



**HAL**  
open science

# On denoising satellite altimeter measurements for high-resolution geophysical signal analysis

Yves Quilfen, Bertrand Chapron

► **To cite this version:**

Yves Quilfen, Bertrand Chapron. On denoising satellite altimeter measurements for high-resolution geophysical signal analysis. *Advances in Space Research*, 2021, 68 (2), pp.875-891. 10.1016/j.asr.2020.01.005 . hal-04203139

**HAL Id: hal-04203139**

**<https://hal.science/hal-04203139v1>**

Submitted on 4 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

## On denoising satellite altimeter measurements for high-resolution geophysical signal analysis

Quilfen Yves <sup>1,\*</sup>, Chapron Bertrand <sup>1</sup>

<sup>1</sup> Laboratoire d'Océanographie Physique et Spatiale (LOPS), IFREMER, Univ. Brest, CNRS, IRD, IUEM, Brest, France

\* Corresponding author : Yves Quilfen, email address : [yquilfen@ifremer.fr](mailto:yquilfen@ifremer.fr)

---

### Abstract :

Satellite radar altimeter observations are key to advanced studies in ocean dynamics, particularly those focusing on mesoscale processes. To resolve scales below about 100 km, because altimeter measurements are often characterized by a low signal-to-noise ratio (SNR), low-pass filtering or least-squares curve fitting is generally applied to smooth the data before analysis. Here, we present an alternative method. It is based on Empirical Mode Decomposition (EMD) developed to analyze non-stationary and non-linear processes, which adaptively projects a signal on a basis of empirical AM/FM functions called Intrinsic Modulation Functions (IMFs). Applied to a Gaussian noise signal, the EMD provides a set of IMFs with a predictable distribution of noise energy that can be exploited by wavelet-inspired threshold methods to provide an efficient approach to data denoising. The EMD method performs a local analysis of the SNR, does not require a priori assumptions about the underlying geophysical signal, e.g., its degree of smoothness or its compliance with a particular mathematical framework. The signal is simply assumed to be the sum of a piecewise-smooth deterministic part and a stochastic part. The proposed EMD-based denoising approach is therefore well suited for mapping non-linear features, such as strong gradients, and extreme values without significant smoothing. Using Jason-2, Cryosat-2, and Saral/AltiKa significant wave height measurements, the method provides an effective means of mapping overlooked geophysical variability of sea state at scales between about 100 km and 25 km, a range largely impacted by low SNR. Below 25 km, a spectral hump caused by inhomogeneities in the altimeter footprint essentially dominates the signal. In addition, the EMD method provides a consistent approach for long-term noise analysis and monitoring under global and local conditions. The proposed method is a step forward that will enable better exploitation of the unique set of altimeter observations that now covers more than 25 years.

**Keywords :** Altimeter Measurement Noise, Empirical Mode Decomposition, Mesoscale Variability

## **1 – Introduction**

For more than 25 years, satellite altimeter measurements have supplied sea surface height (SSH), significant wave height ( $H_s$ ), and radar cross-section ( $\sigma_0$ ) measurements, building a unique dataset on these aspects of the ocean. Although this data offers potential for significant progress in various research-oriented and operational applications, its use can be hampered by technological differences between the sensors used and the different processing and re-tracking

algorithms applied to the analysis of altimeter signals. Efforts are being made to cross-calibrate measurements from the different satellite missions, in particular to improve their usefulness for climate studies, as in the European Space Agency's (ESA) Climate Change Initiative (CCI) project. One of the main limitations is that altimeter measurements are often characterized by a low signal-to-noise ratio (SNR), which can have a significant impact on geophysical analysis at spatial scales below 100 km. The main sources of noise are the altimeter waveform re-tracking algorithms, instrumental noise and geophysical variability in the altimeter footprint (Sandwell and Smith, 2005; Thibault et al., 2010, Dibarboure et al., 2014). In addition, because the recovered parameters are obtained from the same waveform re-tracking algorithm, they can have highly correlated errors, i.e., standard MLE4 processing produces four estimated parameters with correlated errors (SSH, Hs, sigma0, and off-nadir angle). Fine-scale ocean dynamics analysis therefore requires preliminary noise filtering, and low-pass or smoothing filters (e.g., Lanczos, running mean, or Loess filter) are frequently used. Each of these filters efficiently smooths noisy data but results in a systematic loss of small-scale ( $< 100$  km) geophysical information or the creation of artifacts in the geophysical variability analyzed. In addition, such filters only remove some of the high-frequency noise, which is also the result of approaches that infer a correction to eliminate correlated errors from other aspects of the waveform data (Zaron and de Carvalho, 2015; Quartly, 2019; Tran et al., 2019), leaving a substantial amount of low-frequency noise in the data. Such low-frequency noise remaining in the signals after smoothing will still interfere with the precise measurement of geophysical peak height, width, and position (O'Haver, 2019).

To overcome these difficulties, we present an alternative approach providing an adaptive noise elimination method for satellite altimeter measurements. It is based on the non-parametric Empirical Mode Decomposition (EMD) method developed to analyze non-stationary and non-linear signals (Huang et al., 1998; Wu and Huang, 2004; Huang and Wu, 2008). EMD is a scale decomposition into a limited number of amplitude and frequency modulated functions (AM/FM), among which the Gaussian noise distribution is predictable (Flandrin et al., 2004). It can therefore provide the basis for

a noise elimination approach, optimized and applied to a variety of simulated deterministic signals, with results often superior to those of wavelet-based techniques (Kopsinis and McLaughlin, 2009). For instance, to identify the occurrence of rogue waves, Hadjihosseini et al. (2014) used EMD to filter high-frequency variability related to the background sea state in a 1-Hz surface elevation record taken by an ultrasonic wave gauge in the Sea of Japan. Recently, EMD analysis has been successfully applied to altimeter data to analyze wave-current interactions (Quilfen et al., 2018; Quilfen and Chapron, 2019), known to predominate at scales below 100 km (Arduin et al., 2017). These studies showed how EMD analysis is suitable for processing non-stationary and non-linear signals in low SNR conditions, enabling a better exploitation of current and past satellite altimetry missions. More specifically, it can help to consistently recover power spectra of  $H_s$  and SSH variability, as well as to provide an accurate representation of non-linear features, such as strong gradients, and extreme values (Quilfen et al. 2018; Quilfen and Chapron, 2019). As the EMD denoising method was only briefly presented in these two previous articles, it is the purpose of the present study to present it in detail, especially since it is being used for processing of the upcoming sea state CCI and Copernicus Marine Environment Monitoring Service (CMEMS) products.

Particular attention is paid to comparing the results obtained from EMD with those obtained using a smoothing technique based on a Lanczos low-pass filter. The latter method is very popular among oceanographers, due to its simplicity and favorable characteristics. It is widely used for altimeter noise filtering, as performed in the current altimeter sea level and sea state products distributed by CMEMS (Taburet et SL-TAC team, 2018; Taburet et al., 2019). A Lanczos filter is a reformulated low-pass filter that incorporates a multiplicative factor to attenuate spectral ringing and signal energy leakage in the pass band (caused by overshoot ripples known as the Gibbs' phenomenon, Emery and Thomson, 2014). In practice, the signal is reconstructed by adding the low-frequency Fourier components to those within a frequency window, centered on the cut-off frequency, which are weighted. A side effect of the Lanczos filter is excessive signal smoothing in the presence of strong gradients and extreme values. Designing a practical Lanczos filter is thus not easy, because it requires setting of both the cut-off wavelength and the number of weighted Fourier terms required to achieve the desired roll-off

between the stop and pass bands. It always becomes a compromise. Signals filtered with a Lanczos filter are therefore strongly constrained by the inherent Fourier mathematical framework decomposition, while EMD filtered signals are based on a different assumption. The EMD decomposition assumes that the observed variable is the sum of a piecewise-smooth deterministic background part and a stochastic part.

Hereafter, for demonstration purposes and because it is implemented to process data distributed to users, this study focuses on the EMD-based method for processing Hs measurements. However, it also applies to SSH and sigma0 measurements. Section 2 presents the data used, section 3 the principles of EMD, its algorithm and properties, and section 4 describes the EMD-based denoising algorithm applied to altimeter measurements. In section 5, the results are presented for a few illustrative cases and the variances of both the resulting filtered signals and associated noise are analyzed. In addition, the uncertainty parameter associated with the denoised Hs is presented and characterized. Finally, section 6 provides a summary.

## **2 – Data**

The multi-mission altimeter dataset processed at IFREMER by Queffeuilou and Croizé-Fillon (2013) was analyzed from January 2014 to December 2016. This dataset contains inter-calibrated Hs and sigma0 measurements for all altimeter missions from the ERS-1 mission starting in 1991 and lasting until April 2017 for the Saral/AltiKa, Cryosat-2 and Jason-2 missions. The present study uses measurements from these last three altimeters. Each altimeter performed measurements at nadir along the satellite track, and standard processing delivers 1 Hz data with ground sampling that varied slightly from 6 to 7 km depending on the altimeter. The study uses quality-controlled, unfiltered and non-resampled, along-track Hs measurements. Data editing performed on the Queffeuilou and Croizé-Fillon (2013) dataset was based on testing the quality flags of the AVISO Geophysical Data Record products, which are the primary source, and on along-track data screening to eliminate anomalous Hs values that deviate too far from adjacent values. However, this did not eliminate all data contaminated by land or sea ice transitions, rain, slicks or very light winds (sigma0 blooms). Outliers associated with such

conditions are therefore likely to have remained in the analyzed dataset, causing the so-called spectral hump, which adds a spatially coherent error to the Gaussian white noise, especially at scales below about 25 km (Dufau et al., 2016; Dibarboure et al., 2014). Improved data editing will be a necessary to mitigate these errors.

### **3 – Empirical Mode Decomposition principles and algorithm**

EMD is a method to decompose signals into a limited number of scale-dependent components, called intrinsic modulation functions (IMFs), each modulated in amplitude and frequency. An IMF represents a simple oscillatory mode as a counterpart to a simple harmonic function derived from Fourier analysis. However, instead of imposing a constant amplitude and frequency in a simple harmonic component, an IMF can have an evolving amplitude and frequency. The decomposition is based on the local characteristic sampling scale of the data and is therefore applicable to non-linear and non-stationary processes. IMFs have well-behaved Hilbert transforms, from which instantaneous frequencies can be calculated. Thus, any event can be well localized in both the time and frequency domains (Huang et al., 1998).

Unlike a more conventional decomposition method such as the wavelet technique, which projects the signal considered onto a basis of predefined wavelet functions, IMF expansion depends only on the analyzed signal. IMF calculation is performed using an iterative procedure called sifting. By construction, IMFs share the following properties: they are zero mean, all their maxima and minima are respectively positive and negative, they have the same number (or + / - 1) of zero-crossings and local extrema, and they are modulated in both amplitude and frequency (AM/FM). EMD belongs to the class of multi-scale, orthogonal and adaptive decompositions. The EMD method ranks IMFs by construction from the first, containing the finest scales, to the last, containing the global trend.

EMD adaptively decomposes a signal  $x(t)$  into a small number  $L$  of IMFs

$h^{(i)}(t)$ ,  $1 \leq i \leq L$ , so that :

$$x(t) = \sum_{i=1}^L h^{(i)}(t) + d(t) \quad (1)$$

where  $d(t)$  is a residue that is a non-zero mean slowly varying function with only a few extrema.

The IMF number,  $L$ , depends on the length of the record and typically varies from 1 to 10 for the lengths analyzed in the altimeter dataset.

