the northern part of the storm, where the spectrum is broader, around box A (light blue vertical line), the standard deviation due to wave groups is around half the measured standard deviation (i.e. wave groups represent a quarter of the measured variance). On the other hand, for the southern part, around box B (dark blue), the wave

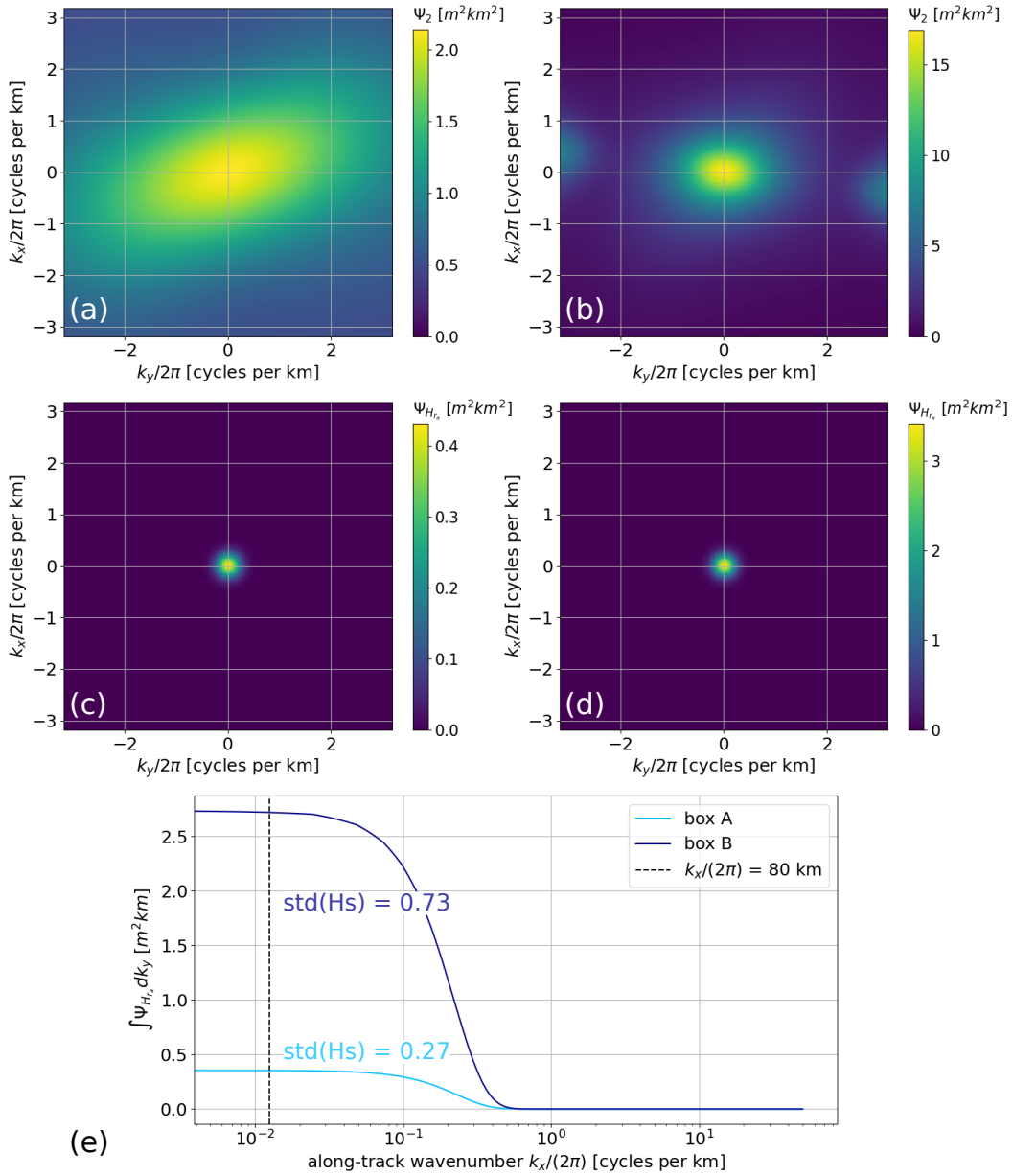

Figure 5. (a,c) corresponds to our chosen northern CFOSAT box, and (b,d) to the southern box. Top line: envelope squared spectrum $\Psi_2(k_x, k_y)$ from convolution. Middle line: $\Psi_{H_{r_a}}(k_x, k_y)$, spectrum of H_{r_a} (including the equivalent altimeter filtering). Bottom line: 1D along-track spectrum obtained by integrating over the cross-track axis, in light blue for the northern box and dark blue for the southern box.

height variability is strongly dominated by wave groups — more than half the observed variance is explained by wave groups.

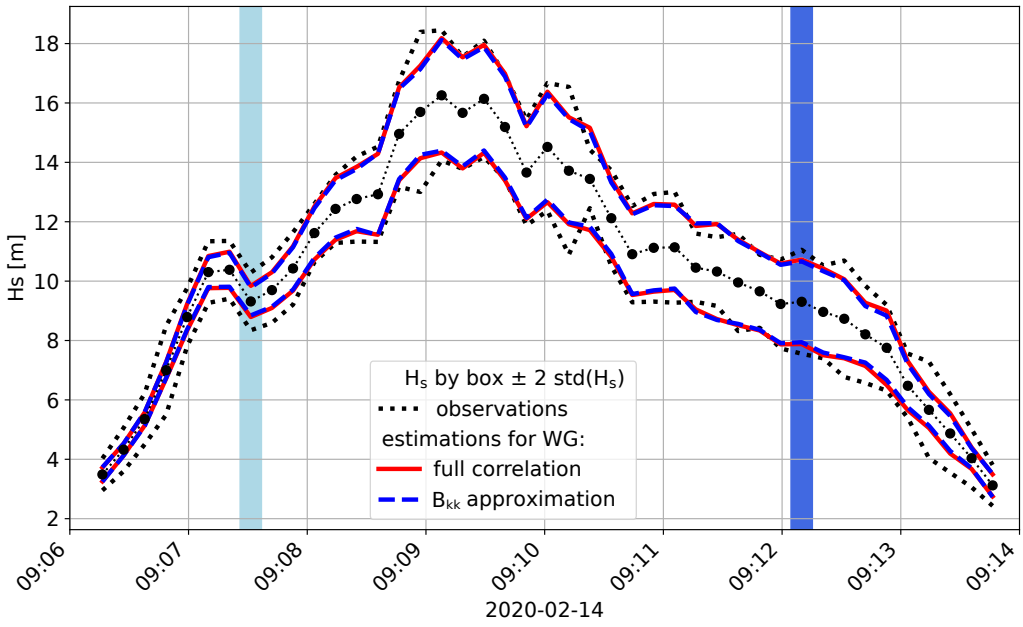

Figure 6. Values of measured $\widehat{H}_s$, averaged over boxes — black circles — , and corresponding std($\widehat{H}_s$, 80km) — black dotted lines — ; as provided in the L2 as a function of sampling time (UTC), for the CFOSAT track shown in Fig. 1. Estimations of std(H_{r_a}, 80km) are also represented — in red and blue — using the two methods summarized at the end of section 2.

Alternatively we can approximate the full integral of the convolution by its value at $k = 0$ using the peakedness Q_{kk}, giving a faster estimate of the variability due to wave groups. The first two steps are the same, then

- Instead of the full convolution, compute only Q_{kk} defined from eq. (16), which can be re-written as,

$$Q_{kk} = \frac{\sqrt{\iint_{\mathbb{R}^2} E^2(k_x, k_y) \mathrm{d}k_x \mathrm{d}k_y}}{\iint_{\mathbb{R}^2} E(k_x, k_y) \mathrm{d}k_x \mathrm{d}k_y} \qquad \text{(Step 3bis)}$$

- skip steps 4–6 to estimate std(H_{r_a}, d_1) directly using eq. (17),

$$\mathrm{std}(H_{r_a}, d_1) \simeq Q_{kk} H_s \sqrt{(4-\pi)\left[2/r_a^2 - 4k_1/(\sqrt{\pi} r_a)\right]}. \qquad \text{(Step 7bis)}$$

This std(H_{r_a}, d_1) estimated through Q_{kk} is also shown in Fig. 6. The values are only slightly overestimated compared to the full correlation calculus, therefore, Q_{kk} could be a useful parameter when working with wave groups.