The IMFs are calculated successively, the first containing the shortest scales by the construction of the algorithm. Each IMF is estimated using an iterative process called sifting (illustrated in Figure 1), which is successively applied to the residual multicomponent signal  $x^{(i)}(t)$  with :

$x^{(i)}(t) = x(t), i = 1$  to derive the first IMF

$$x^{(i)}(t) = x(t) - \sum_{j=1}^{i-1} h^{(j)}(t), i \geq 2 \text{ to derive the IMFs of rank } > 1 \quad (2)$$

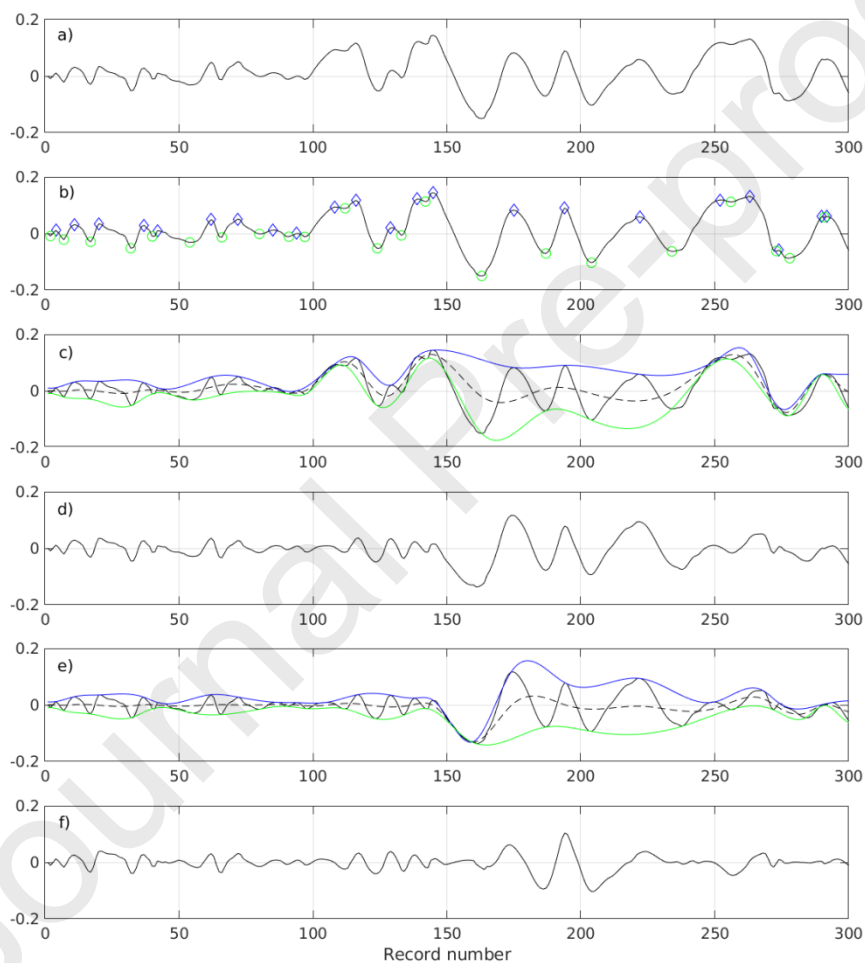
For the estimation of an IMF of rank  $i$   $h^{(i)}(t)$ , during the  $(n + 1)^{\text{th}}$  sifting iteration, a temporary IMF estimate  $h_n^{(i)}(t)$ , called proto-IMF, gets improved according to the following steps:

- 1) Find the local maxima and minima of  $h_n^{(i)}(t)$ .
- 2) Interpolate, using natural cubic splines, along the maxima and minima of  $h_n^{(i)}(t)$  found in the first step in order to form an upper and a lower envelope.
- 3) Calculate the average of the two envelopes.
- 4) Obtain a refined estimate  $h_{n+1}^{(i)}(t)$  of the proto-IMF by subtracting the average found in the previous step.
- 5) Repeat the process from step 1 unless the stopping criterion has been met.

The sifting process to estimate an IMF of rank  $i$  is illustrated in Figure 1. The signal called “arbitrary input” in panel (a) can be the original signal  $x(t)$  for the estimation of the first proto-IMF of the first IMF, or the residual signal for the estimation of higher ranking IMFs, or a proto-IMF obtained during the sifting process for any IMF estimation. Panel (a) presents a smooth input signal for clarity. Panels (b) to (e) show the steps of the sifting process for two iterations, i.e., searching for extrema in panel (b), determining the upper and lower envelopes in panel (c), and estimating a refined proto-IMF by removing the average of the upper and lower envelopes in panel (d). Panel (e) is like panel (c) but an iteration later, showing that the average of the two envelopes is approaching 0. Panel (f) shows the



IMF resulting from the sifting process. This IMF signal is removed from the input signal in panel (a) for estimation of the rank  $i+1$  IMF, or for estimation of the residue if the number of extrema is too small. As can be seen by comparing panels (a) and (f), the sifting process has isolated the shortest scales as an AM/FM function.



upper and lower envelopes (blue and green lines) and their average (dashed line) of the proto-IMF superimposed on the new input; and (f) the IMF resulting from the sifting process.

To best eliminate the noise in Hs or SSH measurements, the stopping criterion should be met when the number of iterations of the sifting process has reached a predefined number. A value of eight iterations is used to comply with the optimization work performed by Flandrin et al. (2004) and by Kopsinis and McLaughlin (2009). Other stopping criteria may be used. The reason for choosing a given number of iterations is the behavior of the energy distribution of an EMD-processed Gaussian noise according to the IMF ranking, as detailed in section 4. Figure 2 gives an example of an AltiKa Hs record and associated IMFs. The shortest scales are mapped in the first IMF, the second panel from the top in Figure 2, to likely contain mostly noise information. The bottom panel shows the last IMF and the residual. For this data record, six IMFs and the residual explain the total signal.

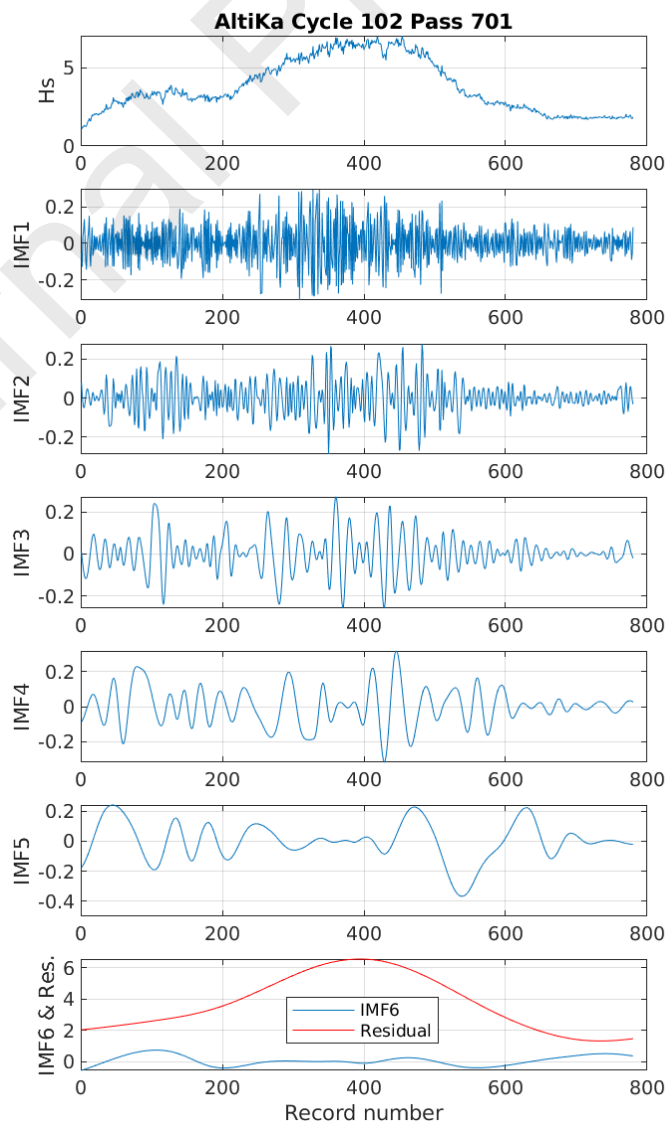


Figure 2: EMD expansion in L IMFs and residue for a segment of AltiKa Hs measurements (October 7, 2016). (upper panel) raw Hs ; (intermediate panels) L-1 IMFs ; (bottom panel) last IMF and residue. As obtained, six IMFs plus the residue describe the total signal.

#### **4 – Data denoising from Empirical Mode Decomposition**

Most denoising methods consist of three steps: projecting the signal on a basis of functions capable of usefully distributing and distinguishing the underlying signal and noise; then making a threshold selection (e.g., hard or soft for wavelet filters, frequency or wavelength cut-off for low-pass filters) in the transformed domain; and finally recovering the signal by back-projecting the modified coefficients into the initial domain. Wavelet and Fourier analyses are often used to effectively decompose signals. With Fourier low-pass filtering, the density energy contained in scales below a given threshold is eliminated. As Gaussian noise contamination can extend over the entire harmonic range, this operation likely eliminates the useful signal at scales below the prescribed cut-off scale and leaves noise above this prescribed scale. This makes it impossible to determine an optimal cut-off wavelength since the actual characteristics of the SNR vary over space and time. Low-pass filtered

signals are also subject to spectral ringing (i.e., forced oscillations are generated) and Fourier analysis is not the most effective method for decomposing non-linear and non-stationary signals.

Wavelet decomposition is a more powerful technique, suitable for decomposing non-stationary signals. However, it requires the design of the basis of functions on which the signal is projected, i.e., the choice of the mother wavelet and the number of vanishing moments. This design can be optimized for well-characterized signals but is difficult for Hs or SSH measurements whose geophysical variability and associated noise can have very different signatures. This is the main reason why wavelet shrinkage methods have not yet been successfully used to eliminate noise from altimeter measurements. In this context, the adaptive EMD decomposition provides a simplified framework that overcomes the need to design the basis of functions.

Once the signal is broken down into a set of IMFs, a denoising strategy inspired by those used for wavelet techniques can be applied. Indeed, the analysis to be carried out takes advantage of 1) the well-behaved and predictable distribution of Gaussian noise energy with the IMF basis, 2) the legacy of decades of wavelet-based denoising techniques, and 3) an ensemble average approach to estimate a robust noise-free signal.

#### ***4.1 Characteristics of the Gaussian noise spectrum with an IMF basis***

Reporting on numerical experiments, Flandrin et al. (2004) showed that in the case of pure fractional Gaussian noise, the first IMF possesses the characteristics of a high-pass filter while the higher order modes behave similarly to a dyadic filter bank, which is reminiscent of what is observed in wavelet decompositions. A bank of dyadic filters is obtained by decomposing the input signal into octave frequency bands in different channels, for which, as descending the frequency scale, the successive frequency bands have half the width of their predecessors. For the EMD decomposition of a white noise, the logarithm (base 2) of the number of zero-crossings (which is an approximate indication of the average frequency of each mode) is shown to decrease linearly with the increasing rank of the IMF, with a slope close to -1. Flandrin et al. (2004) further found that, for IMFs ranking  $>$

1, the IMF's power spectra are self-similar, i.e., the power spectra of all IMFs could be reduced to a single curve when properly renormalized. The result is a well-defined behavior of the decrease of noise energy with increasing IMF ranking. The Gaussian noise variance projected onto the IMF basis can be modeled as follows:

$$\text{var}(h_n(t)) \sim 2^{(\alpha-1)n} \quad (3)$$

where  $n > 1$  is the IMF rank and  $\alpha$  depends on the Hurst exponent of the fractional Gaussian noise (i.e.,  $\alpha = 0.5$  for an uncorrelated noise, e.g., white noise;  $\alpha \neq 0.5$  for an autocorrelated noise).

With the EMD basis, noise energy decreases rapidly with the increasing IMF rank:  $\sim 59, 20.5, 10.3, 5.2, 2.6\%$  of total energy for the first five IMFs, respectively. The first four IMFs represent  $\sim 95\%$  of the noise energy.

The reasoning behind the EMD denoising approach is derived from this result. Since the IMF energy resulting from the decomposition of a noise signal is predicted from Eq. (3), then in real cases of signals containing both information and additive Gaussian noise, a significant difference between the predicted noise energy and the energy of the IMF signal indicates the presence of useful information. Eq. (3) then gives the expected noise energy in each IMF to determine the different thresholds below which signal fluctuations are associated with noise.

As an illustration, Figure 3 shows how the EMD filter bank distributes a white noise signal and the Jason-2 altimeter Hs signal in the Agulhas Current region. The standard deviation of noise was adjusted to fit the Hs background noise at scales  $< 20$  km. In agreement with Flandrin et al. (2004), for a white noise, the EMD filter bank is composed of a high-pass filter, IMF1, and a dyadic filter bank for higher ranking IMFs. It is interesting to note that a similar structure is observed when EMD is applied to the Hs along-track signal. Hs IMF1 resembles the noise IMF1, confirming that it contains mainly noise. A higher power spectral density at scales  $> 20$  km is associated with a higher noise modulation, since the noise amplitude depends on Hs and therefore varies along the altimeter track and

with very large geophysical Hs gradients, which are common in the Agulhas current region. Pure noise and Hs higher ranking IMFs share the same frequency ranges and, in the logarithm representation used, the SNR increases very rapidly with the increase of IMF rank. This gives the practical rule for denoising since the signal modulation in each IMF can be compared to the noise energy expected in a noise-only IMF of the same rank. Hs IMF2 and IMF3 show comparable energy at the 70-km scale, below which IMF2 energy dominates. On average, Hs IMF2 is noise only at a 30-km scale, and the SNR increases to  $\sim 0.5$  at a 50-km scale which means that some geophysical signals can be retrieved locally under conditions where the SNR is still larger.

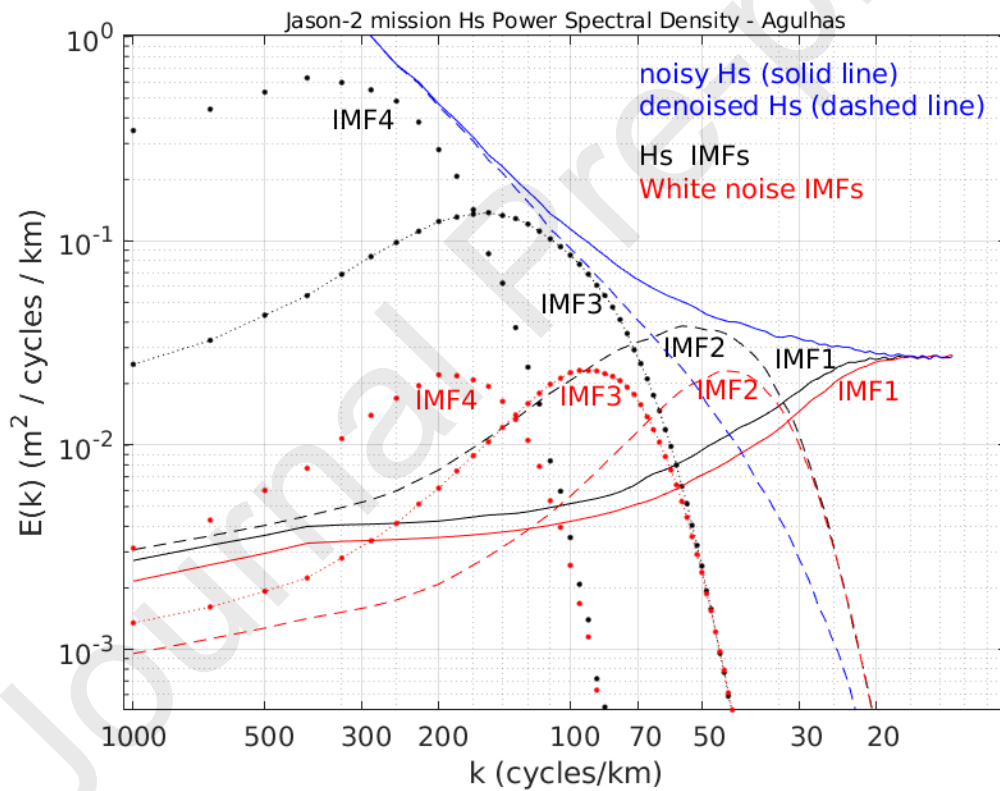


Figure 3: Mean Power Spectral Density (PSD) of the first four IMFs for white noise (red curves) and Jason-2 Hs along-track measurements (black curves), and mean PSD of the corresponding noisy (solid blue line) and denoised (dashed blue line) Jason-2 Hs measurements. The PSD is the average of PSDs computed over all data segments covering the years 2014 to 2016, and the Agulhas region ( $10^{\circ}\text{E} - 35^{\circ}\text{E}$ ;  $45^{\circ}\text{S} - 33^{\circ}\text{S}$ ).

For the altimetry dataset, the white noise hypothesis ( $H = 0.5$ ) is made, following studies showing a near-white noise spectrum (Zaron and de Carvalho, 2015; Xu and Fu, 2012; Sandwell and Smith, 2005). The reference noise energy level is then estimated from the first IMF, which contains mainly noise. The theoretical noise energy level of each IMF of rank  $n > 1$  is then given by:

$$E_n = \frac{E_1}{0.719} 2.01^{-n} \quad (4)$$

The two constants may vary slightly depending on the number of sifting iterations used to calculate the IMFs. The values in the above equation are optimized and correspond to eight sifting iterations (Flandrin et al., 2004). For the estimation of  $E_1$ , a robust estimator is the absolute median deviation (MAD) from zero, calculated as follows:

$$E_1 = (\text{median}|n_1(t)|/0.6745)^2 \quad (5)$$

where  $n_1$  is the noise contained in IMF1, evaluated from a wavelet analysis as described in the following section. The formulation based on a scaled median value is well suited for estimating noise variance, with the factor 0.6745 being a scaling for estimating the noise standard deviation for a Gaussian distribution (Johnstone and Silverman, 1997). It is robust to cases where the analyzed signal, here  $n_1$ , contains residual values associated with a small amount of geophysical information (large  $H_s$  gradients) or outliers (Mallat, 2009).

#### 4.2 Signal denoising

The denoised signal  $y(t)$  is then reconstructed as :

$$y(t) = \sum_{i=M_1}^{M_2} \tilde{h}^{(i)}(t) + \sum_{i=M_2+1}^L h^{(i)}(t) + d(t) \quad (6)$$

where  $h(t)$  represents the IMFs associated with signal  $x(t)$ , and the tilde over  $h(t)$  indicates the IMFs that are effectively denoised with the threshold technique.  $M_1$  and  $M_2$  are two parameters that give the flexibility to exclude lower-order IMFs than  $M_1$ , if they are thought to contain only noise, and to give optional denoising for higher-order IMFs than  $M_2$ , understood as containing very little noise. Indeed,

in the case of a white noise signal, the fourth component contains four times less noise energy than the second, and the first four IMFs contain 95% of the total noise energy. This strongly “colored” nature of the noise with the EMD basis has the advantage that few IMFs of a real signal are significantly contaminated by noise, with the SNR increasing rapidly with IMF rank as shown in Figure 3. For  $M_1$ , a value of 1 was chosen for the analyzed dataset, because the strong transient or nonlinear features associated with the large Hs or SSH gradients are important geophysical information to be recovered, which are likely to appear in the first IMF. This is illustrated in Figure 4, which shows the raw and filtered Jason-2 Hs measurements (upper panel) with very large gradients ( $\sim 1.5$  m / 10 km) triggered by wave-current interactions. Locally, close to the maximum of Hs, the IMF1 shows a large amplitude modulation well above the average noise level.

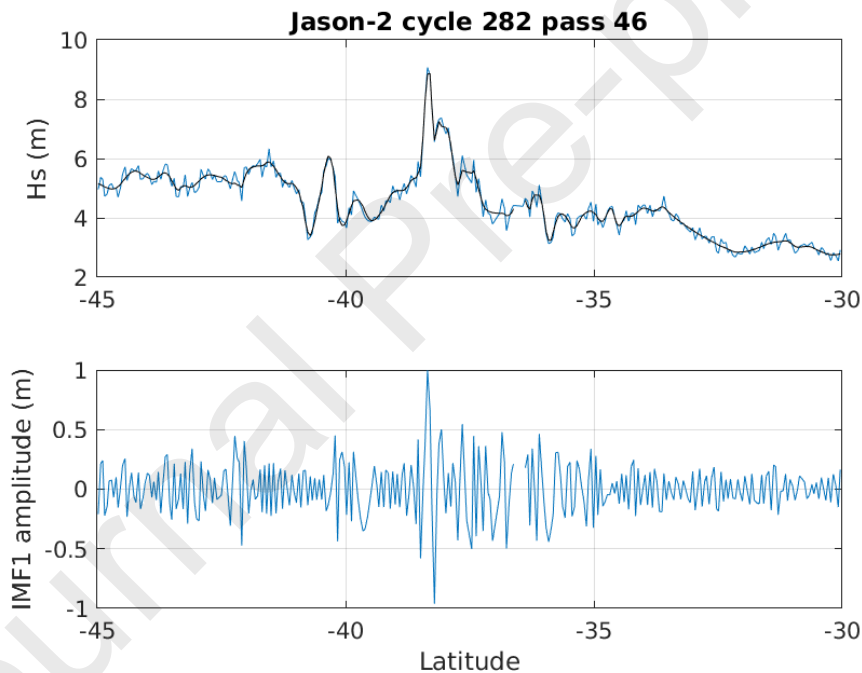


Figure 4: (top) Jason-2 Hs along-track raw (blue) and filtered (black) measurements as a function of latitude; (bottom) associated IMF1 obtained from EMD expansion.