3 Results at the global scale

Beyond the particular case of storm Dennis, for which very large wavelength and narrow spectra lead to a dominant effect of wave groups in $\widehat{H}_s$ variability, one may wonder how important are wave groups in general, and how important can they be compared to other known sources of $\widehat{H}_s$ variability, including winds and currents. To answer this

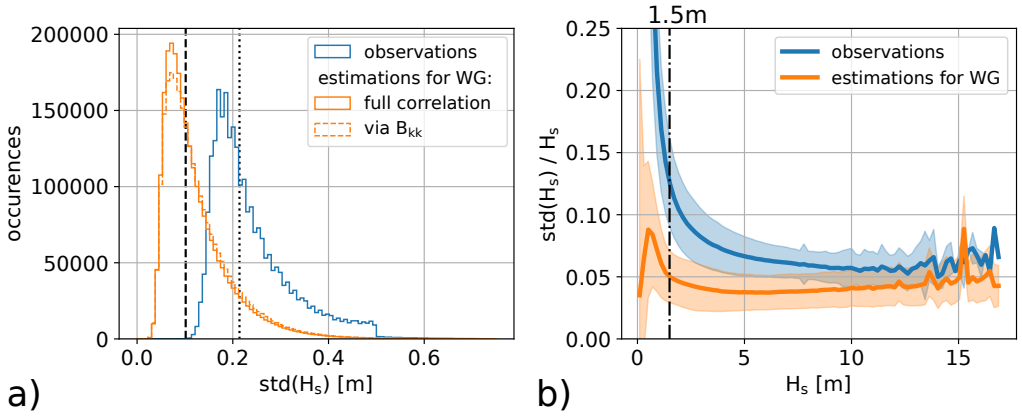

Figure 7. (a) Histograms of std($\widehat{H}_s$, 80km) measured at nadir in blue and our estimate of wave groups contribution std(H_{r_a}, 80km) in orange. (b) Mean — solid lines — and standard deviation — shaded areas — of std(H)/$\widehat{H}_s$ over $\widehat{H}_s$ bins of 0.2 m, both for std($\widehat{H}_s$, 80km) measured at nadir in blue and our estimate of wave group effects std(H_{r_a}, 80km) in orange.

question, we have applied the methodology presented in section 2 for storm Dennis to the full SWIM L2S archive for the years 2020 and 2021, estimating for each of the 2.4 million SWIM L2 box the expected value of std(H_{r_a}, d_1) associated to wave groups as filtered by the altimeter. The distribution of these values is shown in Fig. 7.a, with a typical value around 7 cm, and maximum values around 60 cm. This variability is typically half of the measured standard deviation of $\widehat{H}_s$. We also computed std(H_{r_a}, d_1) as estimated from the spectral peakedness parameter Q_{kk} for the same time period, giving results that are highly correlated to the full convolution, with a Pearson's linear correlation $R = 0.98$.

In practice the group-induced variability of H_{r_a} that that should be present in SWIM data is strongly correlated with the mean value of H_s. In Fig. 7.b we show the statistical distribution (mean and standard deviation) of std(H_{r_a})/mean($\widehat{H}_s$) as a function of H_s. For H_s below 1.5 m the altimeter estimates of H_s are known to have the largest

relative errors (Dodet et al., 2022), which is partially due to a stronger effect of speckle noise, as discussed in Appendix A.3. In that range the wave group variability is 3 times smaller than the observed variability. Above 1.5 m, the relative variability that is expected from wave groups increases with H_s (from 3.5% to 5%), whereas the observed variability decreases from 11% to 5% between $H_s = 1.5$ m and $H_s = 10$ m. The share of the variability expected from wave groups dominates the observed variability for wave heights above 8 m, and probably explains the increase in observed std(H_s)/mean(H_s) for H_s above 12 m. Although there are very few data in that range, it is well known that dominant wave periods are generally higher for higher wave heights (Toba, 1973), corresponding to smaller bandwidths and thus an ever increasing variability due to wave groups.

We now consider the spatial distribution of std($\widehat{H}_s$), and in order to separate the possible effects of different sea states from the general trends associated to local average value of H_s, we have chosen to show a map of the mean value of std($\widehat{H}_s$, 80 km)/mean($\widehat{H}_s$, 80 km) gridded at a resolution of 100 km. Before computing the local mean we have first removed all cases with $\widehat{H}_s < 1.5$ m. Fig. 8.a shows the distribution of $\widehat{H}_s$ variability and compare it to the predicted variability of H_{r_a} in Fig. 8.b.

Note that the range of values are different for both panels because the contribution of wave groups is, on average, half of the measured std($\widehat{H}_s$, 80 km). Both figures have some common patterns with a general increase from the west to the east of the ocean basins consistent with a dominance of swells in the east (Chen et al., 2002) with longer wavelengths and narrower spectra.

Now that we have quantified the variance of $\widehat{H}_s$ associated to wave groups, we can subtract this contribution of wave groups from the total variance of $\widehat{H}_s$ in order to look at the other sources of variability in $\widehat{H}_s$. As shown in Fig. 8.c, the remaining standard deviation of $\widehat{H}_s$ after correction of the effects of wave groups (total variance minus variance due to wave groups) contains a background level of 0.1 to 0.2 m, possibly associated to known artificial effects that include the automatic gain control, speckle noise (Quartly et al., 2001) and true small scale wind variability. Larger localized values are up to 0.3 m. These larger values are co-located with regions of strong ocean circulation mesoscale variability. These same regions match the location of strong H_s gradient in along-track 1 Hz data from SARAL-AltiKa, Jason-2 and Cryosat-2 that have been denoised using an Empirical Mode Decomposition (EMD), by Quilfen and Chapron (2019). Here, we have

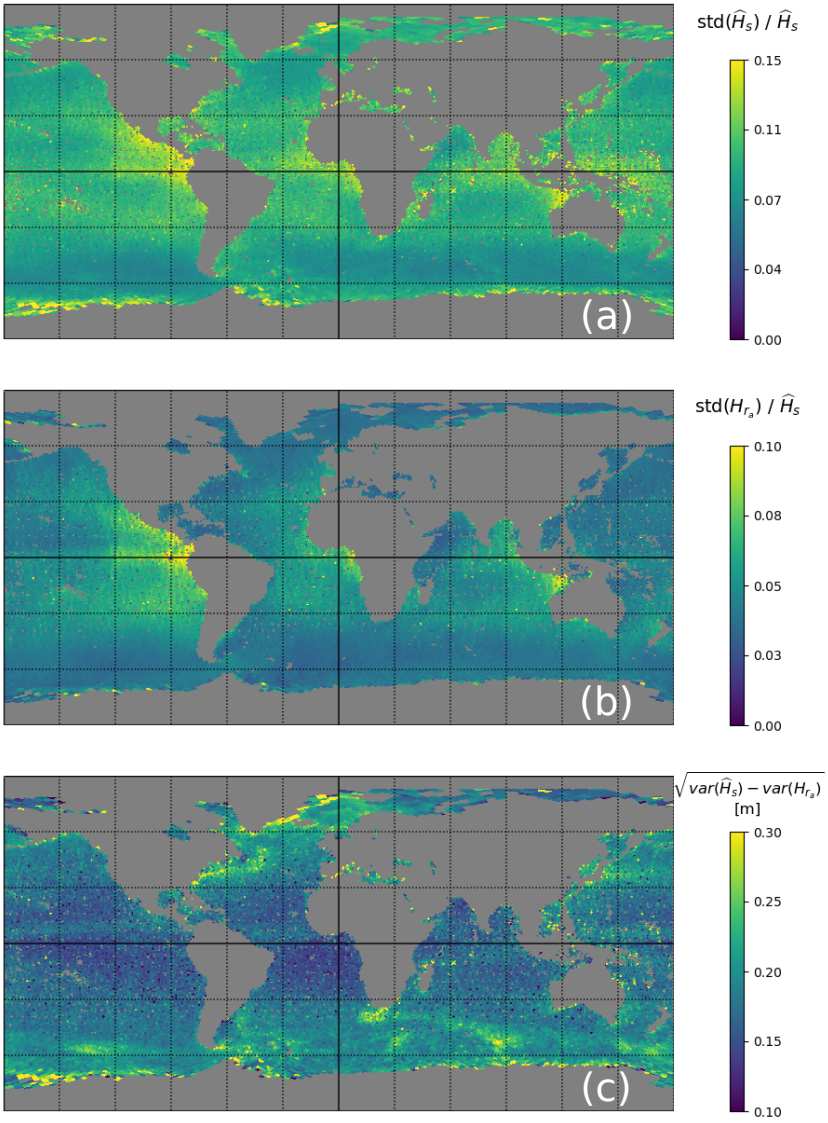

Figure 8. Map of the average of a) std($\widehat{H}_s$, 80 km)/mean($\widehat{H}_s$, 80 km) — upper panel —, b) std(H_{r_a}, 80 km)/mean($\widehat{H}_s$, 80 km) — middle panel — and c) residual standard deviation of $\widehat{H}_s$, in meters, after removing the effect expected from wave groups — lower panel —, for the years 2020 and 2021 for all the SWIM L2boxes with a $\widehat{H}_s$ above 1.5 m. With the wave group contribution std(H_{r_a}, 80km) estimated from SWIM L2S spectra

applied the same EMD filtering to SWIM 4.5 Hz data, in order to remove small scale noise in $\widehat{H}_s$, giving results shown in Fig. 9. The EMD filtered part Fig. 9.c is directly comparable to the wave group signature highlighted in Fig. 8. These maps were constructed using SWIM nadir data from SALP/CAWATAC experimental 4.5 Hz products available

541 on Aviso+ that include both the raw significant wave height estimates $\widehat{H}_s$ and the de-
noised values $\widetilde{H}_s$ using the EMD method. The magnitude and distribution of the expected