For a given noisy input signal, the SNR and robustness of the denoised signal are increased by estimating the final result as an ensemble average of several denoised signals. The denoising algorithm follows steps such as:



- 1) Perform an EMD expansion of the noisy signal  $x(t)$
- 2) Perform the IMF1 noise filtering to separate the IMF1 stochastic noise,  $n_1(t)$ , from the targeted signal and possible outliers. This is done using the wavelet denoising approach described in Huang and Cressie (2000). Indeed, IMF1 is a purely oscillating function with a limited bandwidth, dominated by noise, and sharing the same properties and almost the same frequency range everywhere. It makes an adequate wavelet design much easier and robust
- 3) Perform a reconstruction of signal  $x_s(t)$  by adding the residual signal, after eliminating noise and outlier in IMF1, with higher order IMFs
- 4) Randomly modify the positions of the noise  $n_1(t)$  to obtain a new noise realization  $n_k(t)$  and a new noisy signal  $x_b(t) = x_s(t) + n_k(t)$
- 5) Perform an EMD expansion of  $x_b(t)$
- 6) Carry out the denoising of IMFs ranking from  $M_1$  to  $M_2$  by hard thresholding, where the threshold is based on  $E_n$  (Eq. 4 and 5) and reconstruct a denoised signal  $x_k(t)$  using Eq. (6). The advantage of hard thresholding is that it preserves strong gradients better than soft thresholding (Kopsinis and McLaughlin, 2009).
- 7) Iterate  $k$  times steps 4 to 6 in order to obtain a set of denoised signals
- 8) Make an ensemble average of the  $x_k(t)$  to obtain a robust denoised signal and an uncertainty estimate calculated as the standard deviation of  $x_k(t)$

Step 2 of the proposed algorithm is IMF1 denoising with a double objective: first, to evaluate the zero-mean stochastic noise that will be used to generate the ensemble of noisy signals by random permutation of the noise samples; second, to identify the high SNR signal in IMF1 that will be kept in the reconstructed signal. High flexibility is provided by the use of Huang and Cressie (2000) wavelet shrinkage method. These authors developed a Bayesian approach to denoise various signals, assuming that it can be the sum of the underlying targeted signal composed of a piecewise-smooth deterministic part and a stochastic part, plus a noise as a zero average stochastic part. For the analyzed altimeter

dataset, the signal may also contain outliers, which can be part of IMF1. Outliers can be caused in particular by rain and sigma0 blooms, resulting in corrupted altimeter waveforms and abnormal Hs values, which are difficult to edit. The impact of outliers on wavelet analysis can often result in high values of the coefficients at first level (the finest details) that will be targeted as a useful signal by the wavelet shrinkage algorithm. The best way to eliminate these outliers from both the signal and noise estimations, is to analyze the IMF1 with the wavelet bank in two different steps. First, for the estimation of the useful signal in IMF1, the first wavelet scale containing the finest details is systematically rejected in the reconstruction of the signal. It acts as a low-pass filter with a  $\sim 25$  km cut-off wavelength, the wavelength below which the spectral hump dominates (Dibarboure et al., 2014). However, outliers can also corrupt the higher-level wavelet scales, so better editing of the data source would still be necessary. As shown in Figure 4, the strongest geophysical Hs gradients are projected in the largest scale levels of the wavelet decomposition and are therefore not impacted by this low-pass filtering. Second, for the IMF1 noise estimation, all scale levels are used to separate noise, associated with wavelet coefficients below the prescribed threshold, from useful information and outliers, associated with wavelet coefficients above the prescribed threshold. For the wavelet basis design, Symmlet-4 is chosen as the mother wavelet.

Step 2 of the algorithm then provides an estimate of the IMF1 noise,  $E_1$ , in order to calculate the noise energy  $E_n$  in the other IMFs from Eq. (4), and the associated thresholds  $T_n$  to be used in step 6).  $T_n$  are derived from  $E_n$  following the rules that are generally adopted in wavelet denoising techniques. Mallat (2009) showed that the maximum amplitude of Gaussian noise with variance  $\sigma^2$  has a very high probability of being just below  $T = \sigma\sqrt{2\log N}$ , where  $N$  is the number of samples. Indeed, a greater amplitude of noise is likely to occur with increasing sample size. Following this threshold formulation, longer altimeter records also have a higher probability of including higher Hs values for which the associated noise is higher, which requires a higher threshold value. Mallat (2009) has shown that this formulation is too restrictive and that a lower threshold is a better choice for reducing the risk of

filtering the underlying signal. The threshold formulation is then refined by introducing a factor  $A$ , giving for an IMF of rank  $n$  and sample size  $N$ :

$$T_n = A\sqrt{E_n * 2\log N} \quad (7)$$

with  $A < 1$  being a fine adjustment factor, and the  $\log N$  factor is constrained to vary within a limited range to account for a possible change in noise amplitude statistics for longer data records, as explained above. A standard value of 0.7 was determined for  $A$  in the current Hs denoising algorithm. This parameter can be refined if necessary to account for the differences in the different altimeter missions. Besides, the sensitivity of the results to a prescribed threshold is in practice reduced by the ensemble average performed to estimate a denoised robust signal.

A detailed description of the interval hard thresholding operation can be found in Kopsinis and McLaughlin (2009) and subsequent publications dealing with the wavelet thresholding technique. Its adaptation to IMF denoising is illustrated in Figure 5. Black dotted oscillations between the thresholds indicated by the green lines are canceled, and the red dotted line represents the recovered signal.

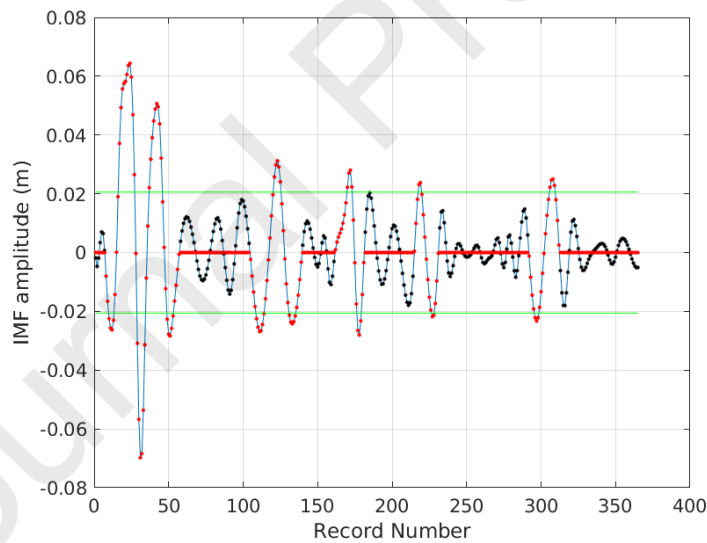


Figure 5: IMF amplitude before (solid blue line) and after (red dots) thresholding. Black dots represent the canceled signal and solid green lines indicate the threshold values.

## 5 – Results

### 5.1 Examples

Figures 6 to 7 show the results for the two altimeter records presented in the previous sections. On the upper panels, raw data (blue solid line), EMD-denoised data (black solid line) and data filtered using the Lanczos filter with a cut-off wavelength of 60 km (red dashed line) are displayed. For the long record in Figure 6, the details do not appear and the two filtered signals seem to be in good agreement. The bottom panel shows the noise that has actually been eliminated. The estimated noise amplitude varies along the track, showing an expected dependence on  $H_s$ , with values more than twice as high for  $H_s > 5$  m as for  $H_s \sim 2$  m.

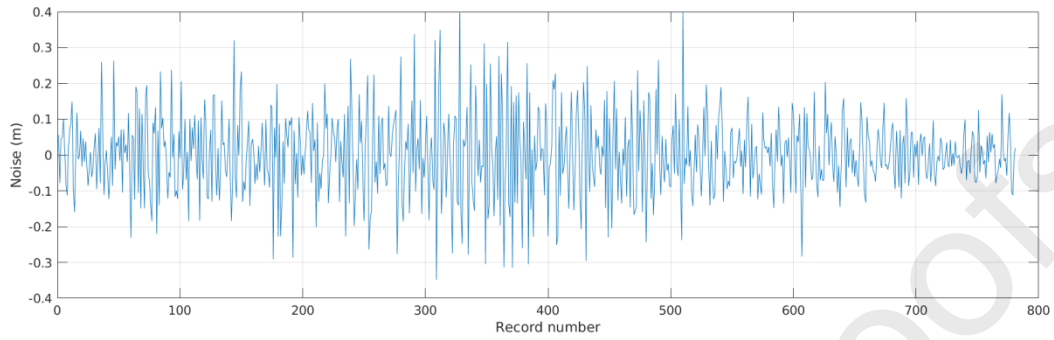
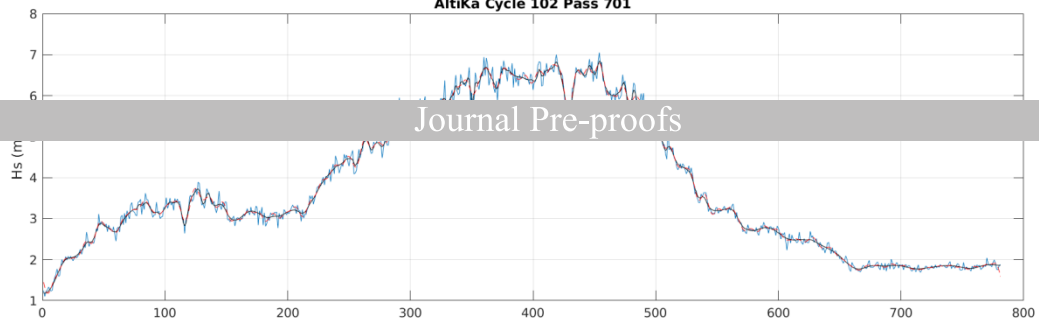


Figure 6: AltiKa 1-Hz along track Hs measurements (October 7,

2016). (upper panel) raw Hs (solid blue line), EMD-denoised Hs (solid black line), Lanczos-filtered Hs (dashed red line). (bottom panel) noise estimated as the difference between raw Hs and EMD-denoised Hs. The abscissa spatial sampling is about 7 km.

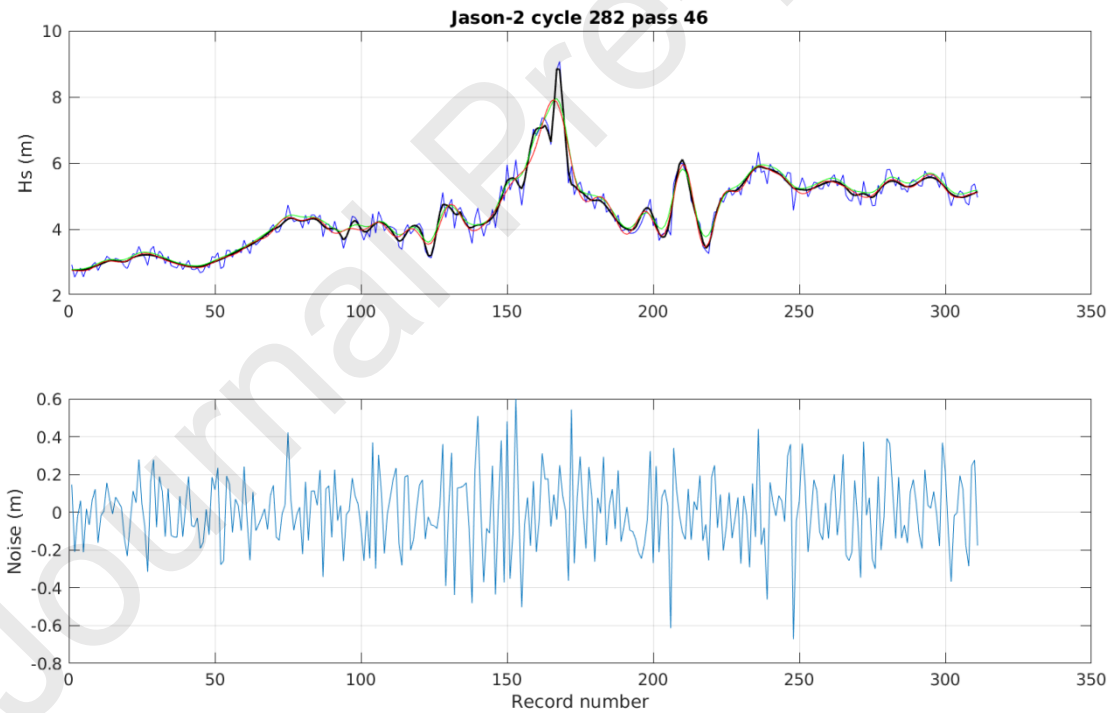


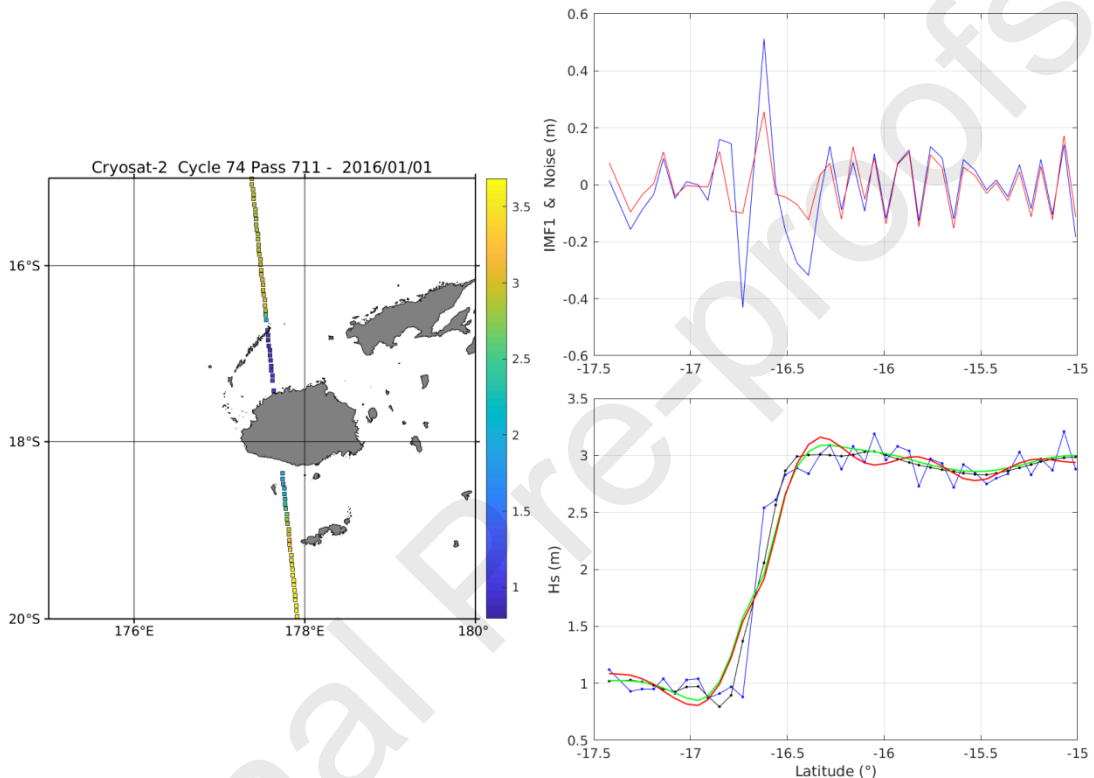
Figure 7: Jason-2 1-Hz Hs measurements (February 29, 2016). (upper panel) raw Hs (solid blue line), EMD-denoised Hs (solid black line), Lanczos-filtered Hs with smooth (solid green line) and sharp (solid red line) roll-

off. (bottom panel) noise estimated as the difference between raw Hs and EMD-denoised Hs. The abscissa sampling was about 5.9 km.

The example in Figure 7 is a particularly good illustration of the capability of EMD-based denoising to map very large gradients and avoid smoothing the maximum values. These large Hs values and associated gradients are the result of energetic wave / current interactions (Quilfen et al., 2018; Quilfen and Chapron, 2019). Since the algorithm is based on measurements extrema detection and removal of the estimated noise, rather than on smoothing, the denoised maximum Hs reaches a peak close to 9 m, about 20 cm less than the raw data. This contrasts with Lanczos filtering at a 60 km cut-off (value used in CMEMS operational products), which significantly modifies gradients and peak values, regardless of how smooth the roll-off is. As tested, a 30 km cut-off for Lanczos filter will effectively restore the observed Hs gradient and maximum, but at a cost of a large number of artifact oscillations produced elsewhere. This expected behavior of low-pass filtering (Emery and Thomson, 2014) is illustrated further in Figures 8 and 9.

Figure 8 shows a Cryosat-2 pass across the island of Fiji, with an abrupt change in sea state on the northern side of the island, between the lagoon area and the high seas. The reason for such a marked Hs gradient is that the lagoon is protected by the Yasawa coral reefs. The data point in this area is missing from the original data as it was corrupted by the ground, but the altimeter track segments on both sides were processed as a continuous segment for illustration purposes. Indeed, high resolution altimetry can provide data very close to the coast and situations involving very rapid changes in fetch conditions will often be encountered. The lower right panel of Figure 8 shows the raw and denoised Hs when using the EMD and Lanczos processing approaches. Although the Lanczos filter is specifically designed to reduce the Gibbs effect, at the cost of gradient smoothing, residual artifact oscillations are still produced on both sides of the gradient and their amplitude depends on the number of coefficients (size of the convolution kernel) used to control the desired roll-off near the cut-off wavelength. The solid green line shows the results for a number of coefficients locally adjusted to best reduce the Gibbs oscillations, and the solid red line shows the results for a number of coefficients

ensuring a sharp cut-off of the filter. This choice has a significant impact, both locally and globally, as discussed in the following section. EMD denoising provides a more realistic estimate of the  $H_s$  profile. Indeed, as shown in the upper right panel of Figure 8, the IMF1 wavelet analysis effectively separates noise from the frontal signature that occurs near  $16.7^\circ$  of latitude. To avoid artifacts and data loss near the edges of the data record for the Lanczos processing, the same mirror padding was applied



for both EMD and Lanczos processing.

Figure 8: Left panel: Cryosat-2 1-Hz  $H_s$  measurements over the Fiji island (January 1, 2016). Right upper panel: IMF1 (blue line) and noise estimated from wavelet analysis of IMF1 (red line). Right bottom panel: raw  $H_s$  (solid blue line), EMD-denoised  $H_s$  (solid black line), Lanczos-filtered  $H_s$  with smooth (solid green line) and sharp (solid red line) roll-off as a function of latitude.

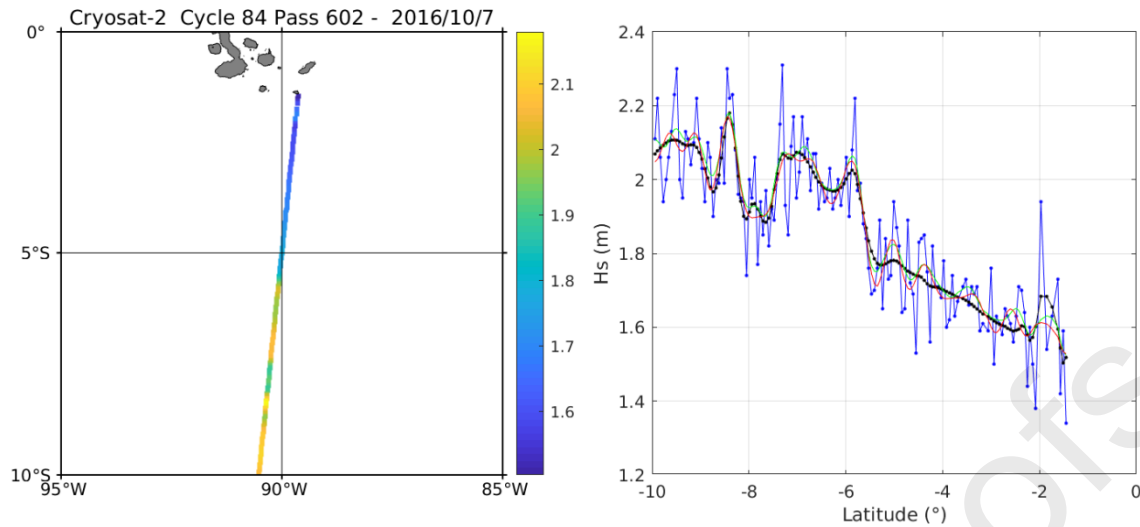


Figure 9: Left panel: Cryosat-2 1-Hz Hs measurements over the Galapagos Islands (October 7, 2016). Right panel: raw Hs (blue solid line), EMD-denoised Hs (solid black line), Lanczos-filtered Hs with smooth (solid green line) and sharp (solid red line) roll-off as a function of latitude.

In addition, the data filtered by the Lanczos filter contain a significant amount of statistically insignificant modulations, constrained by the Fourier analysis; this is illustrated in Figure 9, which shows a Cryosat-2 pass across the Galapagos islands. This is an area where the SAR mode is systematically put into operation, resulting in a 1-Hz Pseudo Low Resolution Mode (P-LRM) noise increase as shown in the right panel of the figure. In this low SNR context, small mesoscale variability is embedded in noise with similar or greater amplitude. In such conditions, Hs derived from EMD varies more smoothly as a result of the analyzed low SNR, while Hs derived from the Lanczos filter shows forced oscillations, for example near 5°S latitude. The EMD result is constrained by the local SNR analyzed while the Lanczos result is strongly constrained by the Fourier analysis framework.

### 5.2 – Analysis of the denoised signal and noise signatures

Figure 10 shows the Hs wavenumber spectra in the Agulhas current region (10 - 35 E ; 45 - 33 S) for the period 2014–2016. The raw measurements spectra flatten at scales < 100 km where the measurements are more significantly contaminated by noise. The differences in the average noise levels of the different altimeters are as expected, with Saral data being less noisy than Cryosat-2 data.



Agulhas is one of the regions where the Cryosat-2 Synthetic Aperture Radar mode is systematically used, resulting in noisier data for the 1-Hz low rate mode. In other regions where SAR mode is not used, Cryosat-2 noise should be lower than Jason-2 noise. The spectra of denoised measurements are almost identical for all three altimeters, which gives a good indication that the EMD processing worked satisfactorily. The PSD obtained from the Lanczos filtering is shown for a sharp (solid green line) and smooth (dashed green line) roll-off at the cut-off wavelength. As expected, the results are very different, which raises the question of the choice of this control parameter. Surprisingly, this has not been addressed in studies using the Lanczos filter for altimeter data filtering. For scales between 300 and 30 km, the PSD slope calculated for data filtered by EMD is not linear. However, it compares favorably with a  $k^{-3}$  law, result in agreement with the study by Ardhuin et al. (2017) based on high resolution numerical simulations of the WaveWatch-III model. Below a scale of about 25 km, the PSD vanishes for the three altimeters as a consequence of the cancellation of IMF1's finest wavelet scales (as discussed at the end of section 4). The results can be analyzed in more detail from the interpretation of the total noise PSD (see Figure 10 left panel), and of the IMF1 noise PSD (see Figure 10 right panel). The total noise is calculated as raw Hs minus filtered, which therefore include outliers. The total noise spectrum exhibits the same behavior for all three altimeters, with a "plateau" for scales below  $\sim 25$  km, at the same PSD level as the raw Hs signals. This part of the signal is completely suppressed as explained above, as it essentially corresponds to the sum of the instrumental Gaussian white noise and the hump artifact (Dibarboure et al., 2014). The PSD of total noise gradually diminishes with the increase in scale, indicating that significant low-frequency noise energy is removed from IMFs of order  $> 1$ , which can be evaluated from the difference between the total noise and the IMF1 noise PSDs (Figure 10, left and right panels).

The IMF1 noise PSD, which takes into account high frequency Gaussian noise at scales below about 25 km, is slightly lower than the raw Hs PSD. This difference should correspond to outlier signatures and is small because the  $\sigma_0$  bloom occurrence is weak in the Agulhas region (Tournadre et al., 2006). A different characterization of the measurement noise, essentially free of outliers, can then be made using the high frequency noise estimated from the processing of IMF1. Indeed, EMD analysis is

an efficient way to isolate the high frequency part of the total noise variance in IMF1, which theoretically represents about 59% of the total noise for a white noise signal. Since the noise variance left in other IMFs is simply related to IMF1 noise variance through Eq. (3), IMF1 noise gives a good representation of the characteristics of the total noise affecting the Hs noisy measurements. It is also the noise figure used to perform the ensemble averaging operation that ultimately gives the denoised signal and associated uncertainty. The results are presented in Figure 11.