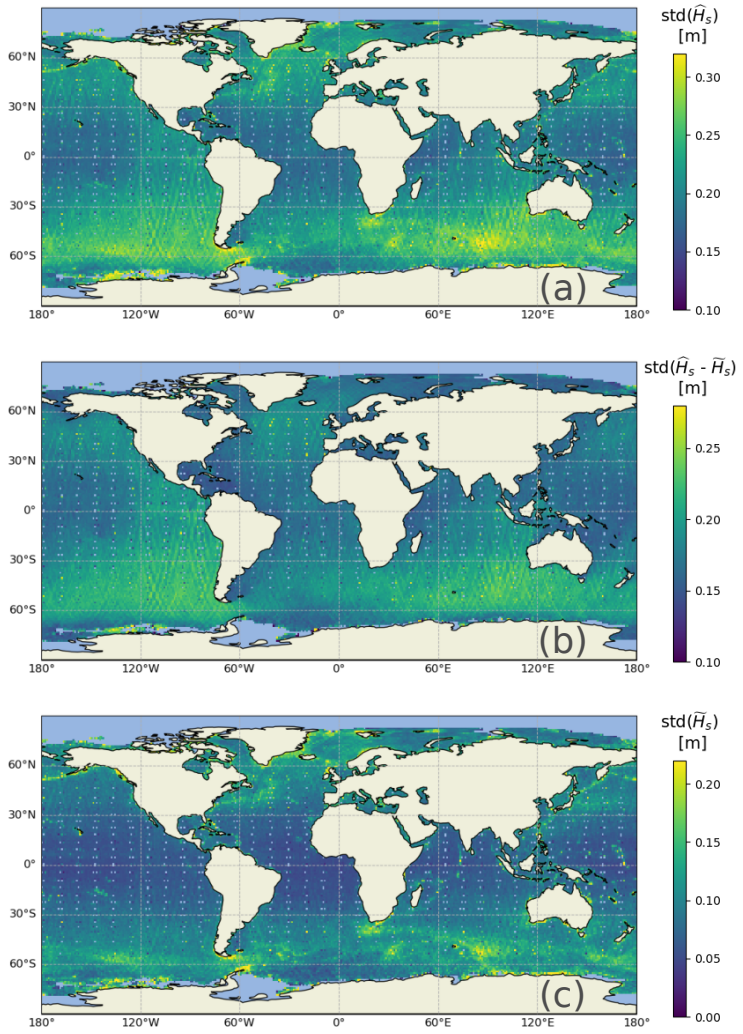

Figure 9. Maps of the average, for the year 2021, of (a) std($\widehat{H}_s$, 80km) computed on original SWIM nadir native values (4.5 Hz), (b) standard deviation of the residual of H_s, defined as the difference between the original and the denoised wave height (c) std($\widetilde{H}_s$, 80km) computed on EMD filtered data $\widetilde{H}_s$.

542

543 effect of wave groups apparently corresponds to the variability that is removed by EMD
544 denoising, without using wave spectrum information. Thus, the uncertainty variable as-
545 sociated to the nadir 4.5 Hz data in the SALP/CAWATAC products which is derived
546 from the standard deviation of the fluctuations removed by EMD should be related to
547 H_s and Q_{kk}. This estimation of the uncertainty may be useful for extrapolating uncer-

tainty estimates based on triple-collocation methods to high values of H_s. For satellite missions other than CFOSAT, the EMD filtered data are available but we do not have measured wave spectra from which Q_{kk} could be estimated to verify this interpretation of the EMD filtering. The analysis of these other missions may use estimates of Q_{kk} from numerical wave models and/or co-location of data with SWIM.

4 Discussion

4.1 Effect of spectral shape

The accurate estimation of wave group contributions critically depends on the accuracy of the spectral shape, in particular the directional width and wavenumber width. Because of the hard wavelength cut-off in the L2 product we have chosen to work with the L2S spectra. Redoing the global analysis with the L2 product generally reduces the expected effect of wave groups. We note that a validation of spectral width from the L2 product was performed by Le Merle et al. (2021), who found that SWIM L2 generally overestimate spectral width compared to buoy data. No such analysis has been performed for the L2S product. It would be also interesting to know how accurate could be the estimation of std(H_{r_a}, d) estimated from model spectra, for the application to other satellite mission that do not measure the wave directional spectrum. The minimum distance d that could be investigated will depend on the wave model resolution.

It should also be reminded that SWIM L2 spectra combine sparse measurements over a 70 km by 90 km box, as illustrated in Fig. 10. Because the wave field has gradients, this combination generally produces a broader spectrum than a more local estimate of the wave spectrum, and this should produce a low bias in our estimate the effect of wave groups. In the example on Fig. 10 the assembly of the L2 spectrum combines data from two neighboring wave azimuths that are observed in distinct regions of the ocean separated by up to 72 km. We kept this assembly to be able to compare results from L2 and L2S data, but it would be more logical, for the case of L2S data, to assemble a spectrum with a spatial continuity of the footprints that correspond to azimuths around the spectral peaks.

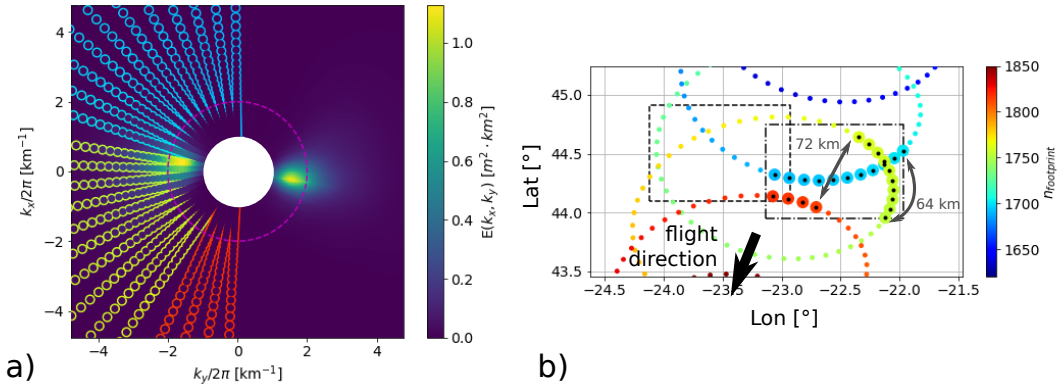

Figure 10. (a) Representation of a two-dimensional spectrum $E(k_x, k_y)$ with values given by the colorbar obtained from assembling L2 1-dimensional spectra $E(k)$ for all azimuths. The color of each circle corresponds to the index of the beam footprint in which the SWIM was making the measurement (b) Geographical layout of boxes — dashed and dash-dotted rectangles — and of centers of the beam footprints on the sea surface. the footprint diameter is about 20 km so that they actually overlap.

4.2 Effects of satellite altitude

The particularly low orbit of CFOSAT at 519 km gives a rather small oceanic footprint that allows for wave groups to be resolved in SWIM data. If we consider the higher altitudes used by other satellite missions, 891 km for the recently launched SWOT or 1340 km for the Topex-Poseidon / Jason / Sentinel 6 series, the oceanic footprint gets bigger and wave groups are more likely to be more smoothed out. In Fig. 11, we illustrate this effect with different altitudes following the method used for Fig. 3. Namely, for the same simulated ocean surface, the wave height is estimated by a least-square fit to the simplest Brown waveform given by eq. (9). As expected, the higher the satellite the lower the variability of $\widehat{H}_s$. It is not clear that this effect of satellite altitude is noticeable in real data that are contaminated by speckle noise and that use different waveform fitting algorithms. More realistic simulations will be needed to compare the behaviour of different instruments and processing chains.

4.3 Expected effect on delay doppler altimeters

We have shown that the variability of $\widehat{H}_s$ at small scale contains some geophysical information and not just random noise related to the measurement. However, the noise