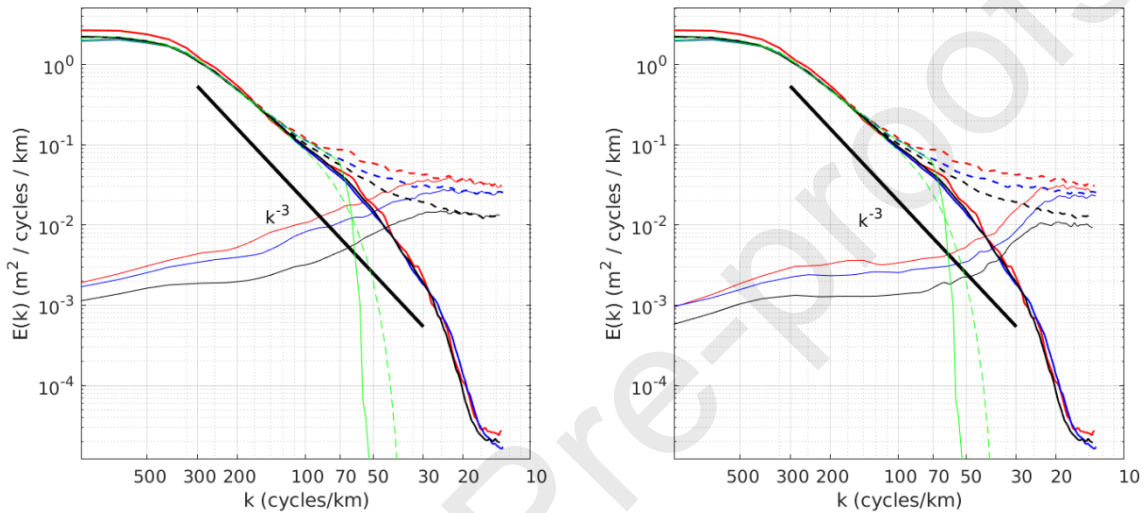


Figure 10: Power spectral density (PSD) as a function of wavenumber (x-axis, labels in km) for altimeter Hs measurements: SARAL/Altika (black), Cryosat-2 (red) and Jason-2 (blue), period 2014–2016, Agulhas region. Dashed lines show raw 1 Hz data, thick solid lines show filtered data, and thin solid lines show the signal removed by the whole process (left panel) and for the IMF1 noise (right panel). The solid (dashed) green line gives the PSD for Jason-2 data processed with a Lanczos filter of 60 km cutoff wavelength with sharp (smooth) roll-off. The solid black line gives the  $k^{-3}$  law between 300 km and 30 km wavelength (right).

As mentioned earlier, outliers affect the EMD analysis. A more thorough verification of the data is therefore carried out before the noise analysis. The data are rejected when the  $\sigma_0$  is greater than the 99.5<sup>th</sup> percentile, when the IMF1 absolute value is greater than the 99.9<sup>th</sup> percentile, and when the distance to the coast is less than 50 km. The  $\sigma_0$  threshold allows the editing for  $\sigma_0$ s, and corresponds to 18.9, 18.7, 19.4 dB for Cryosat-2, Jason-2, and Saral/AltiKa, respectively. The standard deviation of IMF1 noise in one-meter Hs bins is shown in Figure 11. The results are consistent with

simulations performed with recent retracking algorithms (Thibault et al., 2017), showing an increasing standard deviation of noise with the increase in  $H_s$  and an overall lower standard deviation of noise for Saral/AltiKa. The noise levels of Cryosat-2 and Jason-2 are of similar amplitude although this global analysis includes areas where the Cryosat-2 SAR mode is used, for which noise in 1-Hz low resolution mode is higher. Figure 11 (lower panel) shows the probability density function (PDF) of the IMF1 noise amplitude for the three altimeters, and that obtained from Monte-Carlo simulations of the IMF1 amplitude for Gaussian noise with zero-mean and standard deviation of 12 cm (green dots) and 7 cm (magenta dots). Two hundred realizations of a Gaussian noise with 512 samples are performed for the Monte Carlo simulation. The good agreement between the measured and simulated PDFs is a good indication that the chosen denoising approach is consistent. A best fit between the altimeters and the simulated PDFs gives an average estimate of the  $H_s$  measurement noise standard deviation at 12., 12.75, and 7. cm for Cryosat-2, Jason-2, and Saral/AltiKa, respectively. Although not strictly comparable, these values are consistent with the 9, 12, and 6 cm values obtained from the standard deviation of the high-resolution measurements used to calculate 1-Hz  $H_s$  (Taburet et al., 2019), since our estimate for Cryosat-2 includes the areas with SAR operations.

The EMD method therefore also allows noise analysis and monitoring over time and under local conditions. The noise characteristics presented should be significantly improved by the application of improved re-tracking algorithms (e.g., Passaro et al., 2014; Thibault et al., 2017), as planned in the ESA sea state CCI project.

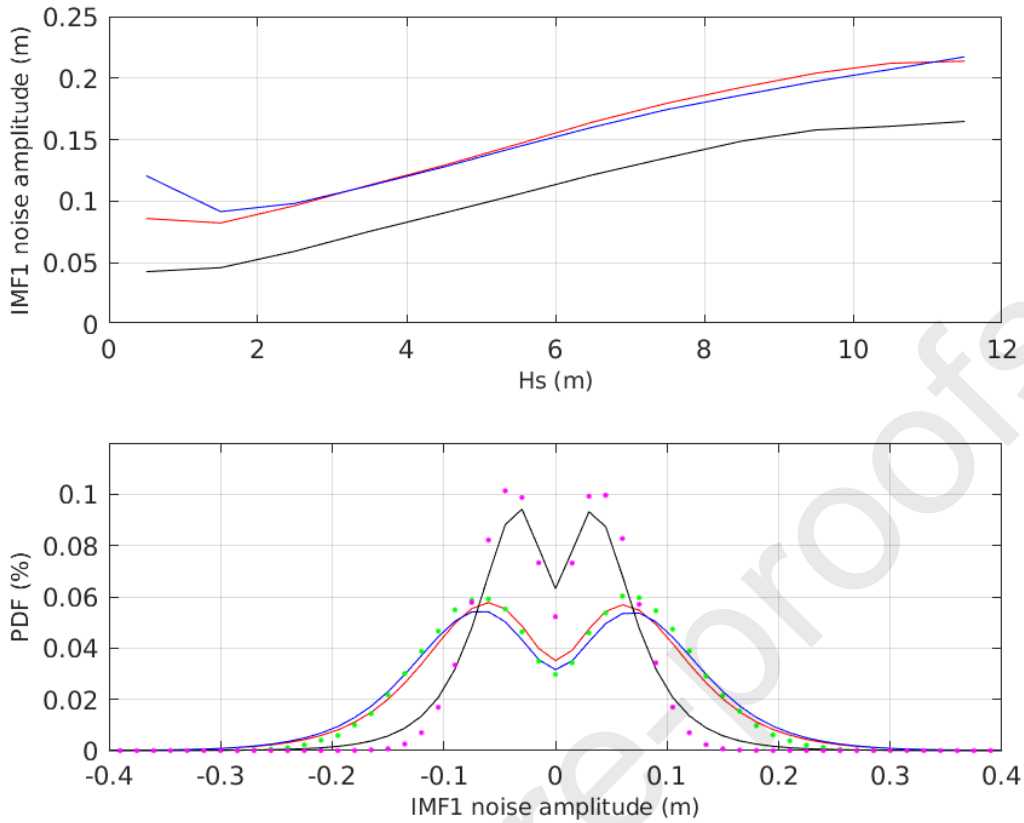


Figure 11: (top) Standard deviation (m) of IMF1 noise amplitude as a function of  $H_s$  (m) in one-meter bins; (bottom) probability density function (%) of IMF1 noise amplitude (m). SARAL/Altika (black), Cryosat-2 (red) and Jason-2 (blue). The dots map the PDF, obtained from Monte-Carlo simulations, of the IMF1 amplitude for gaussian noise with zero-mean and standard deviation of 12 cm (green dots) and 7 cm (magenta dots).

Figure 12 shows the spatial distribution of the IMF1 variance, calculated as the sum of the square amplitudes in  $1^\circ \times 1^\circ$  longitude/latitude boxes. Patterns are clearly apparent in the IMF1 variance of the noisy dataset, the left panels, which are related to a higher noise variance associated with regions with a higher mean sea state. For Cryosat-2, in the center of the left panel, other patterns with high variance show the areas where the altimeter is used in SAR mode. When the EMD is applied to the denoised data, the IMF1 variance (right panels) is associated with the variance of the finest mesoscale geophysical signals (in the range  $\sim 30$  to  $\sim 100$  km). It shows that EMD denoising effectively filters noise in the measurements of the different altimeters, to map the overlooked geophysical variance patterns in a consistent manner. These patterns are mainly associated with

increased sea state variability resulting from wave–current interactions that dominate  $H_s$  variance at scales below 100 km (Ardhuin et al., 2017 ; Quilfen et al., 2018 ; Quilfen and Chapron 2019) .

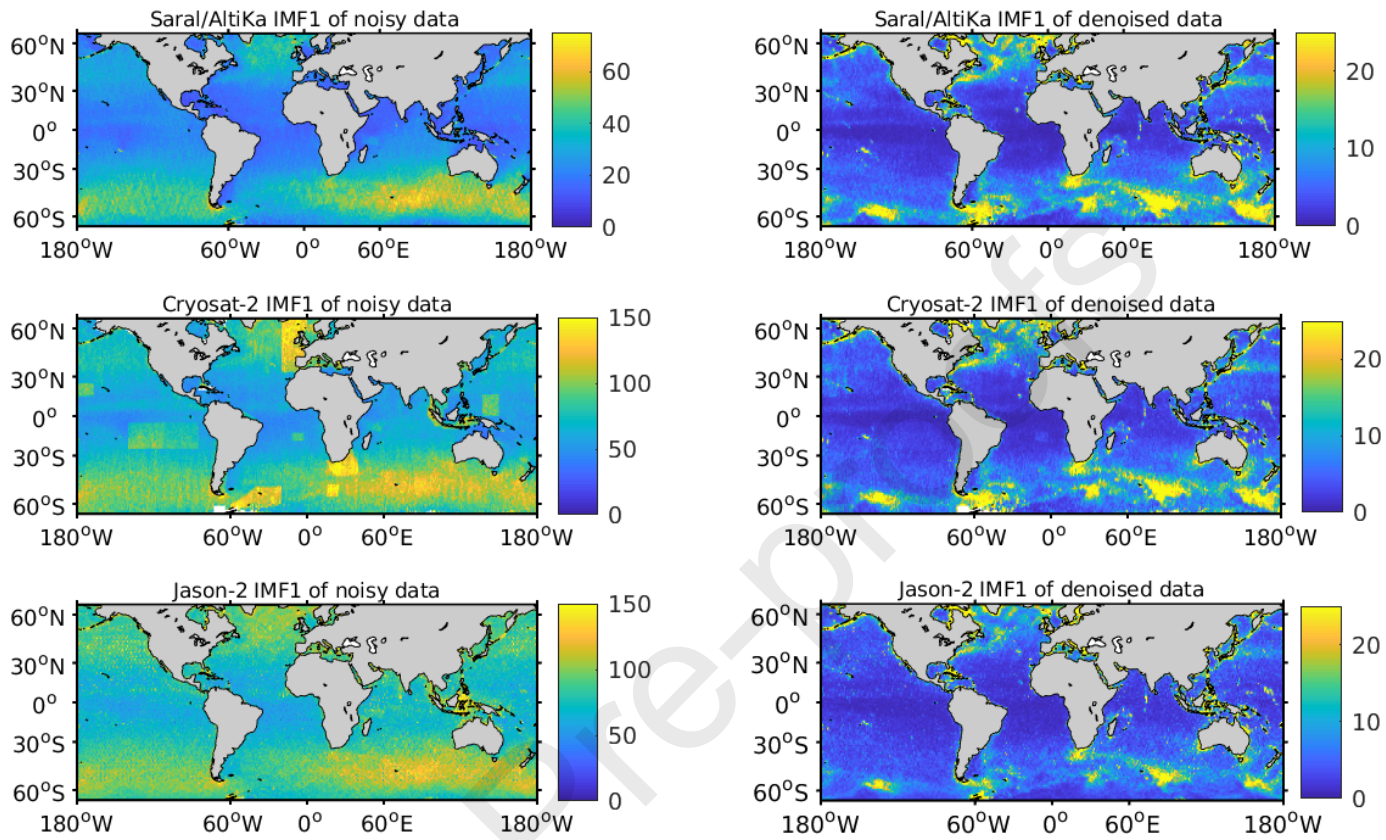


Figure 12 : IMF1 mean square amplitude ( $\text{cm}^2$ ) of  $H_s$  raw (left panels) and denoised (right panels) measurements, period 2014 / 2016. Top: Saral/AltiKa; Middle: Cryosat-2; Bottom: Jason-2.

### 5.3 – Uncertainty associated with the denoised significant wave height estimate

In the previous section, the noise derived from EMD processing and associated with standard 1Hz measurements is analyzed and characterized. However, this does not characterize the uncertainties associated with the “denoised significant wave height” parameter. Such uncertainties could be assessed using conventional approaches, such as comparisons with a reference dataset or a crossover analysis, but this will give an estimate of the average uncertainty rather than an estimate associated with each measurement. The EMD ensemble average approach allows this to be done in line with the estimate of denoised  $H_s$ . Indeed, the uncertainty  $\varepsilon$  can be calculated for each measurement along the

altimeter track as the standard deviation of the ensemble of denoised signals used to calculate the average denoised  $H_s$ . The probability density function of  $\varepsilon$  is shown in Figure 13, top panel. It reflects the results presented in the previous section: the noisier the standard 1 Hz measurements, the greater the uncertainties associated with the denoised measurements. The central panel also shows that  $\varepsilon$  decreases with decreasing significant wave height, except for the lower values for which Jason-2 and Cryosat-2 values are slightly higher. This can be analyzed in more detail by computing  $\varepsilon$  as a function of the radar cross-section,  $\sigma_0$ , which is mainly related to wind speed. At low  $\sigma_0$  values, corresponding to strong winds and large associated  $H_s$ ,  $\varepsilon$  values are larger as expected.  $\varepsilon$  values decrease with decreasing  $\sigma_0$ , reaching a minimum close to 12dB, then increase with decreasing wind speed. Uncertainties associated with sea state measurements are also expected to be greater under light wind conditions, but Jason-2 and Cryosat-2  $\varepsilon$  values increase more rapidly. This is probably related to the fact that the light wind zones are climatologically associated with precipitation and surface slicks. Under such conditions, the better Saral/AltiKa SNR combined with the larger rain contamination in the Ka band makes data editing and noise filtering more efficient. Indeed, the number of edited data, before data denoising, is significantly higher for Saral/AltiKa.  $\varepsilon$  is then a practical uncertainty parameter, taking into account the different errors likely to affect the denoised  $H_s$  data and their dependence on local conditions, which should be useful in particular for assimilation in numerical wave models. To further highlight its behavior, Figure 14 shows the mean geographical distribution of  $\varepsilon$  computed over a three-year time period (2014–2016). As expected, the uncertainty is related to the altimeter measurements signal to noise ratio: it is higher for Jason-2 than for Saral/AltiKa, it is higher under high sea state conditions that predominate in southern oceans and in the North Atlantic Ocean, and is higher in the regions of increased sea state mesoscale variability as depicted in Figure 12 (right panel) for which the SNR is effectively lower.

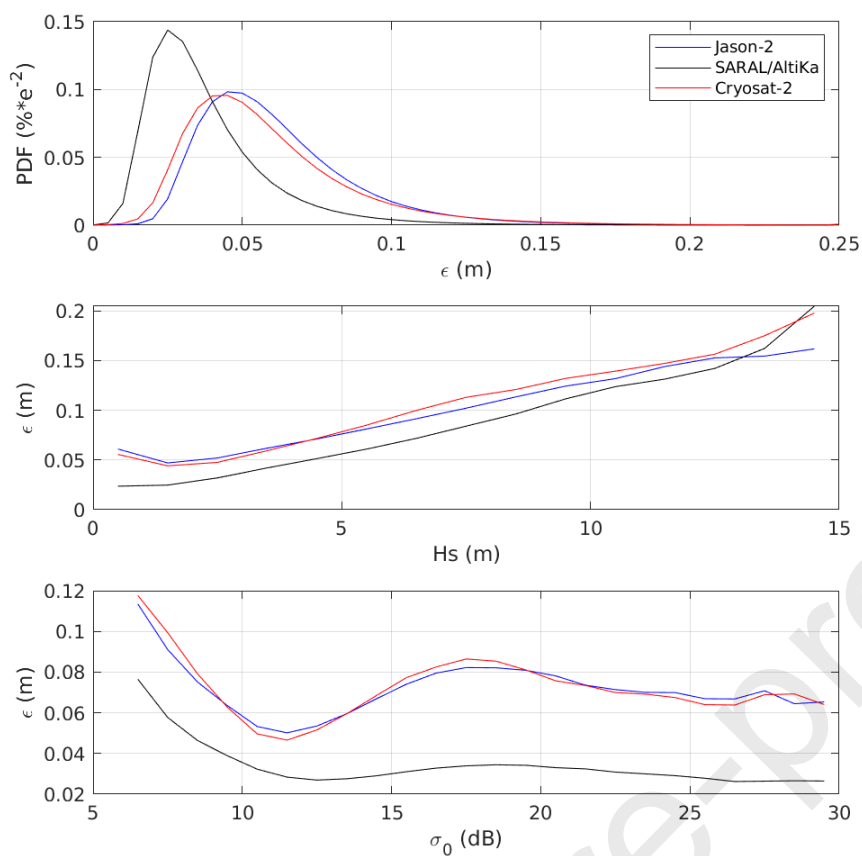


Figure 13: Statistical behavior of the  $\epsilon$  uncertainty parameter (meters) for SARAL/Altika (black), Cryosat-2 (red) and Jason-2 (blue): (top panel) probability density function; (middle panel)  $H_s$  dependency; (bottom panel) radar cross-section, ( $\sigma_0$ ) dependency.

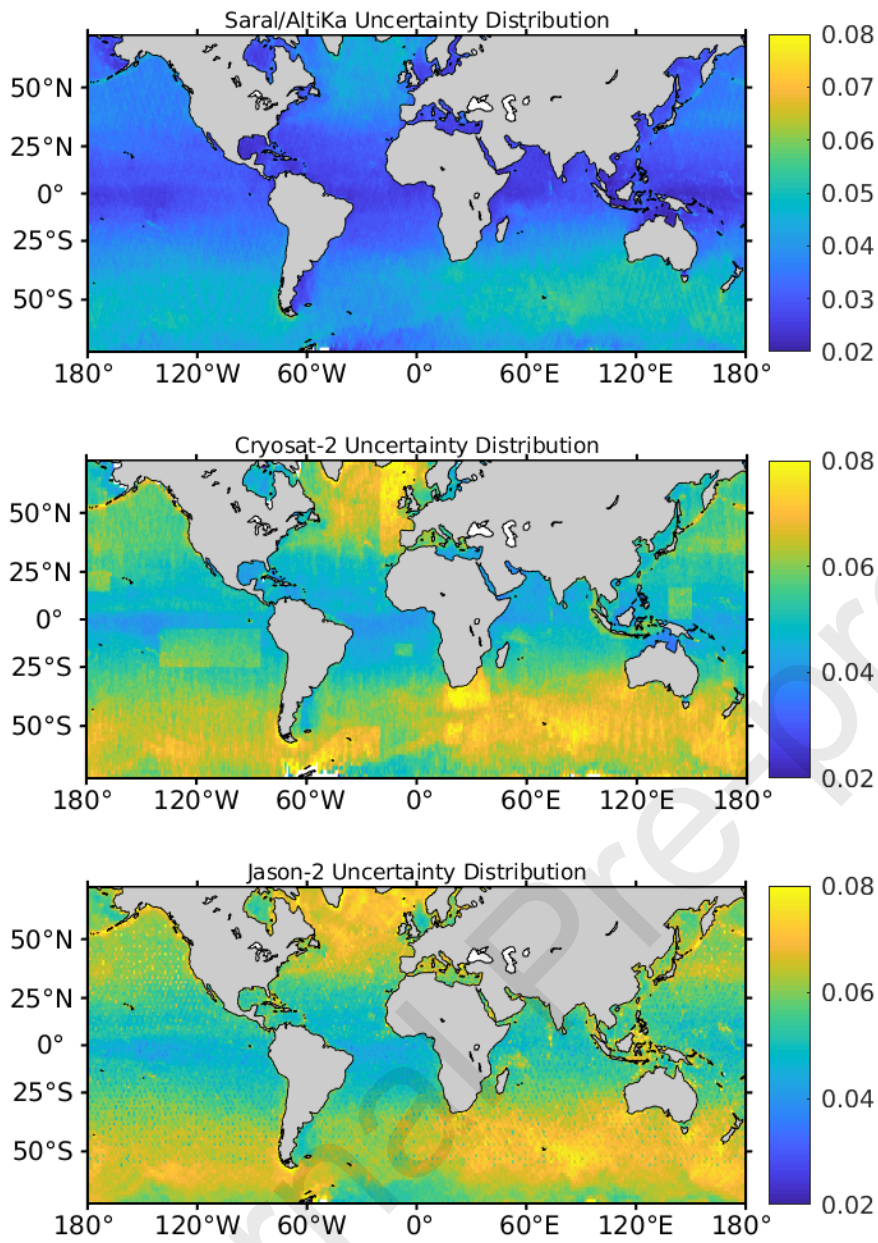


Figure 14: Geographical distribution of the  $\epsilon$  uncertainty parameter (meters) for Jason-2 (upper panel), Cryosat-2 (intermediate panel, and SARAL/Altika (lower panel).



## 6 – Summary

An adaptive method to efficiently eliminate and characterize altimeter data noise was defined and applied. Our method is essentially built on Empirical Mode Decomposition (EMD), which projects a signal on a basis of empirical AM/FM functions, called Intrinsic Modulation Functions (IMFs), whose properties offer a new way of analyzing and denoising along-track altimeter measurements. When applied to a Gaussian noise signal, EMD provides a set of IMFs with a predictable energy distribution. This property was therefore used to define rules for altimeter data denoising, which are based on existing wavelet filtering techniques. The chosen approach simply assumes that the  $H_s$  along-track variability is the sum of an underlying signal composed of a piecewise-smooth deterministic part and a stochastic part, allowing the use of a Bayesian analysis, and a zero-mean Gaussian noise. The stochastic hypothesis for the wave field is a framework common to many studies, as well as in traditional wave models such as WAVEWATCH-III, to take into account the random nature of the wave field.