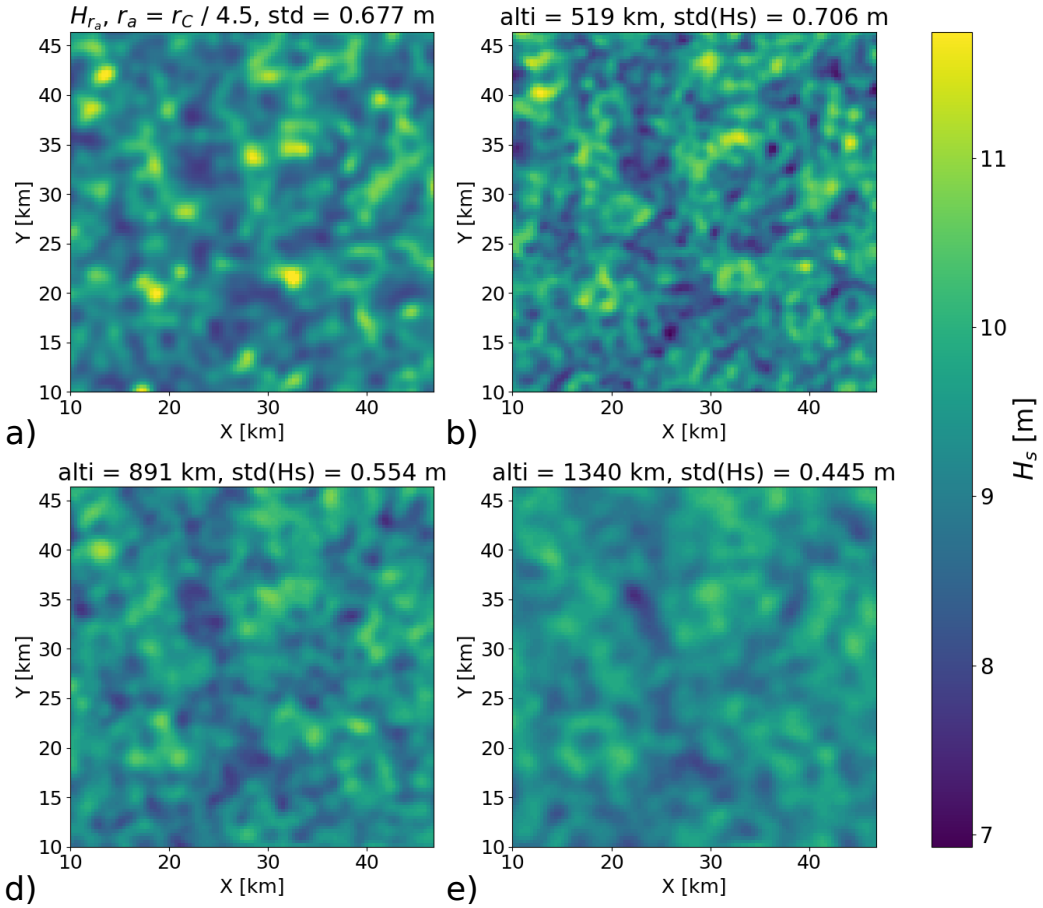

Figure 11. Maps of wave heights obtained a) by smoothing the envelope with Gaussian filter of scale r_a = 619 m; or by simulating altimeter waveforms without speckle and using the same least-squares fit (Appendix A), for different altitudes: b) 519 km, c) 891 km, d) 1,340 km, with corresponding r_a of 619 m, 790 m and 940 m.

for Delay-only altimeters is probably dominated by the speckle noise in the waveforms (Sandwell & Smith, 2005; Quartly et al., 2019). Doppler processing of recent altimeter instruments starting with Cryosat-2 and Sentinel-3 can strongly reduce this speckle noise by forming and combining independent looks of the same sea surface (Egido et al., 2021). It will therefore be interesting to study the effect of wave groups in these measurements of wave height and sea level. Waves can also be resolved directly in the sea level estimates when data is processed at very high resolution (Altiparmaki et al., 2022; Villas Bôas et al., 2022). If the Doppler induced by orbital velocities is neglected, the delay-Doppler measurement is similarly based on the convolution of a surface elevation distribution with a flat surface response (Ray et al., 2015). Only the flat surface response is different from

the Delay-only processing. We thus expect that wave groups will have similar distortion of the waveforms and contributions to estimates of wave heights and sea level. The blurring effect caused by range bunching will now be confined to the direction perpendicular to the track, with maximum effect of a H_s perturbation located off the satellite track (depending on fitting algorithm), possibly also at a distance of the order of $r_C/2$. Because Delay-Doppler altimeters can actually resolve the along-track variability caused by wave groups instead of averaging it, we expect that $\widehat{H}_s$ fluctuations caused by wave groups are much larger in Delay-Doppler altimetry, together with their spurious effect on sea level estimates. This may explain the relative smaller reduction of std($\widehat{H}_s$, 7km) at large H_s which is found when Doppler resolution is enhanced to reduce the speckle effect, and a typical value of std($\widehat{H}_s$, 7km) for Delay-Doppler Sentinel 3A data which that is around 0.7 m for $H_s = 7$ m (Egido et al., 2021), twice the typical value for SWIM data. This will be the topic of further studies.

4.4 Wave groups and satellite measurements uncertainties

Up to now, the uncertainty of satellite measurements has been determined by the triple-collocation method (Abdalla et al., 2011; Dodet et al., 2022), with the practical result that the uncertainty of altimeter data, either denoised or integrated along-track into super-observations, is of the order of 7% of H_s. However, that error contains representation errors (the co-located in situ data does not sample the same space and time frame), and cannot be extrapolated beyond the range of the co-located dataset, typically wave heights below 8 m. So what can we say about the largest measured wave heights of 20.1 m (Hanafin et al., 2012)? Can we use the measured variability of $\widehat{H}_s$, for example the 4.5 Hz or 20 Hz data that is used to make a 1 Hz average, to refine our estimate of the uncertainty of this average? In the present paper we have shown that wave groups are responsible for random fluctuations in the estimates $\widehat{H}_s$, that are generally proportional to H_s but with an effect that depends on the bandwidth of the spectrum, which is generally narrower for larger wave periods. As a result the variability associated to wave groups can be the dominant source of fluctuations in $\widehat{H}_s$ measurements for severe storm conditions. Even though the measurement fluctuations are weakly correlated to the actual wave height variations (as demonstrated in Fig. 3) their magnitudes are strongly correlated. Hence the measured fluctuation std($\widehat{H}_s$, 7km) contains both uncorrelated speckle noise effect, that can be expected to be reduced by $1/\sqrt{N}$ when averaged from N Hz to

1 Hz, and a true geophysical spatial variability associated to wave groups (and variable fetch, currents, etc.) that will only partially average out. We expect that an uncertainty model for averages of $\widehat{H}_s$ measurements may take into account wave groups explicitly. In the case of SWIM, directional wave spectra can be used to separate the actual variability of the 4.5 Hz data into wave group effects and noise plus other geophysical effects. For other altimeters, one may use empirical correlations between spectral bandwidth, wave height and wind speed. For this information to be useful for a theoretically-based uncertainty estimate, which is much needed for wave heights above 8 m, one may extend the parameterization of speckle effects proposed in Appendix A.3, to the actual target waveform and cost function used in the retracking algorithm.

4.5 Considerations on the satellite resolution

The present work should be useful for the exploration of the resolution limits of satellite altimeters and other remote sensing systems that use radar or optical imagery (Kudryavtsev et al., 2017). As processing methods are refined to produce higher resolution near the coast (Passaro et al., 2021) and the ice edge (Collard et al., 2022), some of the high resolution data will be dominated by wave groups. The associated variance of H_s may provide some constraint on the shape of the directional wave spectrum, but the detailed fluctuations are probably of little value for most applications as groups will travel at speeds of the order of 10 m/s and persist for only a few minutes. The contribution of wave groups to the variability of wave heights measured by altimeters is thus a real effect that contains part of the true variability of wave heights at the scale of the altimeter footprint. Methods developed to remove noise in the data, such as the data-driven Empirical Mode Decomposition (EMD) used by (Quilfen & Chapron, 2019) appear to remove the effect of wave groups. An investigation of the variabilty of wave heights at the smallest scales cannot be based on denoised data alone, because they miss a large part of the true variability.

In locations where H_s varies sharply such as over coral reefs, mud banks or across the sea ice edge, the high resolution wave heights will contain other effects, and these are particularly interesting. Some caution should be used when interpreting these sharp gradients. As we have found out, the maximum wave height will generally be displaced from the location of its true maximum. This displacement is smallest for the SWIM instrument, thanks to the low orbit of CFOSAT, which makes it a particularly interest-

ing instrument for studying small scale wave height variations, in spite of its rather low rate for the nadir beam (4.5 Hz instead of 20 Hz for Jason), and the absence of Doppler processing.

5 Conclusion

In this paper, we took advantage of the low orbit altitude of CFOSAT, and the low noise level of the nadir beam of the SWIM instrument to study the along-track variability of wave height. The directional wave spectra measured by off-nadir beams on SWIM has been complementary to study the relationship between wave spectra and along-track H_s variability. After giving a theoretical estimate of the standard deviation of H_s associated to wave groups as a function of the wave spectrum and satellite altitude, we computed this estimate for 2 years of CFOSAT data using L2S products. We found that the standard deviation of H_s associated to wave groups is generally about half of the standard deviation of H_s measured over a 80 km distance, explaining 25% of measured H_s variance. This ratio of variances increases for storms and in the presence of long swells, it was found to be larger than 75% in 3% of the cases.

The residual variability after subtracting the estimated effect of wave groups from the measured H_s variance correlates positively with along-track variance of filtered signals, which brings out regions of strong currents. Studying these features is of major interest and requires some filtering out of the smaller scales. In turn, these smaller scales can be of interest, at least from a statistical point of view, as they are related to wave groups and the generation of infragravity waves and extreme sea level at the coast.

The main novelty of the present work was to expose the limitations of the theoretical Brown (1977) model that is used to retrieve wave heights and sea levels from altimeter data. The Brown model assumes a Gaussian sea level distribution that is spatially-uniform at the scale of the footprint, which is correct when averaged over long enough scales along-track. For an individual measurement, the footprint may not be large enough for the Brown model to be valid, and we have demonstrated that the effect of wave groups on the waveforms, is equivalent to introducing a range-dependent wave height. This limitation is common to all altimeters that provide estimates of wave height and sea level based on theoretical waveforms, including Low Rate Mode (LRM) and Synthetic Aperture Radar Mode (SARM) processing. The resulting waveforms have distorted shapes

that can be similar to the effect of non-Gaussian sea level statistics (Srokosz, 1986; Rodriguez, 1988), but with much larger distortions. To our knowledge this non-uniformity effect had never been considered. We have shown that the wave group effect averages out to zero over large scales when waveforms are fitted with a simple least square cost function, but that is not necessarily the case for the more sophisticated methods that are often used. We expect that further work will expand on our approach to consider spurious effects on sea level estimates as well as applications to recent delay-Doppler altimeters.

Appendix A Non-homogeneous H_s and waveform retracking

In this analysis we keep the most simple model of altimeter measurement that is also used in section 2: we neglect antenna pattern, thermal noise and mispointing effects, and neglect the Earth sphericity. These assumptions are meant to simplify the algebra as much as possible while keeping the essential features of non-homogeneity in wave heights. Likewise we have used the most simple cost function when fitting the waveform, while maximum likelihood methods are generally used with real data (Rodriguez, 1988; Halimi, 2013). We also start by ignoring speckle noise. The analysis performed below is easily extended to consider the third parameter which is usually estimated in retracking wave forms, that is the Normalized Radar Cross Section.

A1 Wave groups and H_s estimate

We consider a small perturbation Δ_H of H_s over an area A, localized around a range $h+R_0$. The original normalized Brown waveform of eq. (9) corresponds to the histogram of the ocean area per unit range, divided by $2\pi h$ so that it varies between 0 and 1, with h the satellite altitude. The perturbation to the waveform is equivalent to removing the original Gaussian distribution of surface elevation with $\sigma_H = H_s/4$ over the area A, and replacing it by a new Gaussian with $\sigma' = (H_s+\Delta_H)/4$, over the same area A, and divide by the normalization factor $2\pi h$. We define the parameter $a = A\Delta_H/(8\pi h)$, which should be small compared to H_s^2. For a small change in H_s, this change in waveform is proportional to the derivative of the Gaussian distribution with respect to σ_H and we find that the waveform is now

$$w(R) = w_B(R,\sigma_H) + a\frac{\mathrm{e}^{-(R-h-R_0)^2/(2\sigma_H^2)}}{\sqrt{2\pi}}\frac{(R-h-R_0)^2-\sigma_H^2}{\sigma_H^4} + O(a^2) \qquad \text{(A1)}$$

We note that a smaller change Δ_H over a larger area A changes the waveform in the same way as a larger change over a smaller area, provided that a is the same. For simplicity we redefine the Chelton footprint diameter as $r'_C = \sqrt{2H_sh}$, and we find that taking an area of radius $\alpha r'_C$ gives $a = \alpha^2\Delta_H H_s/4$.

The shape of these simulated distorted waveforms is illustrated in Fig. A1.

With the distortion shown here, fitting a Brown waveform would give a wave height of $H_{s,\mathrm{fit}} = 12.6$ m for $R_0 = 2.5$ m and $H_{s,\mathrm{fit}} = 10$ m for $R_0 = 0$, which is a strange way to average the $H_s = 13$ m over part of the footprint and 10 m in the rest of the

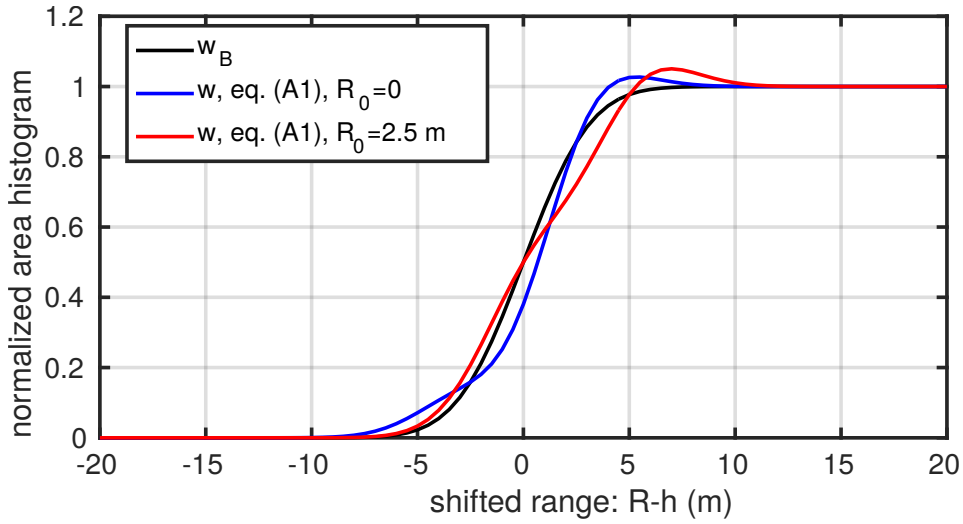

Figure A1. Example simulated waveforms in the presence of a localized change in H_s around the range $h + R_0$, for $H_s = 10$ m. The perturbations use $a = 1.875$ m^2, that would correspond to Δ_H=3 m over an area of radius $r'_C/4$, a perturbation that is neither small nor localized.

footprint. Fig. A2 shows that such perturbations are of the order of the deviations from the mean waveform in the case of the Box B SWIM waveforms, and are absent in Box A. The main difference between the simulated waveforms and the true waveforms is the speckle noise that is of the order of 10 % for both box A and box B.

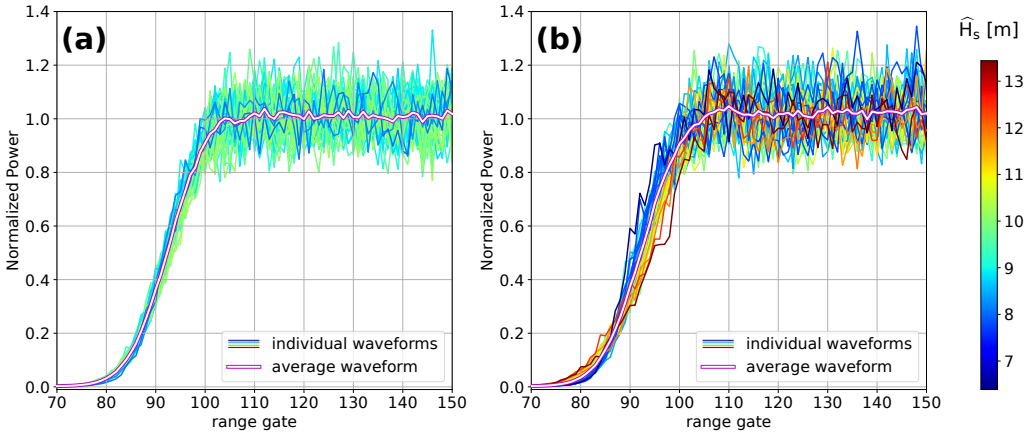

Figure A2. Ensembles of CFOSAT/SWIM waveforms in (a) box A and (b) box B. These are the L1A product, already corrected for the antenna pattern, and thus directly comparable to Fig. A1. Individual waveforms are color-coded with the estimated wave height. The white line represents the average waveform.

For small values of the perturbation a, the deviation in the fitted H_s can be computed analytically. For simplicity we will assume that the waveforms are defined for $-\infty < R < \infty$, and the sum of the difference squared between $w(R)$ and $w_B(R, \sigma'_H)$, when integrated from $R = -\infty$ to $R = \infty$ is the following cost function,

$$\begin{aligned} C &= \int_{-\infty}^{\infty} \left\{ [w_B(R,\sigma_H) - w_B(R,\sigma'_H)] + [w(R) - w_B(R,\sigma_H)] \right\}^2 \mathrm{d}R \\ &\simeq \int_{-\infty}^{\infty} \left\{ (\sigma_H - \sigma'_H) \frac{\partial w_B(R,\sigma_H)}{\partial \sigma_H} + [w(R) - w_B(R,\sigma)] \right\}^2 \mathrm{d}R \\ &= (\sigma'_H - \sigma_H)^2 \frac{1}{4\sqrt{\pi}\sigma_H} + (\sigma'_H - \sigma_H) \frac{aR_0}{8\sqrt{\pi}\sigma_H^5} \mathrm{e}^{-R_0^2/(4\sigma_H^2)} (R_0^2 - 6\sigma_H^2) + \frac{3a^2}{8\sqrt{\pi}\sigma_H^3}. \end{aligned}$$

Fitting σ'_H corresponds to solving $\partial C / \partial(\sigma'_H - \sigma_H) = 0$. We note that error terms that are either not a function of $(\sigma'_H - \sigma_H)$ or odd functions of R have no impact on the fitted value. For example the a^2 term in eq. (A1) does not contribute any difference to the fit.

We find that the fitted value differs from the background value H_s by a factor proportional to a and function of R_0/H_s ,

$$H_{s,\mathrm{fit}} = H_s + \frac{A}{\pi h} \frac{\Delta_H}{H_s} \underbrace{\left[2 \frac{R_0}{H_s} \left(6 - \left(\frac{4R_0}{H_s} \right)^2 \right) \mathrm{e}^{-4R_0^2/H_s^2} \right]}_{J(R_0/H_s)}, \tag{A2}$$

with the function J in brackets having a maximum close to 2 for $R_0 \simeq H_s/4$, as shown in Fig. A3.

We note that this perturbation is zero for $R_0 = 0$, meaning that a localized change at the center of the footprint does not modify the estimated H_s. This lack of impact on $H_{s,\mathrm{fit}}$ comes from the fact that the perturbation of the waveform (the second term in eq. (A1) is an odd function of range and thus orthogonal to the even functions that are the Brown waveforms with zero epoch $w_B(R, \sigma_H)$. The maximum perturbation of $H_{s,\mathrm{fit}}$ occurs for H_s perturbations at a range R_0 close to σ_H, i.e. corresponding to a distance from nadir of $r'_C/4$. Eq. (A2) gives results that are fairly robust for finite values of a/H_s^2, and would predict a wave height of 12.9 m in the case $R_0 = 2.5$ m shown in Fig. A1.

We now consider the average effect of the perturbation by computing the average over R_0, taking all values of R_0 from 0 to nH_s, which corresponds to averaging over an

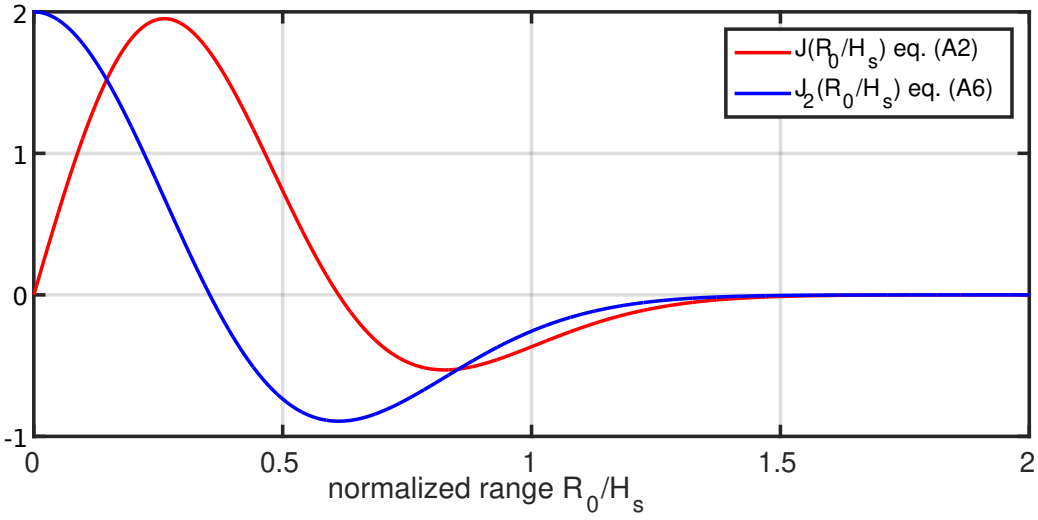

Figure A3. Functions $J(R_0/H_s)$ and $J_2(R_0/H_s)$ corresponding to the term in square brackets in eqs. (A2) and (A8). The maximum of J is at $R_0/H_s = 0.5\sqrt{0.5(3-\sqrt{6})} \simeq 0.26$, where J takes a value close to 1.96. This location corresponds to a distance from nadir approximately $\sqrt{0.26}r'_C \simeq r'_C/2$.

area $B = \pi n {r'_C}^2 = 2n\pi H_s h$. The integral of the function in brackets is

$$I = \int_0^\infty 2\frac{R_0}{H_s}\left(6 - 16\frac{R_0^2}{H_s^2}\right) \mathrm{e}^{-4R_0^2/H_s^2} \mathrm{d}R_0 = 0.5H_s. \tag{A3}$$

As a result, the average effect of a Δ_H change over an area $A = \pi\alpha^2 {r'_C}^2 = 2\pi\alpha^2 h H_s$ is, when n is large,

$$\delta_{H,\mathrm{alti}} = \frac{1}{nH_s}\int_0^{nH_s} (H_{s,\mathrm{fit}} - H_s)\,\mathrm{d}R_0 = \frac{1}{2n}\frac{A}{\pi h}\frac{\Delta_H}{H_s} = \frac{\alpha^2}{n}\Delta_H. \tag{A4}$$

This average effect of the localized perturbation of H_s is the same as a true area average, which is the perturbation times the ratio of the areas A and B, namely $\delta_H = \Delta_H A/B$. In other words, the perturbation is amplified if located at $0.15 < r/r'_C < 0.34$ from nadir, by a factor J that is up to 2. Otherwise the perturbation is attenuated, so that on average it is equal to the true perturbation. This averaging property and the unbiased estimate of $H_{s,\mathrm{fit}}$, with a perturbation that changes sign when Δ_H changes sign, are specific to the simple least squares used here. For example, fitting the logarithm of the waveform produces a biased estimator and a non-zero response for $R_0 = 0$. Hence the results presented here are specific to the fitting method.

In practice, distributed anomalies of H_s are not only a function of the distance from nadir, so that a local estimate of H_s will combine positive and negative anomalies Δ_H

777 that are located at the same distance from nadir, and will partially cancel. This explain
778 that our best fit for r_0 is $r_C/4.5$, smaller than the $r_C/2$ which is a more typical scale of
779 the footprint. Instead of retracking the simulated altimeter data, we can reproduce the
780 H_s variability by first summing the Δ_H anomalies for a given r, compute the $H_{s,\text{fit}}$ anomaly
781 for that r using eq. (A2) and then sum those anomalies for all r, as demonstrated in Fig. A4.
782 This procedure is equivalent to a spatial filter $\mathcal{J}(r)$ that is built from the J function, con-
783 verting the range $h + R_0$ to a horizontal distance from nadir $r = \sqrt{2hR_0}$,

784 $$\mathcal{J}(r) = g_{r_C}(r) + J \otimes (Id - g_{r_C})(r) \tag{A5}$$

785 where, g_{r_C} is a Gaussian filter with width r_C, and Id is the identity function. The es-
786 timated H_s thus comes from

787 $$H_s(x) = 4\sqrt{\frac{2}{\pi}}(\mathcal{J} \otimes \eta)(x) \tag{A6}$$

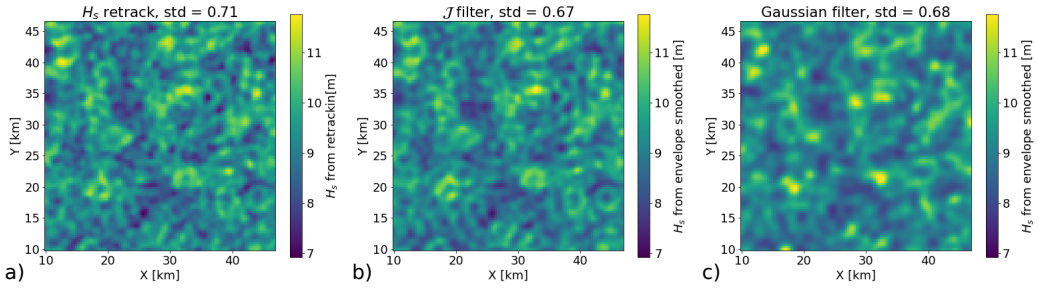

Figure A4. Equivalent to Fig. 3: (a) H_s surface obtained from retracking and H_s surfaces obtained from convoluting the envelope with (b) a spatial filter built from the J function, (c) a gaussian filter with $r_a = r_C/4.5$.

788 ## A2 Wave groups and sea level estimate

789 While perturbations at nadir do not change the H_s estimate, they would change
790 the mean sea level z_e (the epoch is $-z_e$) when using a 2-parameter waveform

791 $$w_{B2}(R, \sigma_H, z_e) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{(R + z_e) - h}{\sqrt{2}\sigma_H}\right)\right]. \tag{A7}$$

792 In the case shown in Fig. A1 with $R_0 = 0$ the estimated mean sea level is $z = -37$ cm.
793 We thus expect wave groups to contribute to fluctuations in the estimated sea level at
794 the scale of groups. The estimation of that effect follows the same method used above.

795 Fitting $w_{B2}(R, \sigma'_H, z_e)$ to our waveform $w(R)$ given by eq. (A1) is obtained by minimiz-
796 ing a modified cost function, that is the same as C but with one extra term $z_e \partial w_{B2}/\partial z_e$
797 inside the curly brackets, giving two extra non-zero terms proportional to z_e^2 and z_e. We
798 note that the cross-term proportional to $z_e(\sigma'_H - \sigma_H)$ is an odd function of R and thus
799 integrates to zero. After integration over R we get the cost function,

800 $$C_2 = C + \frac{z_e^2}{2\sqrt{\pi}\sigma_H} + \frac{a z_e}{4\sqrt{\pi}\sigma_H^2} \mathrm{e}^{-R_0^2/(4\sigma_H^2)} \left(\frac{R^2}{\sigma_H^2} - 2 \right).$$

801 Taking the derivative of C_2 with respect to z_e gives

802 $$z_e = -\frac{A\Delta_H}{8\pi h H_s} \underbrace{\left[\left(2 - 16 \frac{R_0^2}{H_s^2} \right) \mathrm{e}^{-4R_0^2/H_s^2} \right]}_{J_2(R_0/H_s)}. \tag{A8}$$

803 The function J_2 is plotted in Fig. A3. Hence z_e has the strongest deviation when the
804 wave height perturbation is centered at nadir, and the sign of the deviation is opposite
805 to Δ_H: i.e. a wave group centered at the nadir would give a spurious lower sea level. On
806 average the z_e deviation has a zero mean when R_0 is varied. As a result of the differ-
807 ent shapes of J and J_2, there is no simple correlation of the H_s and z_e perturbations,
808 contrary to the correlations induced by speckle noise in the waveform measurement (Sandwell
809 & Smith, 2005).

810 There is some correlation for R_0/H_s between 0.7 and 1.2 which may contribute to
811 anti-correlation of sea level anomalies and wave heights at scales around r_C, and thus
812 may persist in 1 Hz data. We insist that these are spurious sea level variations. In deep
813 water these spurious oscillations are much larger than the fraction of a millimeter asso-
814 ciated to true sea level variations with bound infragravity elevation that is anti-correlated
815 with the envelope of kilometer-scale wave groups (Ardhuin et al., 2004). The spurious
816 sea level oscillations described are also probably generally larger in amplitude than the
817 larger scale (20-km wavelength) true sea level variations associated to free infragravity
818 waves that have no phase correlation with the local envelope (Ardhuin et al., 2014). In
819 shallow water, the real sea level fluctuations can be more important.

820 ### A3 Speckle noise

821 Random fluctuations in the electromagnetic power measured by the radar combine
822 an additive thermal noise that can often be neglected and a multiplicative noise that is
823 caused by the Rayleigh fading of the interfering reflections off a random sea surface (Quartly
824 et al., 2001). In fact speckle is to the radar power what wave groups are to the wave en-

ergy. A good model for the speckle is a multiplicative random noise, so that the measured waveform for each range is multiplied by a factor $(1+\varepsilon(R))$ with $\varepsilon(R)$ following a χ^2 distribution with $N(R)$ degrees of freedom depending on the number of pulses averaged and the pulse repetition frequency (Quartly et al., 2001).

For the retracking, the effect of this speckle perturbation is one additional term $\varepsilon(R)w(R)$ inside the curly brackets of the cost function. Expanding the square and expressing the integral, it gives two terms, one proportional to $(\sigma'_H - \sigma_H)$ that is relevant to the H_s estimate and the other proportional to z_e fit, so that the cost function is now,

$$C_3 \simeq C_2 - 2(\sigma'_H - \sigma_H)\int_{-\infty}^{\infty}\varepsilon(R)w(R)\frac{\partial w_{B2}}{\partial \sigma_H}\mathrm{d}R - 2z_e\int_{-\infty}^{\infty}\varepsilon(R)w(R)\frac{\partial w_{B2}}{\partial z_e}\mathrm{d}R, \quad \text{(A9)}$$

with

$$\frac{\partial w_{B2}}{\partial \sigma_H} = -\frac{R-h+z_e}{\sigma_H^2\sqrt{2\pi}}\mathrm{e}^{-(R-h+z_e)^2/(2\sigma_H^2)}, \quad \text{(A10)}$$

and

$$\frac{\partial w_{B2}}{\partial z_e} = \frac{1}{\sigma_H\sqrt{2\pi}}\mathrm{e}^{-(R-h+z_e)^2/(2\sigma_H^2)}. \quad \text{(A11)}$$

The estimated wave height that gives $\partial C_3/\partial(\sigma'_H - \sigma_H) = 0$ thus has an extra term induced by speckle noise,

$$H_{s,\mathrm{fit}} = H_s + \frac{A}{\pi h}\frac{\Delta_H}{H_s}J(R_0/H_s) + 16\sqrt{2}H_s\int_{-\infty}^{\infty}\varepsilon(u)\left(1+\mathrm{erf}(2\sqrt{2}u)\right)u\mathrm{e}^{-8u^2}\mathrm{d}u, \quad \text{(A12)}$$

with $u = (R-h+z_e)/H_s$. The speckle-induced perturbation of $H_{s,\mathrm{fit}}$ is a weighted sum of random fluctuations with zero mean. In practice we can consider $\varepsilon(R)$ to be Gaussian, and the variance of the speckle perturbation is the sum of the variances associated to each range R times the weight squared. To get some useful order of magnitude we may take the variance of $\varepsilon(R)$, which is $1/N(R)$, to be constant at $1/N$. For large values of H_s, the discretized waveform is well approximated by the continuous form and the part of the variance of $H_{s,\mathrm{fit}}$ induced by the speckle is approximately 5.0 H_s/N, with a standard deviation $2.24\sqrt{H_s/N}$. Using the value $N = 512$ for the number of pulses of the SWIM nadir beam that we may assume to be independent, and $H_s = 2$ m, this gives a standard deviation of 0.14 m, broadly consistent with the background level in Fig. 8.c. However, we note that the magnitude of the variability of $H_{s,\mathrm{fit}}$ will depend on the method used to fit the waveform. In the case of the SWIM data, the adaptive method that is used is based on a maximum likelihood (Tourain et al., 2021). It is probably more robust to speckle noise perturbations than the least square estimate used here, in particular for this instrument that has a relatively high signal to noise ratio.

Acknowledgments

All CFOSAT data are provided by courtesy of CNSA and CNES. This research was made possible by support from ESA as part of the Sea State CCI project. MDC is currently supported by a postdoctoral grant from the Centre National d'Études Spatiales (CNES). MDC would like to thank JF Piolle from Ifremer for the data management support.

Open Research

The L2 SWIM dataset used here corresponds to the files reprocessed by CNES (2020) in version 5.1.2 and made available by CNES on the ftp server of AVISO+ (ftp-access.aviso.altimetry.fr, directory cfosat/swim_l2_op05), accessible to anyone after registration.

The L2P SWIM dataset used here corresponds to the files reprocessed by CNES/CLS (2021) in version 1.2 and made available by CNES on the ftp server of AVISO+ (ftp-access.aviso.altimetry.fr, directory cfosat/swim_l2p_box_nrt/), accessible to anyone after registration.

The L2S SWIM dataset used in this paper corresponds to the files reprocessed by Ifremer / CERSAT (2022) in version 1.0 and available at
https://data-cersat.ifremer.fr/projects/iwwoc/swi_l2s

The denoised CFOSAT data used for Fig. 9 were processed by Copernicus Marine Service / Centre National D'Etudes Spatiales / Collecte Localisation Satellite (2022) and distributed by AVISO+ with support from CNES (https://doi.org/10.24400/527896/a01-2022.013). The data are available on the ftp server of AVISO+ (ftp-access.aviso.altimetry.fr) after registration and selection of "Wave experimental products".

References

Abdalla, S., Janssen, P., & Bidlot, J.-R. (2011). Altimeter near real time wind and wave products: Random error estimation. *Marine Geodesy*, *34*, 396–406. doi: 10.1080/01490419.2011.585113

Alday, M., Accensi, M., Ardhuin, F., & Dodet, G. (2021). A global wave parameter database for geophysical applications. part 3: Improved forcing and spectral resolution. *Ocean Modelling*, *166*, 101848. doi: 10.1016/j.ocemod.2021.101848

Alday, M., & Ardhuin, F. (2023). On consistent parameterizations for both dom-

inant wind-waves and spectral tail directionality. *J. Geophys. Res.*, *128*, e2022JC019581. doi: 10.1029/2022JC019581

Altiparmaki, O., Kleinherenbrink, M., Naeije, M., Slobbe, C., & Visser, P. (2022). Sar altimetry data as a new source for swell monitoring. *Geophys. Res. Lett.*. doi: 10.1029/2021GL096224

Ardhuin, F., Chapron, B., & Elfouhaily, T. (2004). Waves and the air-sea momentum budget, implications for ocean circulation modelling. *J. Phys. Oceanogr.*, *34*, 1741–1755. doi: 10.1175/1520-0485%282004%29034%3C1741%3AWATAMB%3E2.0.CO%3B2

Ardhuin, F., Rawat, A., & Aucan, J. (2014). A numerical model for free infragravity waves: Definition and validation at regional and global scales. *Ocean Modelling*, *77*, 20–32. doi: 10.1016/j.ocemod.2014.02.006

Ardhuin, F., Stopa, J. E., Chapron, B., Collard, F., Husson, R., Jensen, R. E., ... Young, I. (2019). Observing sea states. *Frontiers in Marine Sci.*, *6*, 124. doi: 10.3389/fmars.2019.00124

Arhan, M., & Ezraty, R. (1978). Statistical relations between successive wave heights. *Oceanol. Acta*, *1*, 151–158.

Badulin, S. I. (2014). A physical model of sea wave period from altimeter data. *J. Geophys. Res.*, *119*, 856–869. doi: 10.1002/2013JC009336

Borge, J. C. N., Lehner, S., Niedermeier, A., & Shulz-Stellenfleth, J. (2004). Detection of ocean wave groupiness from spaceborne synthetic aperture radar. *J. Geophys. Res.*, *109*, C07005. doi: 10.1029/2004JC00298

Brown, G. S. (1977). The average impulse response of a rough surface and its applications. *IEEE J. Oceanic Eng.*, *2*(1), 67–63. doi: 10.1109/JOE.1977.1145328

Chelton, D. B., Walsh, E. J., & MacArthur, J. L. (1989). Pulse compression and sea level tracking in satellite altimetry. *J. Atmos. Ocean Technol.*, *6*, 407–438. doi: 10.1175/1520-0426(1989)006⟨0407:pcaslt⟩2.0.co;2

Chen, G., Chapron, B., Ezraty, R., & Vandemark, D. (2002). A global view of swell and wind sea climate in the ocean by satellite altimeter and scatterometer. *J. Atmos. Ocean Technol.*, *19*, 1849–1859.

CNES. (2020). *Wave products from SWIM Level 2 (L2), version 5.1.2* [Dataset]. AVISO+ with support from CNES. Retrieved from `ftp://ftp-access.aviso.altimetry.fr/` (after registration, directory: cfosat/swim_l2_op05)

CNES/CLS. (2021). *Wave products from SWIM Level 2+ Box Off Nadir NRT (L2P), version 1.2* [Dataset]. AVISO+ with support from CNES. Retrieved from `ftp://ftp-access.aviso.altimetry.fr/` (after registration, directory: cfosat/swim_l2p_box_nrt)

Collard, F., Marié, L., Nouguier, F., Kleinherenbrink, M., Ehlers, F., & Ardhuin, F. (2022). Wind-wave attenuation in arctic sea ice: A discussion of remote sensing capabilities. *J. Geophys. Res.*, *127*, e2022JC018654. doi: 10.1029/2022JC018654

Copernicus Marine Service / Centre National D'Etudes Spatiales / Collecte Localisation Satellite. (2022). *Wave experimental products: along-track significant wave height 5hz* [Dataset]. CNES. doi: "10.24400/527896/A01-2022.013"

Dodet, G., Abdalla, S., Alday, M., Accensi, M., Bidlot, J., & Ardhuin, F. (2022). Error characterization of significant wave heights in multidecadal satellite altimeter product, model hindcast, and in situ measurements using the triple collocation technique. *J. Atmos. Ocean Technol.*, *39*, 887–901. doi: 10.1175/JTECH-D-21-0179.1

Dodet, G., Piolle, J.-F., Quilfen, Y., Abdalla, S., Accensi, M., Ardhuin, F., ... Donlon, C. (2020). The sea state cci dataset v1: towards a sea state climate data record based on satellite observations. *Earth System Sci. Data*, *12*, 1929–1951. doi: 10.5194/essd-12-1929-2020

Egido, A., Dinardo, S., & Ray, C. (2021). The case for increasing the posting rate in delay/Doppler altimeters. *Adv. Space Res.*, *68*, 930–936. doi: 10.1016/j.asr.2020.03.014

Halimi, A. (2013). *From conventional to delay/doppler altimetry* (Doctoral dissertation, INP Toulouse, Toulouse, France). Retrieved from `https://theses.hal.science/tel-00951973`

Hanafin, J., Quilfen, Y., Ardhuin, F., Sienkiewicz, J., Queffeulou, P., Obrebski, M., ... Stutzmann, E. (2012). Phenomenal sea states and swell radiation: a comprehensive analysis of the 12-16 February 2011 North Atlantic storms. *Bull. Amer. Meterol. Soc.*, *93*, 1825–1832. doi: 10.1175/BAMS-D-11-00128.1

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

(Suppl. A)

Hauser, D., Tison, C., Amiot, T., Delaye, L., Corcoral, N., & Castillan, P. (2017). SWIM: The first spaceborne wave scatterometer. *IEEE Trans. on Geosci. and Remote Sensing*, *55*(5), 3000–3014.

Hauser, D., Tourain, C., Hermozo, L., Alraddawi, D., Aouf, L., Chapron, B., ... Tran, N. (2021). New observations from the SWIM radar on-board CFOSAT: Instrument validation and ocean wave measurement assessment. *IEEE Trans. on Geosci. and Remote Sensing*, *59*(1), 5–26. doi: 10.1109/tgrs.2020.2994372

Ifremer / CERSAT. (2022). *Global Ocean Directional Wave Parameters Level 2S from SWIM onboard CFOSAT for IWWOC project, version 1.0* [Dataset]. Ifremer, Plouzane, France. doi: 10.12770/12cfed8d-7645-442b-b8ef-a8d08decbaed

Jackson, F. C., Walton, W. T., & Peng, C. Y. (1985). A comparison of in situ and airborne radar observations of ocean wave directionality. *J. Geophys. Res.*, *90*(C1), 1005–1018.

Kudryavtsev, V., Yurovskaya, M., Chapron, B., Collard, F., & Donlon, C. (2017). Sun glitter imagery of surface waves. part 1: Directional spectrum retrieval and validation. *J. Geophys. Res.*, *122*. doi: 10.1002/2016JC012425

Lavrenov, I. V. (2001). Effect of wind wave parameter fluctuation on the nonlinear spectrum evolution. *J. Phys. Oceanogr.*, *31*, 861–873. Retrieved from `http://ams.allenpress.com/archive/1520-0485/31/4/pdf/i1520-0485-31-4-861`

Le Merle, E., Hauser, D., Peureux, C., Aouf, L., Schippers, P., Dufour, C., & Dalphinet, A. (2021). Directional and frequency spread of surface ocean waves from swim measurements. *J. Geophys. Res.*, *126*(7), e2021JC017220. doi: 10.1029/2021JC017220

Longuet-Higgins, M. S. (1984). New integral relations for gravity waves of finite amplitude. *J. Fluid Mech.*, *149*, 205–215. (see also Yu and Wu, J. Fluid Mech., 1987)

Masson, D., & Chandler, P. (1993). Wave groups: a closer look at spectral methods. *Coastal Eng.*, *20*, 249–275.

Passaro, M., Cipollini, P., Vignudelli, S., Quartly, G. D., & Snaith, H. M. (2014). ALES: A multi-mission adaptive subwaveform retracker for coastal and open ocean altimetry. *Remote sensing of Environment*, *145*, 173–189.

Passaro, M., Hemer, M. A., Quartly, G. D., Schwatke, C., Dettmering, D., & Seitz,

F. (2021). Global coastal attenuation of wind-waves observed with radar altimetry. *Nature communications*, *12*, 3812. doi: 10.1038/s41467-021-23982-4

Peral, E., Rodriguez, E., & Esteban-Fernandez, D. (2015). Impact of surface waves on SWOT's projected ocean accuracy. *Remote Sensing*, *7*(11), 14509–14529. doi: 10.3390/rs71114509

Quartly, G. D., Smith, W. H. F., & Passaro, M. (2019). Removing intra-1-Hz covariant error to improve altimetric profiles of σ_0 and sea surface height. *IEEE Trans. on Geosci. and Remote Sensing*, *57*(6), 3741–3752. doi: 10.1109/TGRS.2018.2886998

Quartly, G. D., Srokosz, M. A., & McMillan, A. C. (2001). Analyzing altimeter artifacts: Statistical properties of ocean waveforms. *J. Atmos. Ocean Technol.*, *18*, 2074–2091. doi: 10.1175/1520-0426(2001)018⟨2074:AAASPO⟩2.0.CO;2

Quilfen, Y., & Chapron, B. (2019). Ocean surface wave-current signatures from satellite altimeter measurements. *Geophys. Res. Lett.*, *216*, 253-261. doi: 10.1029/2018GL081029

Ray, C., Martin-Puig, C., Clarizia, M. P., Ruffini, G., Dinardo, S., Gommenginger, C., & Benveniste, J. (2015). SAR altimeter backscattered waveform model. *IEEE Trans. on Geosci. and Remote Sensing*, *53*, 911–919. doi: 10.1109/TGRS.2014.2330423

Rice, S. (1944). Mathematical analysis of random noise. In N. Wax (Ed.), *Noise and stochastic processes* (pp. 133–294). New York: Dover Publications Inc. (published 1954).

Rodriguez, E. (1988). Altimetry for non-gaussian oceans: Height biases and estimation of parameters. *J. Geophys. Res.*, *93*, 14107–14120. doi: 10.1029/JC093iC11p14107

Sandwell, D. T., & Smith, W. H. F. (2005). Retracking ers-1 altimeter waveforms for optimal gravity field recovery. *Geophys. J. Int.*, *163*, 79–89. doi: 10.1111/j.1365-246X.2005.02724.x

Saulnier, J.-B., Clément, A., de O. Falcão, A. F., Pontes, T., Prevosto, M., & Ricci, P. (2011). Wave groupiness and spectral bandwidth as relevant parameters for the performance assessment of wave energy converters. *Ocean Eng.*, *38*(1), 130-147.

Srokosz, M. A. (1986). On the joint distribution of surface elevation and slopes for

a non linear random sea, with an application to radar altimetry. *J. Geophys. Res.*, *91*, 995–1006.

Tayfun, A., & Lo, J.-M. (1989). Wave envelope and related spectra. *J. of Waterway, Port Coast. Ocean Eng.*, *115*, 515–533. doi: 10.1061/(ASCE)0733-950X(1989)115:4(515)

Timmermans, B. W., Gommenginger, C. P., Dodet, G., & Bidlot, J.-R. (2020). Global wave height trends and variability from new multimission satellite altimeter products, reanalyses, and wave buoys. *Geophys. Res. Lett.*, *47*, e2019GL086880. doi: 10.1029/2019GL086880

Toba, Y. (1973). Local balance in the air-sea boundary processes. III on the spectrum of wind waves. *J. Oceanogr. Soc. Japan*, *29*, 209–220. Retrieved from `http://www.terrapub.co.jp/journals/JO/JOSJ/pdf/2905/29050209.pdf`

Tourain, C., Piras, F., Ollivier, A., Hauser, D., Poisson, J. C., Boy, F., ... Tison, C. (2021). Benefits of the adaptive algorithm for retracking altimeter nadir echoes: Results from simulations and CFOSAT/SWIM observations. *IEEE Trans. on Geosci. and Remote Sensing*, *59*, 9927–9940. doi: 10.1109/TGRS.2021.3064236

Tournadre, J. (1993). Time and space scales of significant wave heights. *J. Geophys. Res.*, *98*(C3), 4727–4738.

Villas Bôas, A. B., Lenain, L., Cornuelle, B. D., Gille, S. T., & Mazloff, M. R. (2022). A broadband view of the sea surface height wavenumber spectrum. *Geophys. Res. Lett.*, *49*(e2021GL096699). doi: 10.1029/2021GL096699