The method performs a local SNR analysis for each processed data record. For each record, whose length varies from about 50 km to 1500 km depending on the continuity of the data, the final noise-free altimeter signal is estimated as an ensemble average of several signal realizations. A fine adjustment of the result is possible thanks to a global scaling parameter, making it possible to adapt the thresholds used to separate the noise from the signal. It allows, if necessary, to optimize the results according to the different sensors on-board the different altimetry missions. Complementary to methods based on crossovers analysis, EMD provides also a mean to effectively estimate the noise characteristics for the various altimeter sensors. The noise variance parameter derived from the EMD should therefore be useful for noise monitoring over time and under local conditions and can provide practical information in various applications such as altimeter data assimilation to numerical models.

The proposed method is free of systematic artifacts, preserves the amplitude of spatial gradients and extreme values, and eliminates the noise over the whole frequency range, which is not the case for the boxcar or low-pass filters currently applied to altimetry products distributed to users. Indeed, such filters only remove some of the high-frequency noise, which is also the result of approaches that infer

a correction to eliminate correlated errors from other aspects of the waveform data (Zaron and de Carvalho, 2015; Quartly, 2019, Tran et al., 2019), leaving a substantial amount of low-frequency noise in the data. Low-frequency noise remaining in the signals after smoothing will still interfere with the precise measurement of geophysical peaks height, width, and position (O’Haver, 2019).

The EMD-based denoising method robustness stems from the self-consistent, data-driven approach, and from the ensemble average computed as the resulting denoised signal. The method developed for processing altimeter data benefited from previous optimization work and sensitivity studies carried out by Flandrin et al. (2004) and Kopsinis and McLaughlin (2009). Since the EMD-based denoising method does not contain an implicit smoothing operator, sparse data may still appear noisy in particular when the data editing performed before denoising did not identify large errors. It can occur mainly when altimeter waveforms are corrupted by rain, surface slicks or very light winds generating sigma0 blooms, or during transitions between land, sea ice and sea. However, the use of more efficient data editing and retracking algorithms should mitigate these errors (e.g., Passaro et al., 2014; Thibault et al., 2017).

As reported, signals down to scales of nearly 30 km can be recovered, provided that the signal-to-noise ratio is sufficient. This is particularly the case in the presence of very strong gradients and in main dynamic and/or eddy-rich regions. Although recent studies (Ardhuin et al., 2017, Quilfen and Chapron, 2019) reveal that wave–current interactions can trigger a significant increase in energy below the 100-km scale, very little is known and very few measurements are available for this scale of sea state variability. Until now, the sea state variability at scales below 100 km was often considered to be noise only (Zaron and de Carvalho, 2016).

Analysis of Jason-2, Cryosat-2, and Saral/AltiKa significant wave height power spectral density and of the spatial distribution of variance at short scales shows that the proposed method gives consistent results. As a newly developed tool for altimetry data, the EMD-based system could certainly require further testing and optimization, in particular for better screening of outliers that can shape the energy spectrum and bias noise statistics. The proposed method is a step forward to better exploit this unique

set of sea state observations, which now covers more than 25 years, and is easily applicable to other altimetric observations, such as sea level and radar cross-section.

### Acknowledgments

This study was conducted within the Ocean Surface Topography Science Team (OSTST) activities. A grant was awarded to the SILLAGE project by the TOSCA board in the framework of the CNES/EUMETSAT call CNES-DSP/OT 12-2118. The authors thanks Y. Kopsinis for having made the EMD-denoising software available (visit <http://cgi.di.uoa.gr/~kopsinis/software.html>); the version modified to process altimeter data can also be obtained upon request. The authors are grateful to the two anonymous reviewers who helped to improve the manuscript. Altimeter data were provided by Institut Français de Recherche pour l'Exploitation de la MER (IFREMER) (<ftp://ftp.ifremer.fr/ifremer/cersat/products/swath/altimeters/waves/>).

### References

- Ardhuin, F., Gille, S.T., Menemenlis, D., Rocha, C.B., Rasche, N., Chapron, B., Gula, J., Molemaker, J. (2017). Small-scale open-ocean currents have large effects on wind-wave heights. *Journal of Geophysical Research: Oceans*, **122**, 4500-4517, <https://doi.org/10.1029/2016JC012413>
- Dibarboure, G., Boy, F., Desjonquieres, J.D., Labroue, S., Lasne, Y., Picot, N., Poisson, J.C., Thibault, J.P. (2014). Investigating short-wavelength correlated errors on low-resolution mode altimetry. *Journal of Atmospheric and Oceanic Technology*, **31**, 1337-1362, <https://doi.org/10.1175/JTECH-D-13-00081.1>

Dufau, C., M. Orszynowicz, G. Dibarbouré, R. Morrow, and P.Y. Le Traon (2016). Mesoscale resolution capability of altimetry: Present and future. *Journal of Geophysical Research: Oceans*, **121**, 4910-4927, <https://doi.org/10.1002/2015JC010904>

Emery, W.J., Thomson, E. (2014). Data analysis methods in physical oceanography (3rd edition). Amsterdam, Netherlands: Elsevier B.V., 728pp, <https://doi.org/10.2307/1353059>

Flandrin, P., Rilling, G., Goncalves, P. (2004). Empirical mode decomposition as a filter bank. *IEEE Signal Processing Letters*, **11**, 112-114, <https://doi.org/10.1109/LSP.2003.821662>

Hadjihosseini, A., Peinke J., Hoffmann, N.P. (2014), Stochastic analysis of ocean wave states with and without rogue waves. *New Journal of Physics*, **16(5)**, 053037, <https://doi.org/10.1088/1367-2630/16/5/053037>

Huang, N.E., Shen, Z., Long, S.R., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society London*, **454**, 903-993, <https://doi.org/10.1098/rspa.1998.0193>

Huang, N.E., Wu, Z. (2008). A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Reviews of Geophysics*, **46**, RG2006, <https://doi.org/10.1029/2007RG000228>

Huang, H., Cressie, N. A. (2000). Deterministic/stochastic wavelet decomposition for recovery of signal from noisy data. *Technometrics*, **42**, 262-276.

Johnstone, I. M., Silverman, B.W. (1997). Wavelet Threshold Estimators for Data with Correlated Noise. *Journal of the Royal Statistical Society. Series B (Methodological)*, **59**, pp. 319–351.

Kopsinis, Y., McLaughlin, S. (2009). Development of EMD-based denoising methods inspired by wavelet thresholding. *IEEE Transactions on Signal Processing*, **57**, 1351-1362,

<https://doi.org/10.1109/TSP.2009.2013885>

Mallat, S. (2009). A wavelet tour of signal processing. London: Academic Press

<https://doi.org/10.1016/B978-0-12-374370-1.X0001-8>

O'Haver, T. A. (2019). Pragmatic Introduction to Signal Processing. Available online at

<https://terpconnect.umd.edu/~toh/spectrum/IntroToSignalProcessing2019.pdf>

Passaro, M., Cipollini, P., Vignudelli, S., Quartly, G., Snaith, H. (2014). ALES: A multi-mission subwaveform retracker for coastal and open ocean altimetry. *Remote Sensing of the Environment* , **145**, 173-189, <https://doi.org/10.1016/j.rse.2014.02.008>.

Quartly, G. (2019). Removal of covariant errors from altimetric wave height data. *Remote Sensing*,

**11(19)**, 2319, <https://doi.org/10.3390/rs11192319>

Queffeuilou P., Croizé-Fillon, D. (2013). Global altimeter Hs data set, version 11. Technical report, Laboratoire d'Océanographie Spatiale, IFREMER. Available online:

[ftp://ftp.ifremer.fr/ifremer/cersat/products/swath/altimeters/waves/documentation/altimeter\\_wave\\_merge\\_\\_11.4.pdf](ftp://ftp.ifremer.fr/ifremer/cersat/products/swath/altimeters/waves/documentation/altimeter_wave_merge__11.4.pdf)

Quilfen, Y., Yurovskaya, M., Chapron, B., Ardhuin, F. (2018). Storm waves focusing and steepening in the Agulhas current: satellite observations and modeling. *Remote Sensing of Environment*, **216**,

561-576, <https://doi.org/10.1016/j.rse.2018.07.020>

Quilfen Y., Chapron B.(2019). Ocean surface wave-current signatures from satellite altimeter measurements . *Geophysical Research Letters* , **46**, 253-261 . <https://doi.org/10.1029/2018GL081029>

Sandwell, D.T., Smith, W.H.F. (2005). Retracking ERS-1 altimeter waveforms for optimal gravity field recovery. *Geophysical Journal International*, **163**, 79-89, <https://doi.org/10.1111/j.1365-246X.2005.02724.x>

Taburet, N., Husson, R., Charles, E. (2019). WAVE Thematic Assembly Centre Quality Information Document. CMEMS-WAVE-QUID-014-001, 56 pp. Available online: [http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com\\_csw&view=details&product\\_id=WAVE\\_GLO\\_WAV\\_L3\\_SWH\\_NRT\\_OBSERVATIONS\\_014\\_001](http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=WAVE_GLO_WAV_L3_SWH_NRT_OBSERVATIONS_014_001)

Taburet, N., SL-TAC team (2018). Sea Level TAC-DUACS Quality Information Document. CMEMS-WAVE-QUID-014-001, 66 pp. Available online : [http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com\\_csw&view=details&product\\_id=SEALEVEL\\_GLO\\_PHY\\_L3\\_NRT\\_OBSERVATIONS\\_008\\_044](http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=SEALEVEL_GLO_PHY_L3_NRT_OBSERVATIONS_008_044)

Thibault, P., Poisson, J.C., Bronner, E., Picot, N. (2010). Relative Performance of the MLE3 and MLE4 Retracking Algorithms on Jason-2 Altimeter Waveforms. *Marine Geodesy*, **33**, 317-335, <https://doi.org/10.1080/01490419.2010.491033>

Thibault, J.P., Piras, F., Poisson, J.C., et al. (2017). Convergent solutions for retracking conventional and Delay Doppler altimeter echoes. Presented at Ocean Surface Topography Science Team Meeting, Miami, FL, USA, 23–27 October; Available online: <https://tinyurl.com/ybppndx6>

Tournadre, J., Chapron, B., Reul, N., Vandemark, D. C. (2006). A satellite altimeter model for ocean slick detection, *Journal of Geophysical Research*, **111**, C04004,

<https://doi.org/10.1029/2005JC003109>

Tran, N., Vandemark, D.C., Zaron, E.D., Thibaut, P., Dibarboure, G., Picot, N. (2019). Assessing the effects of sea-state related errors on the precision of high-rate Jason-3 altimeter sea level data. In press in *Advances in Space Research: “25 Years of Altimetry”* special issue

Wu, Z., Huang, N.E. (2004). A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society London*, **460**, 1597-1611,

<https://doi.org/10.1098/rspa.2003.1221>

Xu, Y., Fu, L.L. (2012). Global variability of the wavenumber spectrum of oceanic mesoscale turbulence. *Journal of Physical Oceanography*, **41**, 802–809,

<https://doi.org/10.1175/2010JPO4558.1>

Zaron, E.D., De Carvalho, R. (2016). Identification and reduction of retracker-related noise in altimeter-derived sea surface height measurements. *Journal of Atmospheric and Oceanic Technology*,

**33**, 201-210, <https://doi.org/10.1175/JTECH-D-15-0164.1>

**Declaration of interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: