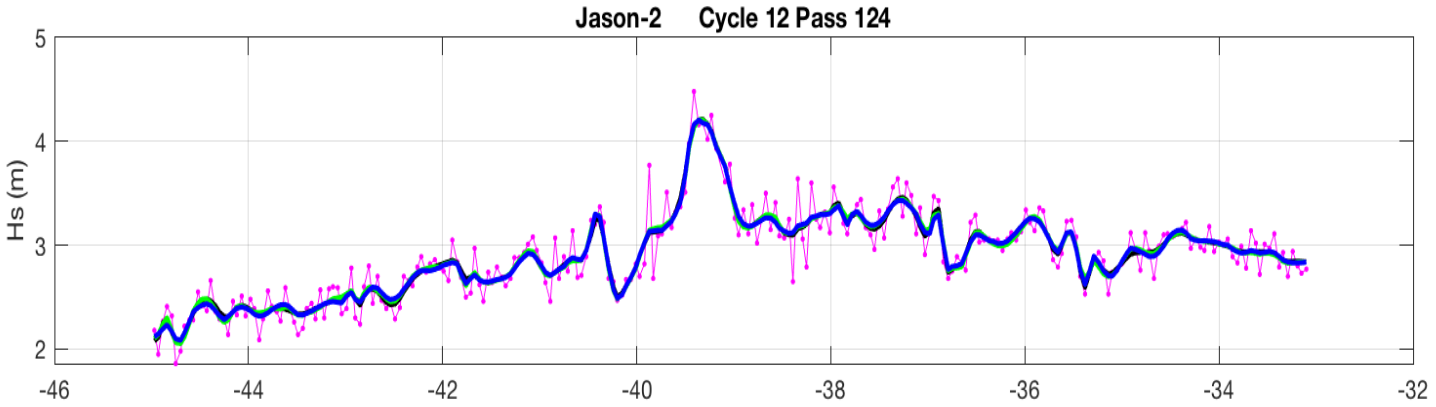
On denoising satellite altimeter measurements for geophysical signal analysis

B. Chapron, Y. Quilfen, G. Dodet



✓ Hs and SSH measurements are significantly contaminated at scales < 150 km

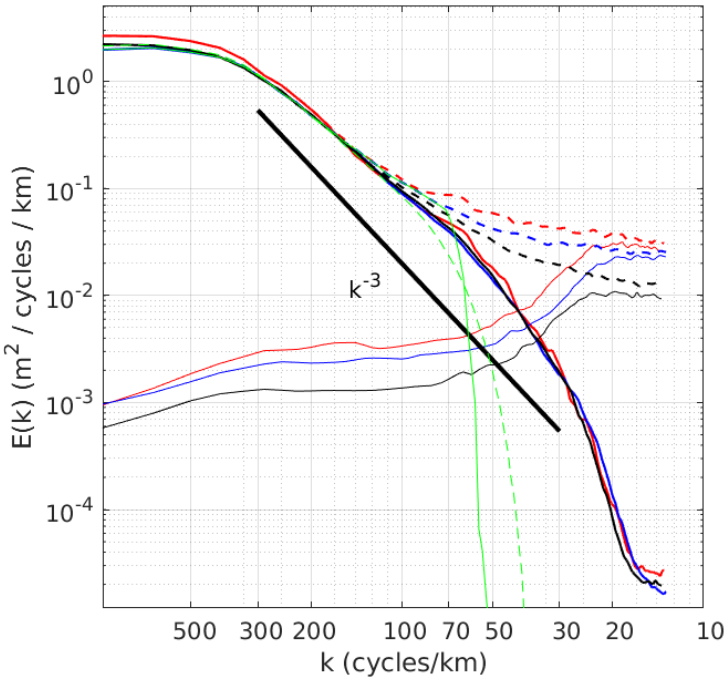
✓ Large scale currents and eddies are a major source of sea state variations at these scales, down to ~ 10 km (Ardhuin et al., 2017; Quilfen et al., 2018, 2019)

✓ Smoothing filters, such as running mean, Loess or Lanczos (the one used to process the actual CMEMS products) show similar limitations:

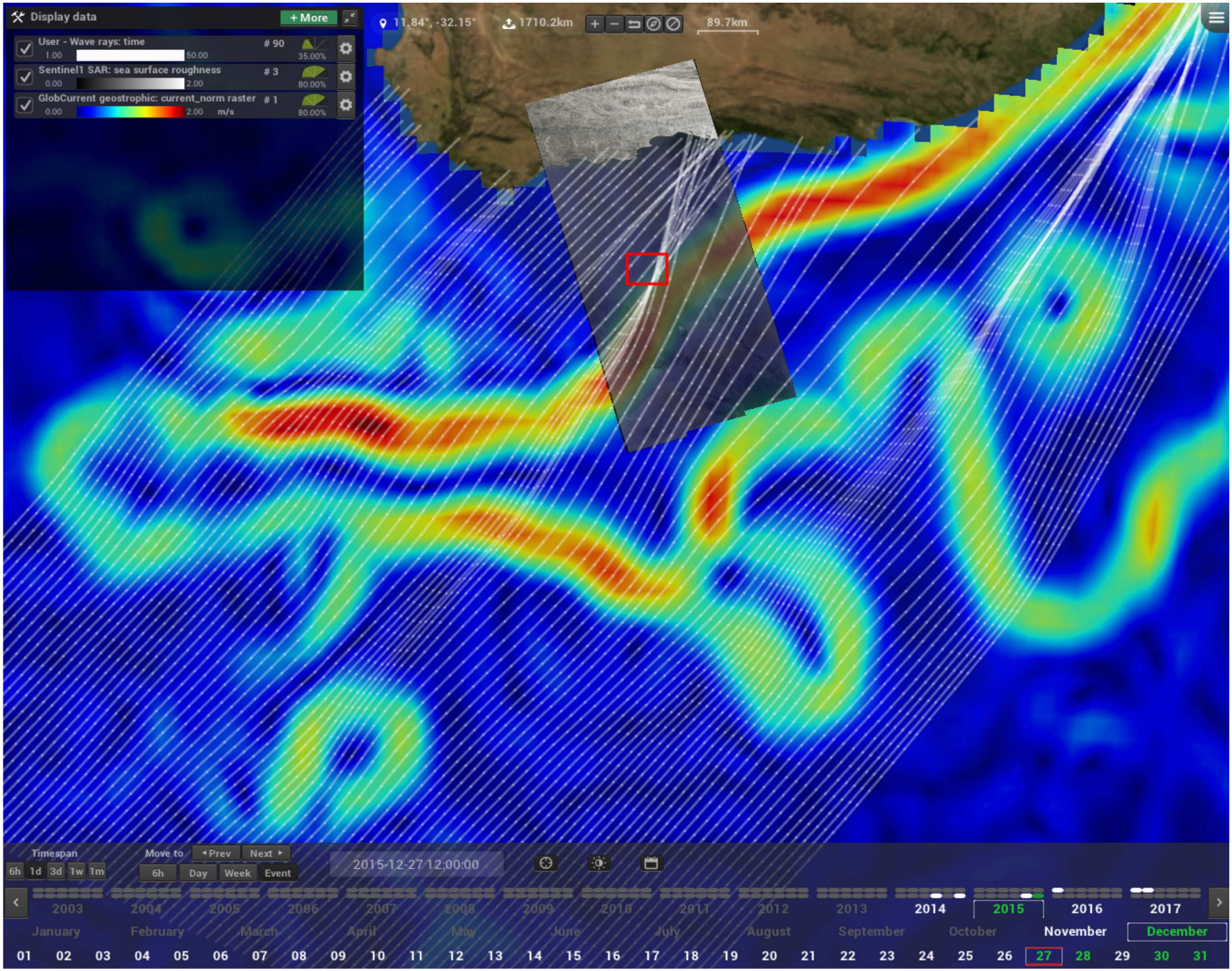
- difficult and inadequate settings of the filter length or cut-off frequency at global scale (lack of data adaptive property)
- spectral ringing or excessive smoothing (a trade-off for the Lanczos filter)
- unphysical representation of the variability spectrum

✓ A non parametric approach, the EMD-based denoising, has been successfully adapted for altimetry data to overcome such drawbacks and help mapping overlooked signals (such as related to wave / current interactions) in low SNR environment

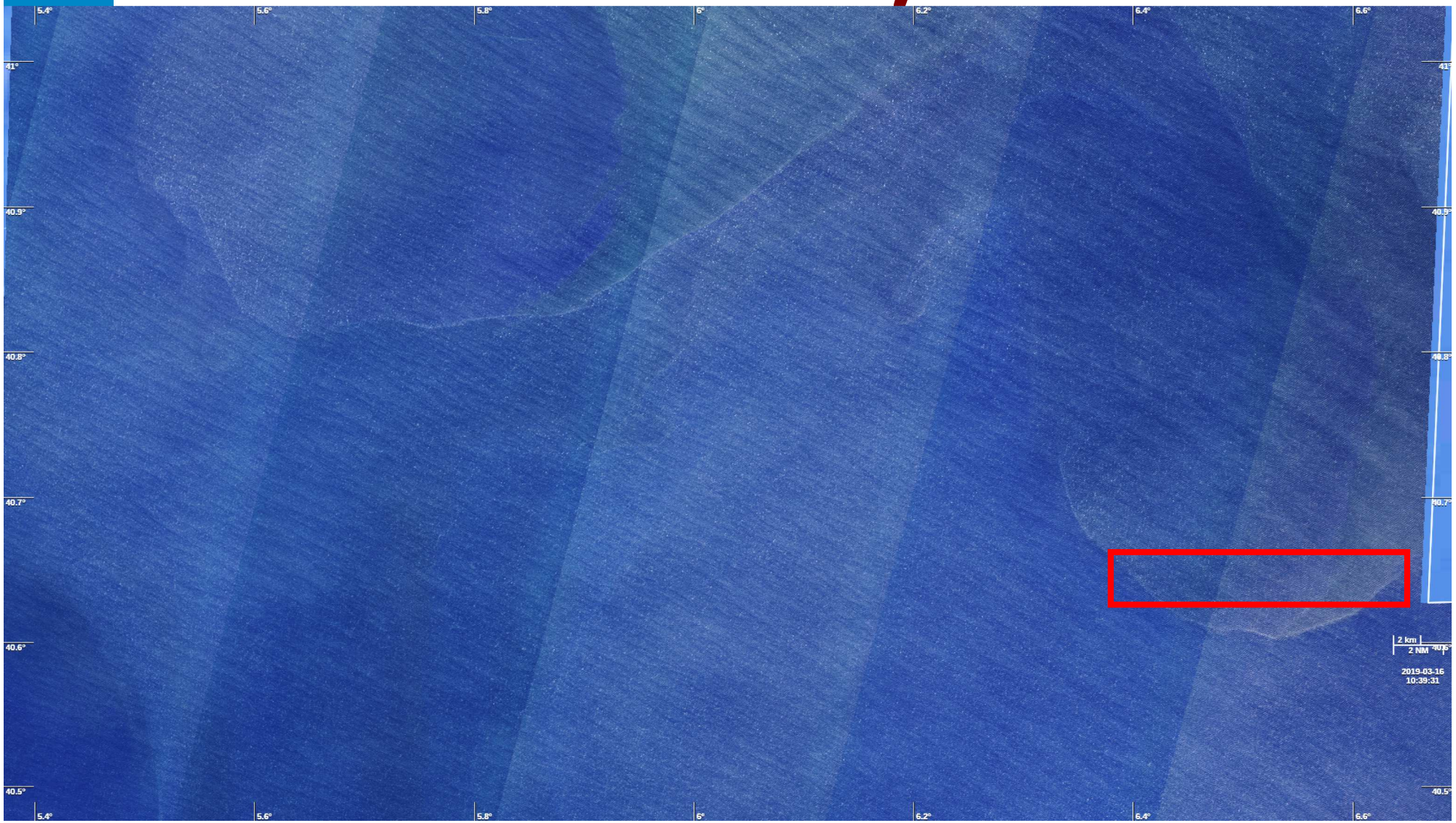
✓ It is actually operated to provide a denoised Hs parameter and associated uncertainty in the sea state CCI actual products and upcoming CMEMS products



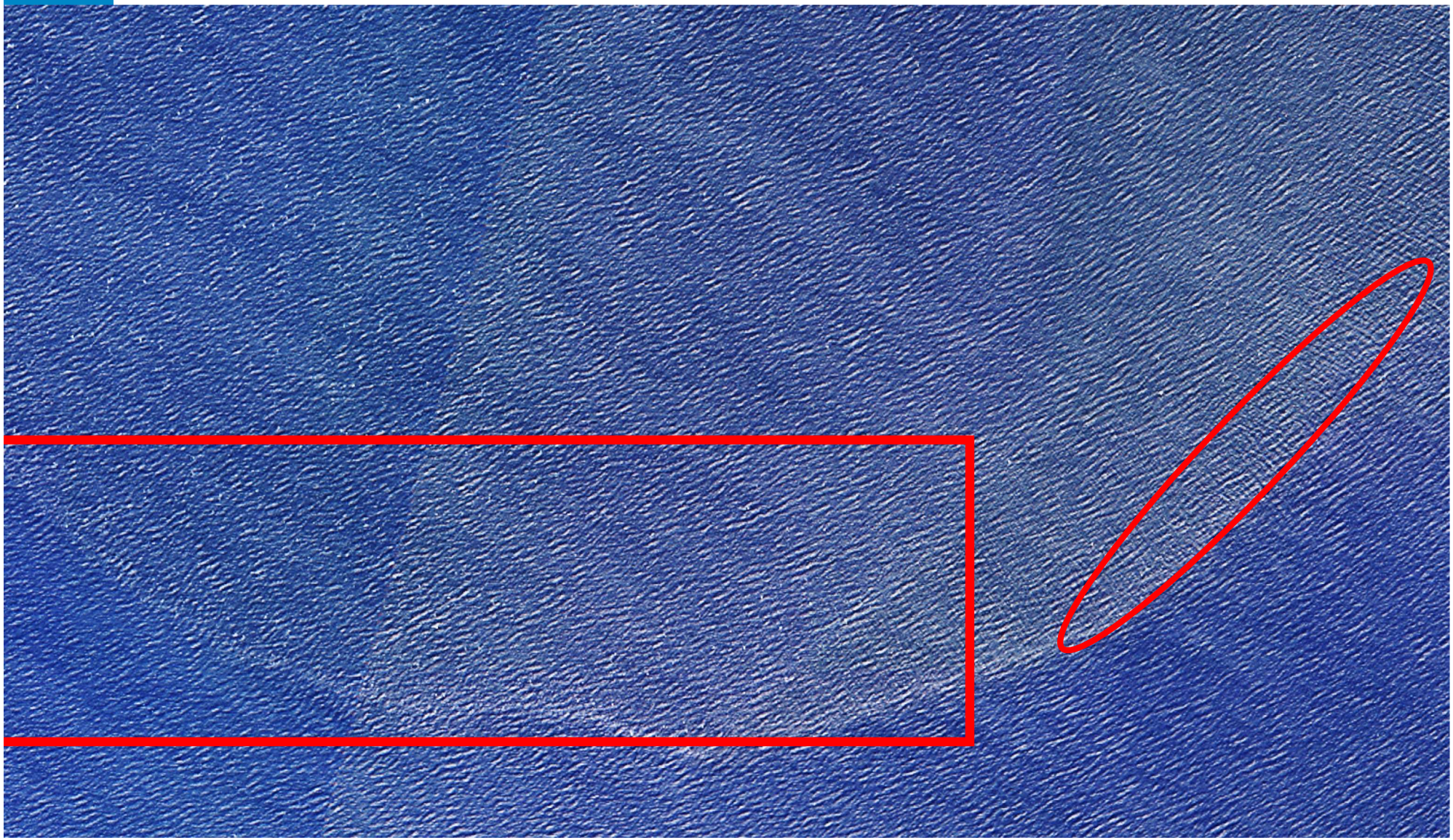
PSD of SWH for Saral (black), Jason-2 (blue), Cryosat-2 (red) noisy (dashed lines), denoised (thick solid lines), and IMF1 noise (thin solid lines) data in the Agulhas region. Green lines gives the PSD for a 65-km Lanczos filter.

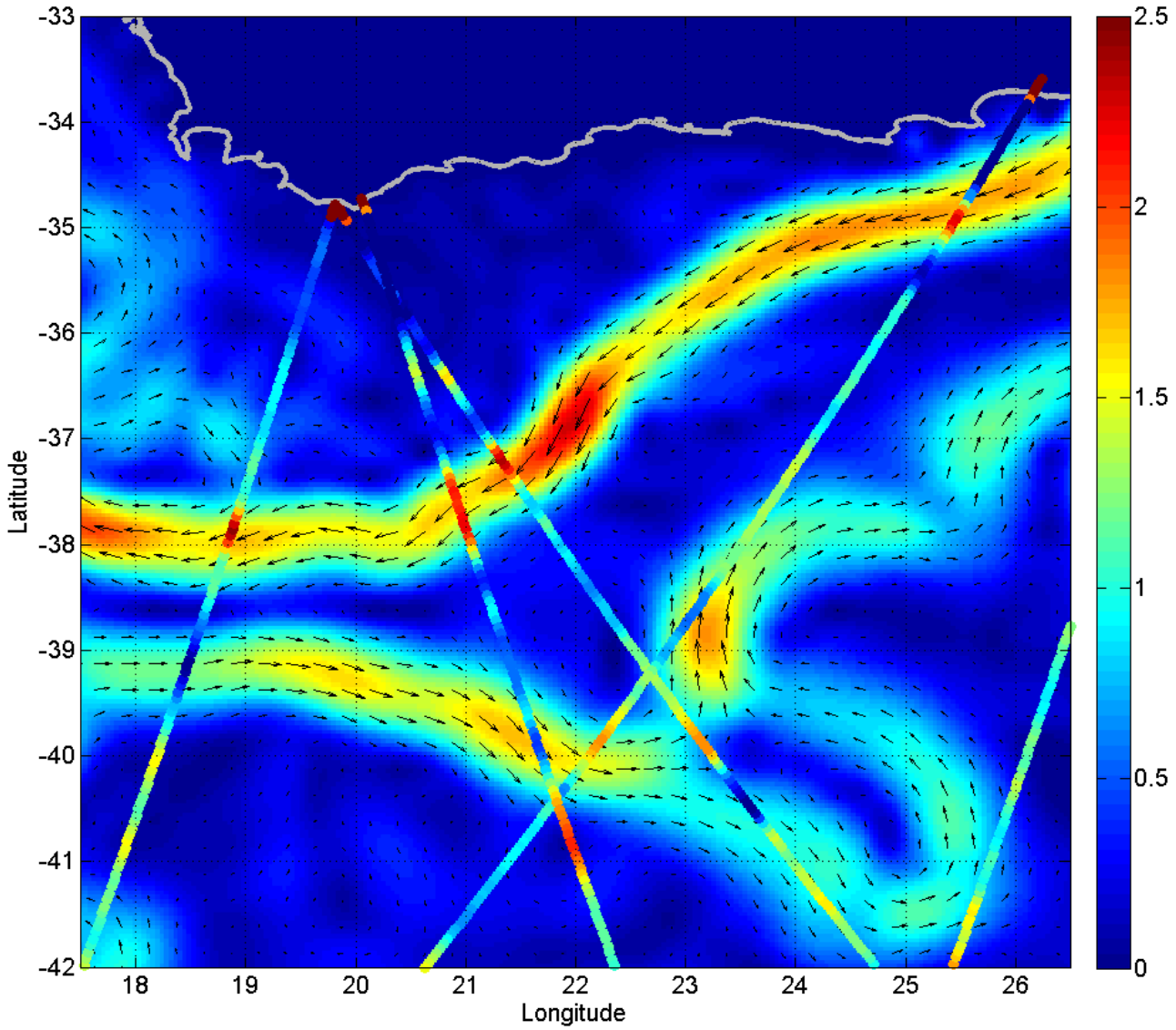

Display data
+ More
User - Wave rays: time
90
1.00
50.00
35.00%
Sentinel1 SAR: sea surface roughness
3
0.00
2.00
80.00%
GlobCurrent geostrophic: current_norm raster
1
0.00
2.00
m/s
80.00%
11.84°, -32.15°
1710.2km
89.7km
Timespan
6h 1d 3d 1w 1m
Move to
Prev
Next
6h Day Week Event
2015-12-27 12:00:00
2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
January February March April May June July August September October November December
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

Sentinel 2 spatial wave field variability



Sentinel 2 spatial wave field variability





Geostrophic surface current velocity corresponding to January, 4rth 2016, and SWH anomalies, , along the altimeter tracks, from a 250 km moving average along the altimeter track.

Empirical Mode Decomposition (EMD) denoising: key features and background

Key features

✓ The method is a self-consistent adaptive approach relying on a data-driven evaluation of the actual SNR

✓ It makes use of an empirical and intuitive decomposition (EMD) of the along-track signals, not requiring compliance with a mathematical framework (such as in Fourier or wavelet analysis)

✓ EMD is therefore suitable to decompose and denoise efficiently non-linear and non-stationary signals, preserving gradients and extreme values

✓ A local ensemble average approach gives robust denoised Hs and estimation of the uncertainty:

Hs_measured = Hs_geophys + measurements errors (white noise, uncorrelated along-track)
Hs_denoised = (Hs_geophys , Hs_uncertainty)

Background

Main references:

- Huang and co-authors (1998, 2008) for the development of the EMD at NASA for purpose of time / frequency analysis
- Flandrin and co-authors (2004) for studies at ENS on the EMD properties in case of Gaussian noise analysis
- Kopsinis and McLaughin (2009) for the development of the EMD-based denoising algorithm and its numerical evaluation

- Quilfen and Chapron (2019), On denoising satellite altimeter measurements for high resolution geophysical signal analysis. Submitted to Advances in Space Research

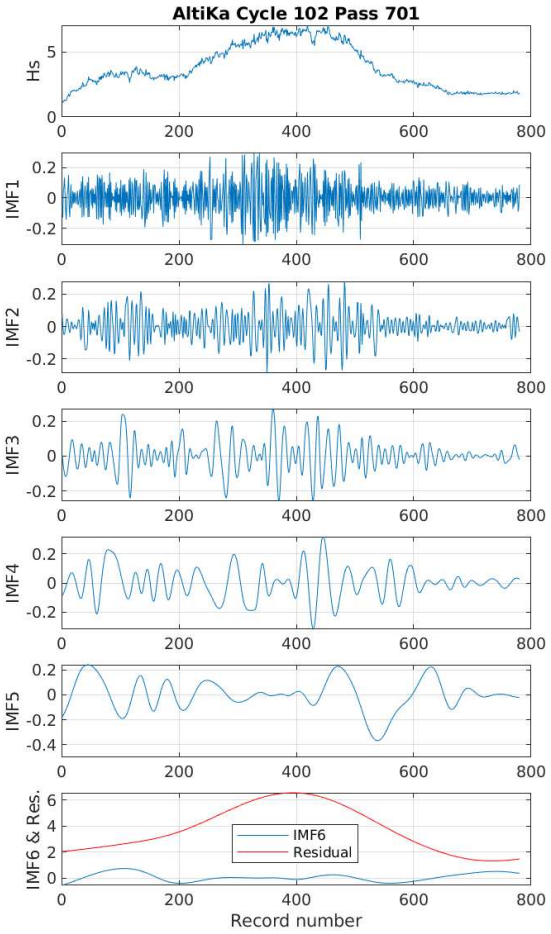
EMD algorithm

✓ EMD adaptively decomposes a signal *y(x)* into a small number L of Intrinsic Modulation Functions (IMF) so that:

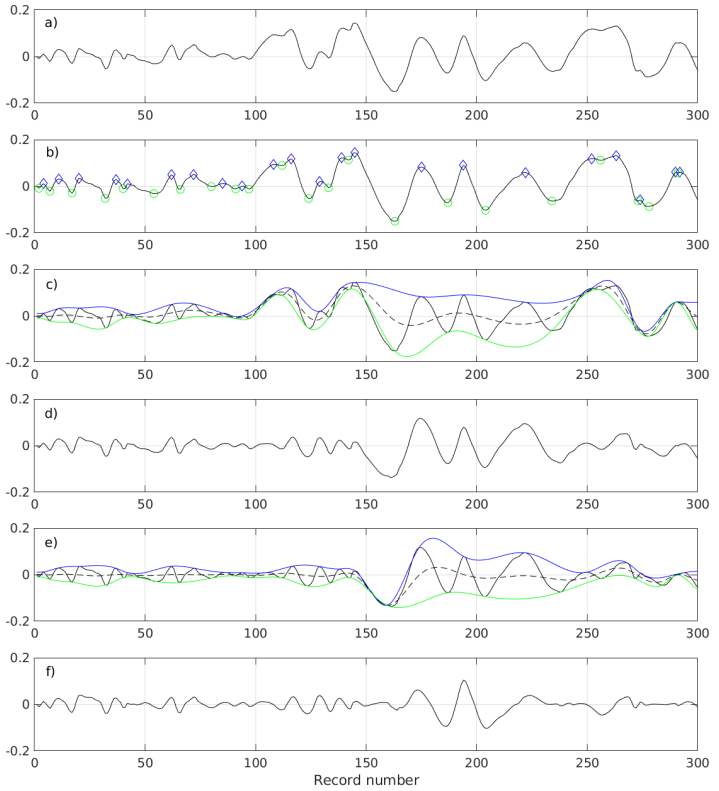
$$y(x) = \sum_{i=1}^{L} h^{(i)}(x) + d(x)$$ with *d(x)* the trend (or residue)

✓ By definition, IMFs $h^{(i)}(x)$ are AM / FM functions with zero mean (+ / - ε) and the same number of zero-crossings and extrema (+ / - 1)

✓ The IMFs are obtained by construction from the first containing the high-frequency modulations to the last representing the trend (or residue), through an iterative algorithm called "sifting process". For any given signal, the sifting iteratively removes the low-frequency modulation until the IMF satisfies the above definition



Example of Saral/AltiKa 1 Hz data and associated IMFs
6 IMFs plus the residual/trend accounts for the full signal

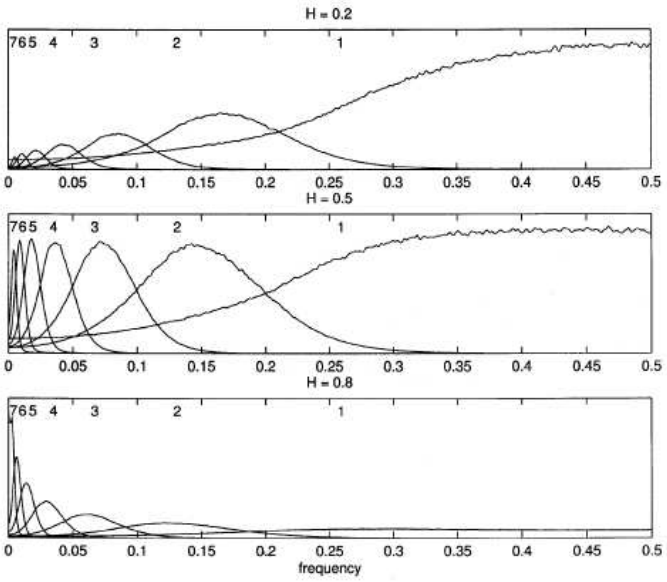


Piece of the sifting process: (a) any input signal; (b,c,d,e) two iterations showing extrema detection, spline-interpolated upper and lower envelops and removal of their mean; (f) the final IMF after convergence

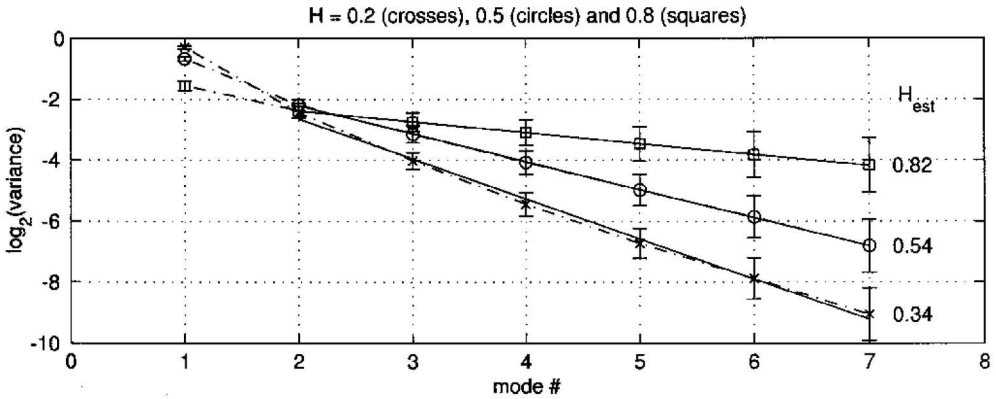
Gaussian noise properties in the EMD basis

- ✓ IMF's properties are analyzed by means of numerical simulations as they are not mathematical functions
- ✓ Flandrin et al. (2004) conducted Monte-Carlo simulations of a fractional Gaussian noise processed with EMD
- ✓ They found IMF1 is a high-pass filter and higher order IMFs form a dyadic filter bank, i.e. the occupied frequency band is halved for consecutive IMFs from lower to higher rank.
- ✓ IMFs spectra are self-similar, for n > 1, and the relationship between IMFs variances is predicted. For white noise $E_n \approx \frac{E_1}{0.719} 2^{-n}$
- ✓ Two features are core to the proposed denoising method:

1) In the EMD basis, the noise energy is decreasing quickly with IMF increasing rank: ~ 60, 20, 10, 5, 2.5 % of the total energy with increasing rank. The 4 first IMFs account for ~ 95% of the noise energy
2) As IMF1 contains mostly noise, *E1* estimation is easy to enable prediction of the expected noise energy in other IMFs, and setting of thresholds for the SNR testing in each IMF



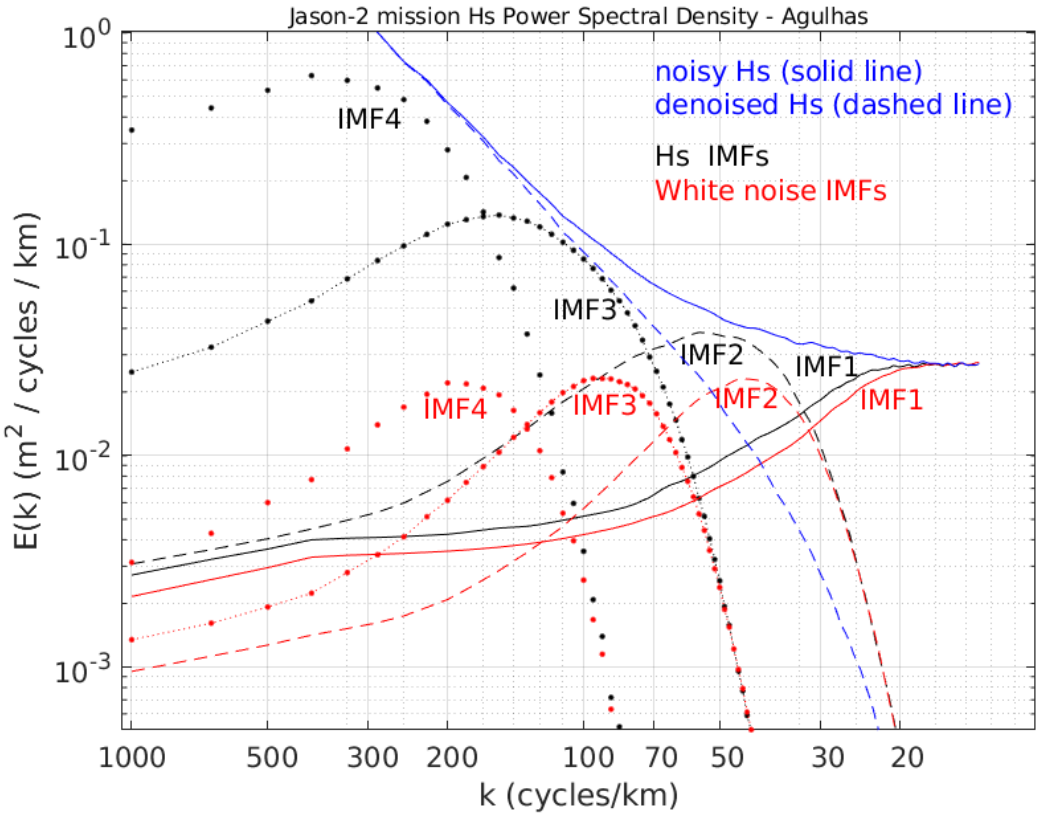
IMFs spectra for correlated (top and bottom) and uncorrelated (middle) Gaussian noise



IMFs variance (log2) distribution for correlated (upper and lower curves) and uncorrelated (middle curve) Gaussian noise

Gaussian noise and Hs spectral density in the EMD basis

- ✓ IMFs spectra for a white noise and 3-yrs of Jason-2 measurements are compared for the Agulhas region
- ✓ The noise standard deviation has been adjusted to fit the Hs noise floor at scales < 20 km
- ✓ Hs and noise IMFs spectra share the same frequency bands, as a result of the EMD filter bank. Hs IMFs can be tested against noise IMFs
- ✓ Hs IMF1 behaves as white noise IMF1 and contains mostly noise, as expected, with slightly increased energy at scales > 20 km (along track modulation of noise energy and large geophysical gradients)
- ✓ Hs IMF2 and IMF3 show comparable energy at 70km scale, below which IMF2 energy dominates
- ✓ On average, IMF2 is noise only at 30km scale and the SNR is ~ 0.5 at 50km scale, meaning that information can be recovered locally.



EMD denoising algorithm principles

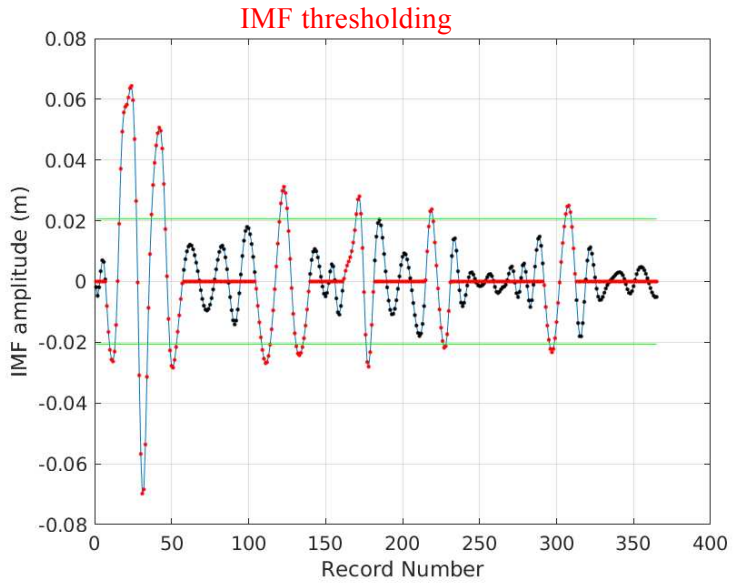
EMD denoising main steps applied to continuous along-track 1 Hz data segments (with a minimum length of ~ 50 kms) :

- ✓ 1) Perform a first EMD expansion of $y(x)$ and denoise IMF1 to obtain the IMF1 noise series $\varepsilon(x)$ and associated energy reference level E_1 (done with a wavelet denoising approach)
- ✓ 2) Compute $y(x) - \varepsilon(x)$ and generate an ensemble of k noisy series $y_k(x)$ by adding random permutations of $\varepsilon(x)$ to $y(x) - \varepsilon(x)$
- ✓ 3) For each $y_k(x)$: perform the EMD expansion, threshold each IMF using thresholds Tn, sum the denoised IMFs to get the denoised $y_{kd}(x)$

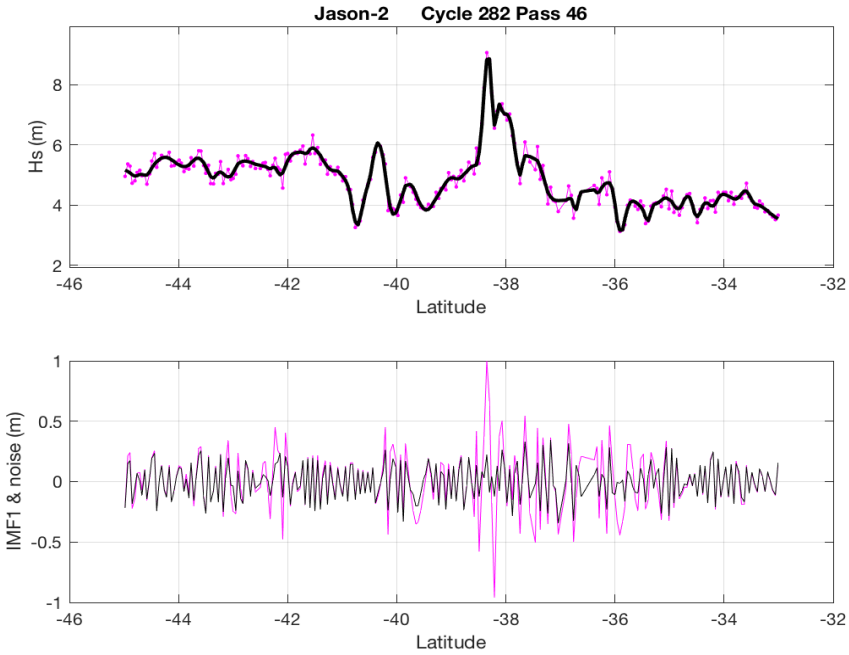
  $T_1 = C * E_1$ and $T_n = C * E_n = C * \frac{E_1}{0.719} 2^{-n}$ for n > 1, with C a constant to be adjusted globally (fine tuning)

- ✓ 4) Compute the ensemble average and standard deviation of $y_{kd}(x)$ to obtain the resulting denoised signal and associated uncertainty $y_d(x), \epsilon(x)$

The ensemble average increases robustness of the denoised signal, mitigating for noise and process uncertainties



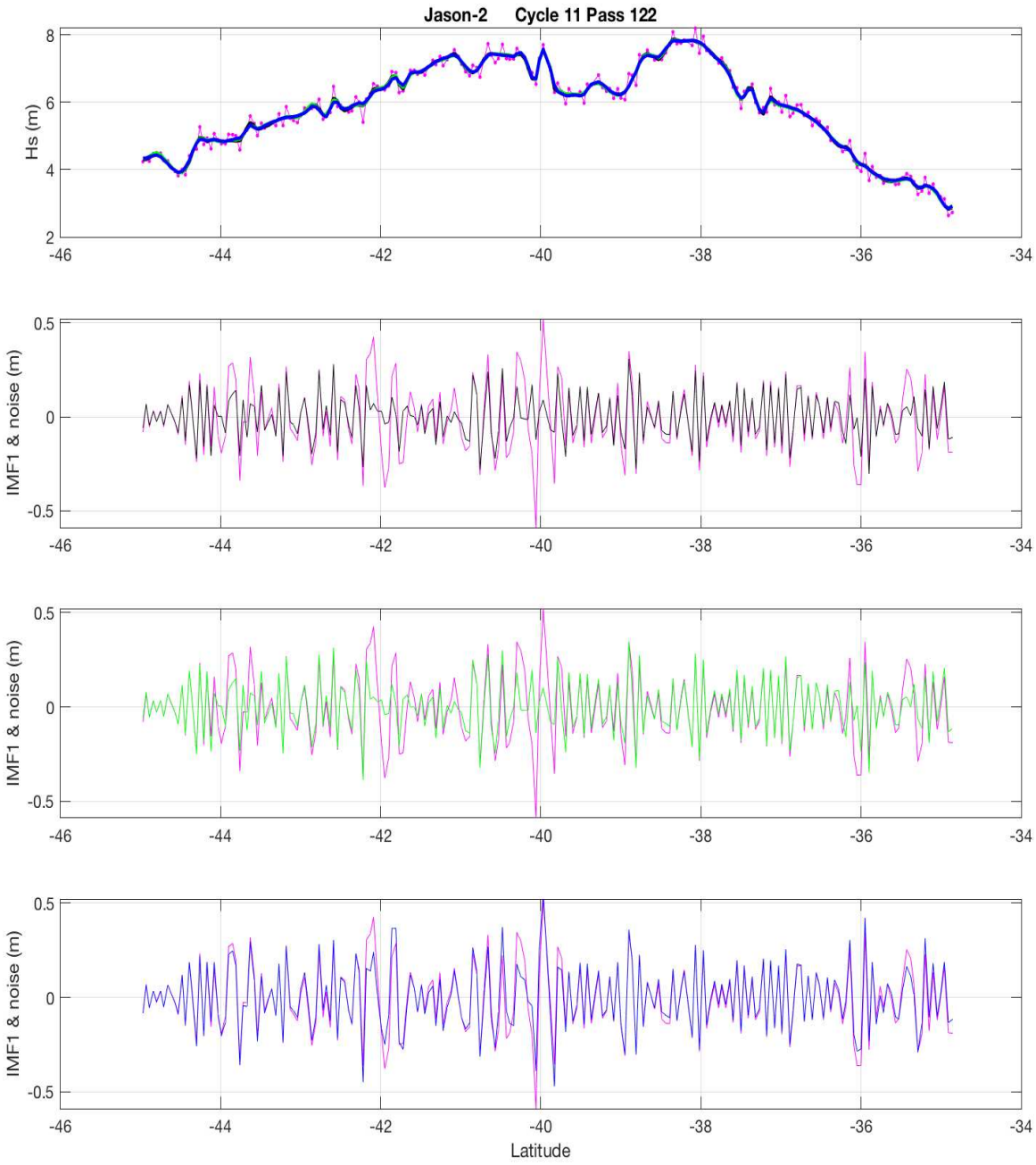
IMF amplitude (blue), canceled modulations (black), recovered useful signal (red), thresholds (green)



Significant wave height noisy and denoised series (top) and IMF1 amplitude and noise series (bottom)

Ensemble average sensitivity to noise and prescribed threshold

Robustness to varying statistics and unexpected behavior of noise (re-tracking errors) and strong geophysical signatures



The three curves obtained with different noise configurations almost coincide

The IMF1 noise $\varepsilon(x)$ is evaluated through wavelet analysis, with prescription of the first wavelet level to be tested for signal extraction

Noise level $E_1 = \left(\frac{median|\varepsilon(x)|}{0.6745}\right)^2$

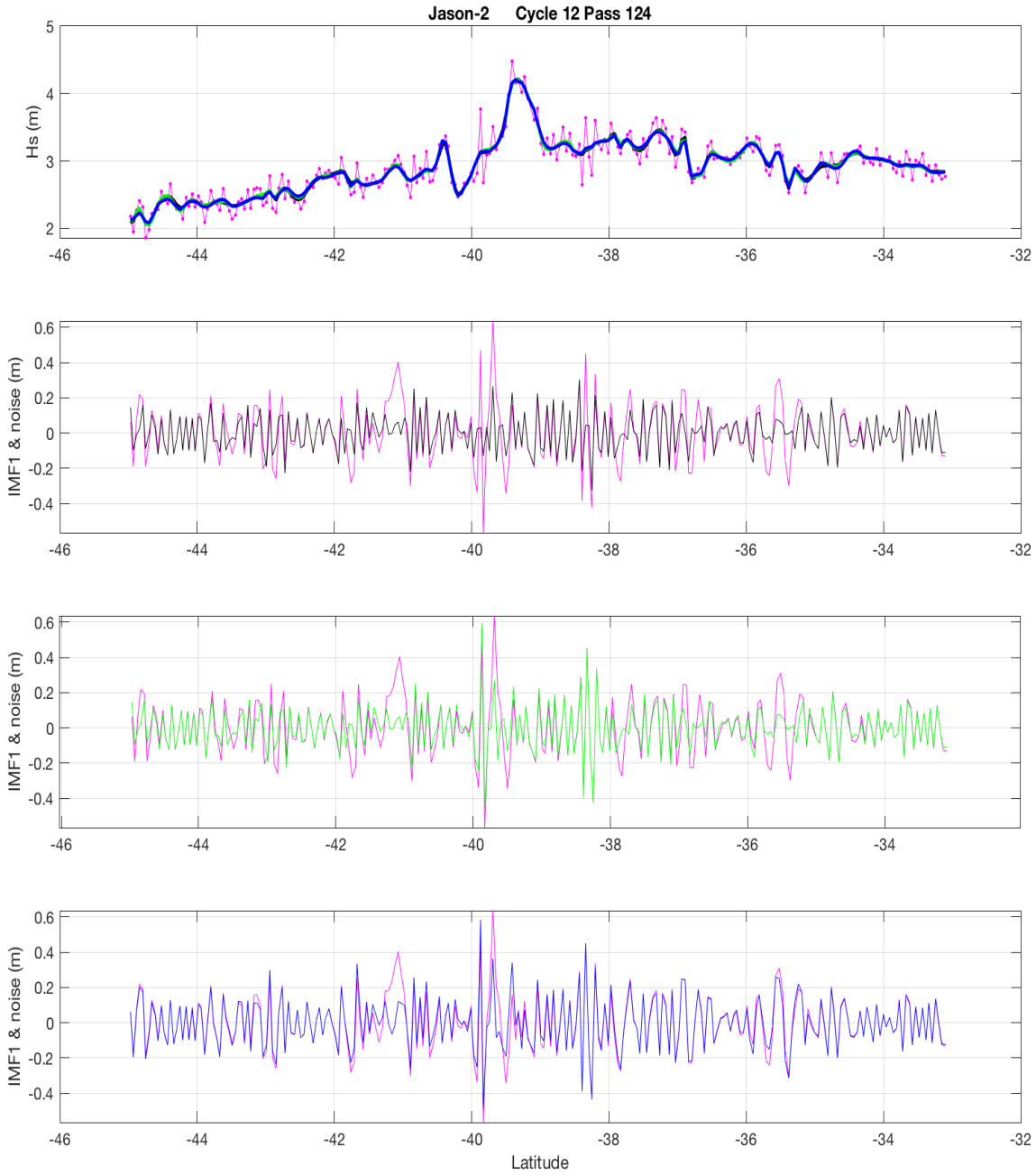
Signal tested starting at 1st level (finest details can contain useful signal, as implemented in the processor). *E1 = 0.0185*

Signal tested starting at 2nd level (first level is noise only). *E1 = 0.0230*

Signal tested starting at 3rd level (two first levels are noise only). *E1 = 0.0349*

Ensemble average sensitivity to noise and prescribed threshold

Robustness to varying statistics and unexpected behavior of noise (re-tracking errors) and strong geophysical signatures



The three curves obtained with different noise configurations almost coincide

The IMF1 noise $\varepsilon(x)$ is evaluated through wavelet analysis, with prescription of the first wavelet level to be tested for signal extraction

Noise level $E_1 = \left(\frac{median|\varepsilon(x)|}{0.6745}\right)^2$

Signal tested starting at 1st level (finest details can contain useful signal, as implemented in the processor). *E1 = 0.0171*

Signal tested starting at 2nd level (first level is noise only). *E1 = 0.0173*

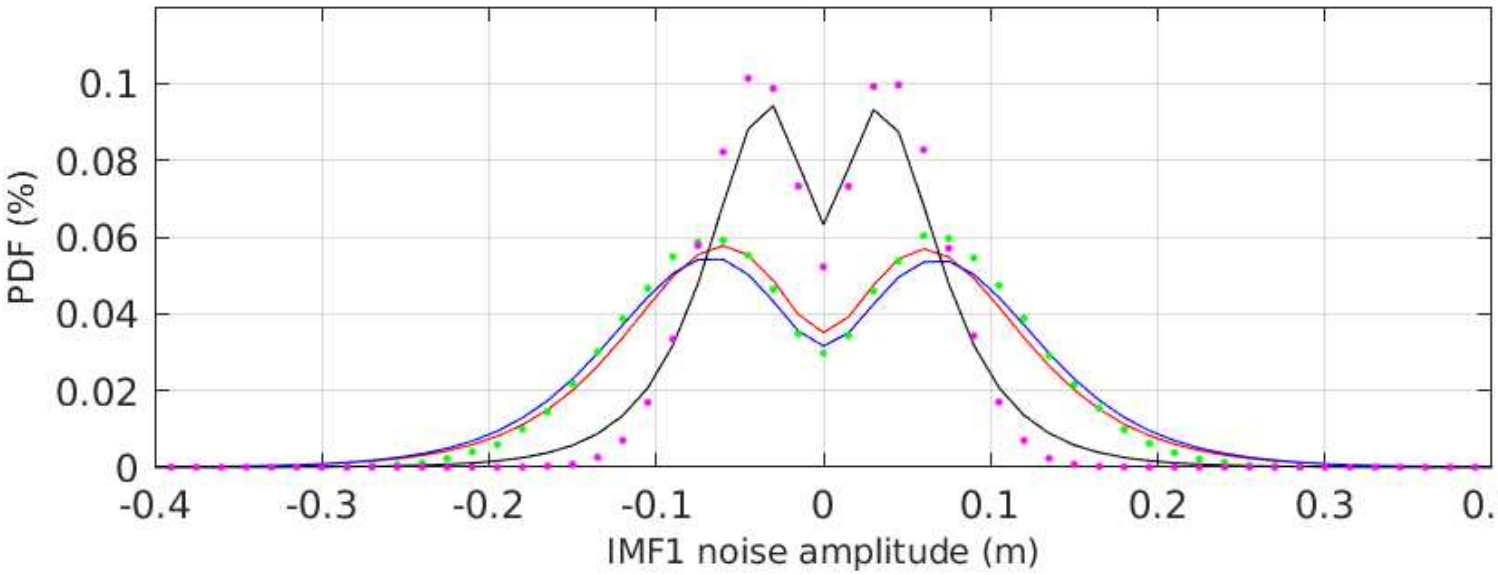
Signal tested starting at 3rd level (two first levels are noise only). *E1 = 0.0244*

Estimated noise vs white noise statistics

✓ The estimated noise statistics are verified through Monte-Carlo simulations

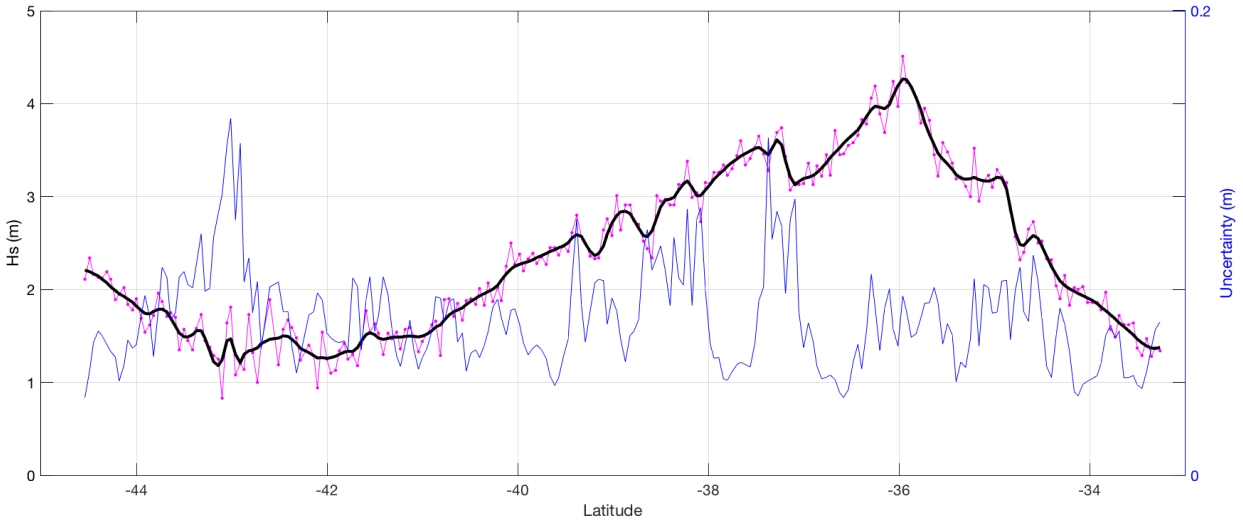
IMF1 noise amplitude PDF well compares with that of a white noise of 7 cm standard deviation for Saral, and 12 cm for Jason-2 and Cryosat-2 (in agreement with other studies such as cross-over analysis)

✓ Noise monitoring can be performed this way



IMF1 noise amplitude PDF for Saral (black), Jason-2 (blue), Cryosat-2 (red); and Monte Carlo simulations of IMF1 for a white noise of 7cm (green dots) and 12cm (magenta dots) standard deviation

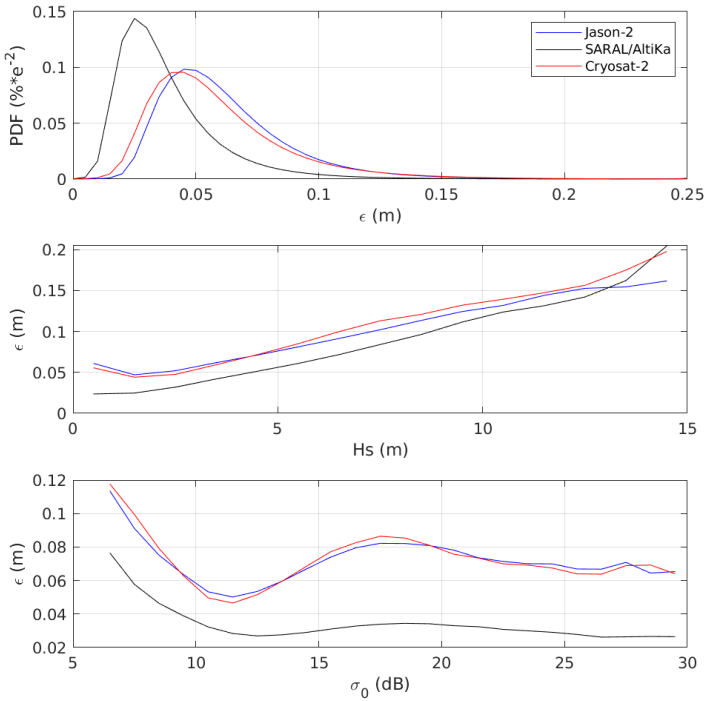
Distribution of the uncertainty associated with denoised Hs



Raw (magenta) and denoised (black) Hs, with associated uncertainty (blue)

Uncertainty $\epsilon(x)$ (associated with denoised Hs) ≈ whole system uncertainty + stochastic variability

- ✓ The system uncertainty is related to the local SNR and EMD filter errors (e.g. uncertainties related to prescribed thresholds).
- ✓ Lower the SNR, larger $\epsilon(x)$. It then depends on each altimeter, on the analyzed Hs, and on the local resolved scales ($\epsilon(x)$ larger at shorter scales)
- ✓ It accounts for errors related to slicks and rainfall, events for which re-tracking errors are larger, occurring mainly in low winds areas. $\epsilon(x)$ apparently increases more at low winds for Jason-2 and Cryosat-2 because less data are discarded by the data editing

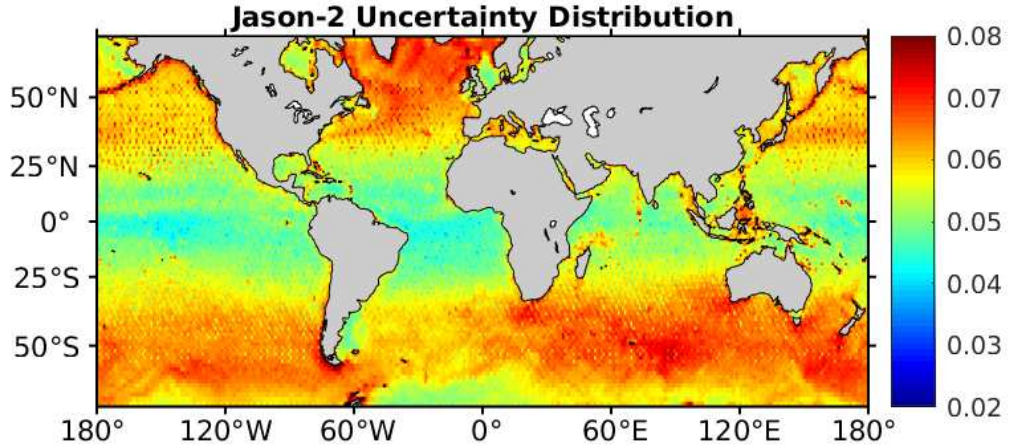


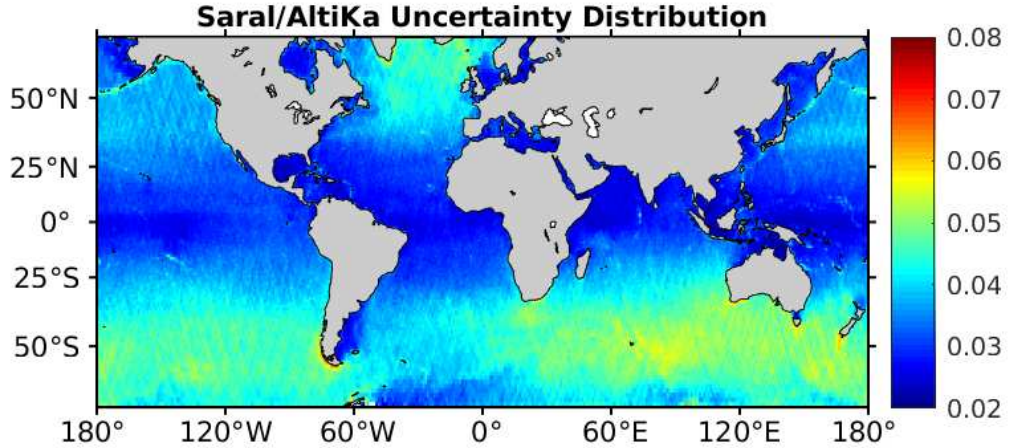
(top) PDF of uncertainty ϵ for Saral (black), Jason-2 (blue), Cryosat-2 (red);
(middle) ϵ as a function of significant wave height
(bottom) ϵ as a function of the radar cross section

Distribution of the uncertainty associated with denoised Hs

Uncertainty $\epsilon(x)$ (associated with denoised Hs) $\approx$ whole system uncertainty + stochastic variability

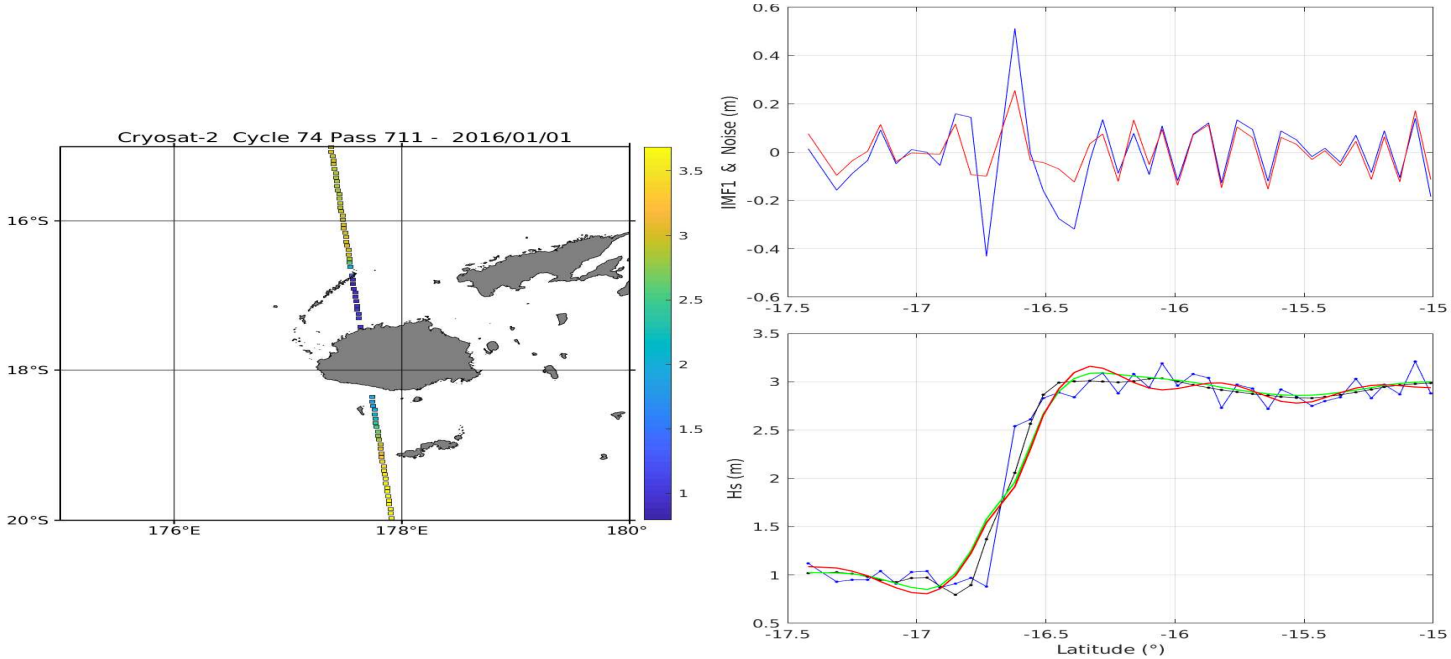
✓ The system uncertainty is related to the local SNR and EMD filter errors (e.g. uncertainties related to prescribed thresholds).

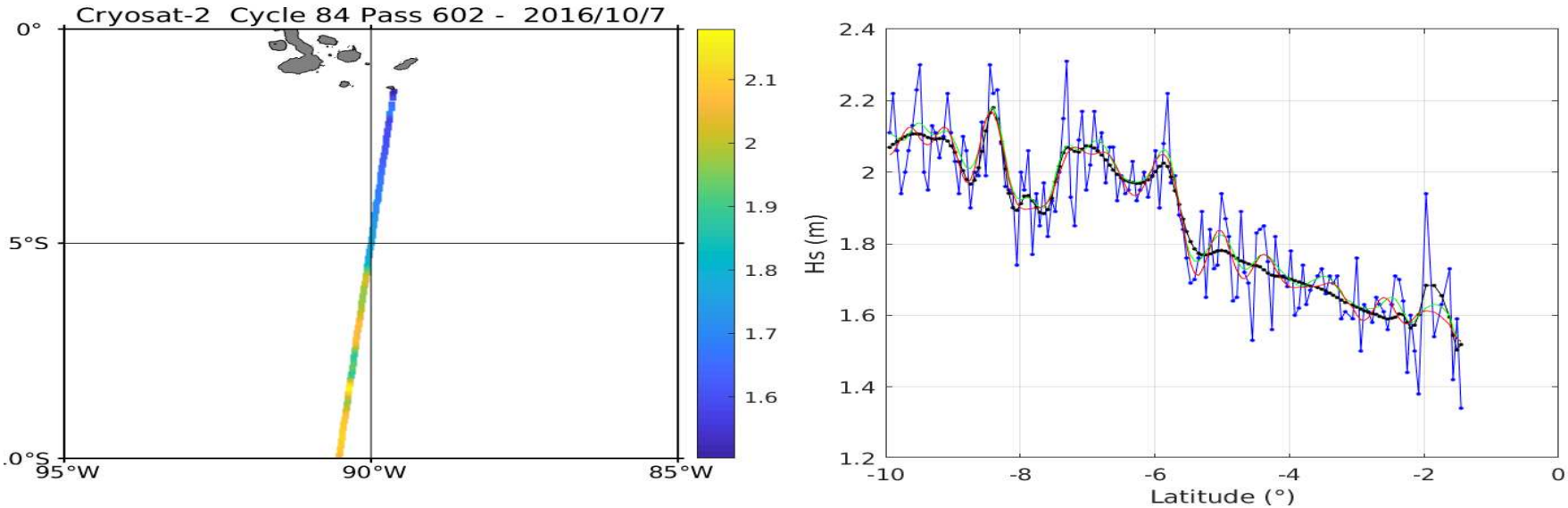




Examples

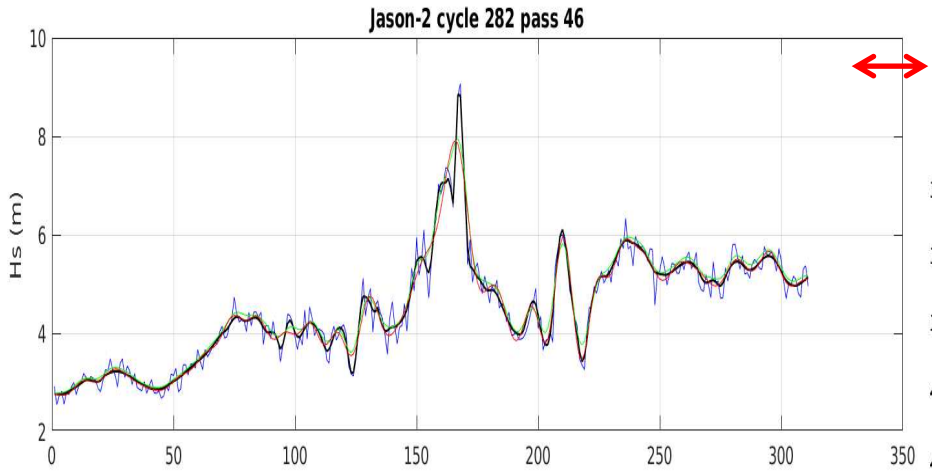
EMD-denoising (black curve) preserves better large gradients while Lanczos filter (60 km cut-off) performs greater gradient smoothing although overshoot ripples still exist whatever the applied roll-off, smooth (green curve) or sharp (red curve)



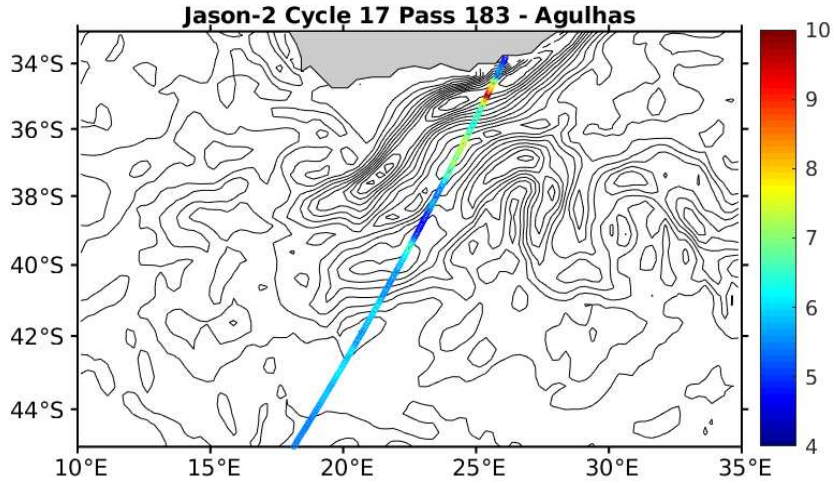


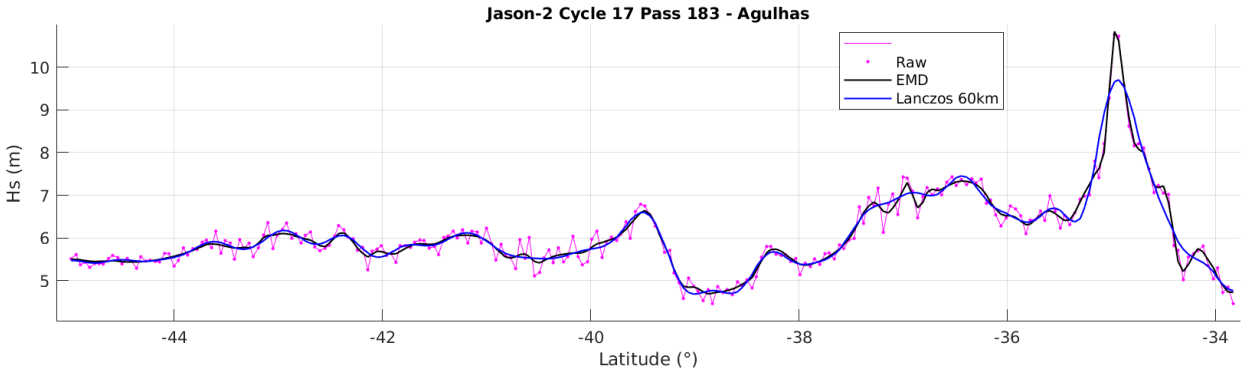
EMD-denoising (black curve) accounts for a too low SNR, to give a smooth signal while the Lanczos filter produces artifact oscillations whatever the applied roll-off, smooth (green curve) or sharp (red curve)

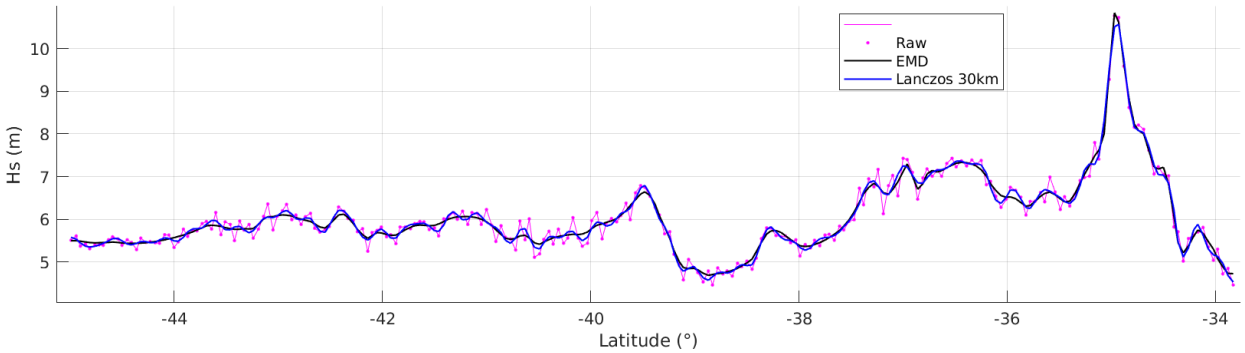
Examples



EMD-denoising (black curve) preserves large gradients and extreme values while Lanczos filter (60 km cut-off) operates excessive smoothing





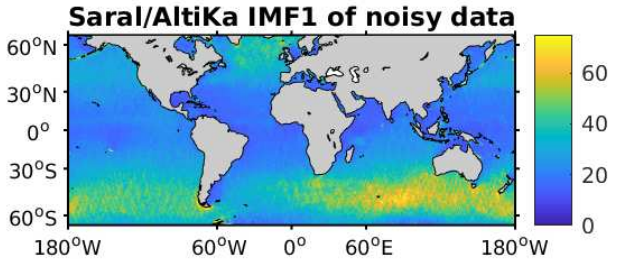


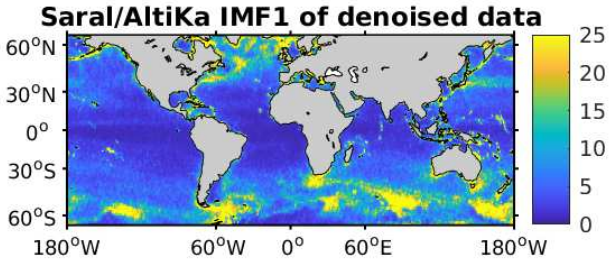
EMD-denoising (black curve) preserves by construction large geophysical gradients, while Lanczos filter needs 30 km cut-off to do so, at the expense of generating a bunch of artifact oscillations where SNR is lower

Global analysis of IMF1 variance: what is gained

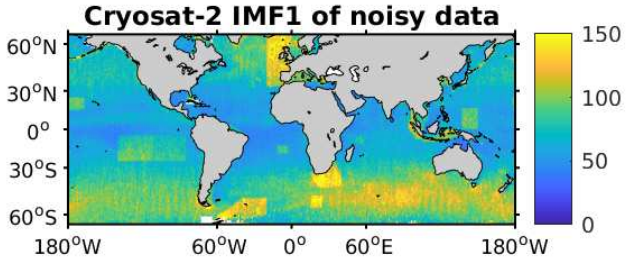
✓ EMD is a scale analysis tool, as Fourier or wavelet analysis, but its empirical nature is best suited to decompose non-linear and non-stationary signals

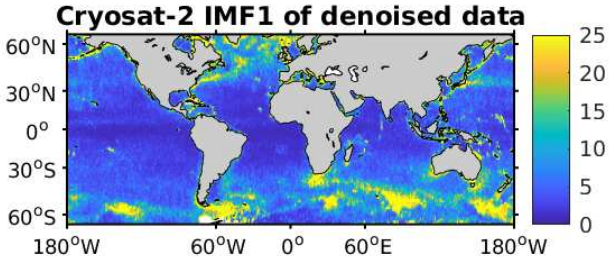
✓ As derived from EMD, IMF1 contains the high-frequency / small-scale content of the input signal. Mapping of the IMF1 variance for noisy measurements (left panels) reveals the noise distribution (e.g. SAR operations for Cryosat-2). Mapping of the IMF1 variance for denoised measurements (right panels) reveals overlooked sea state small mesoscale variability (< 100 km) linked to wave / current interactions, and illustrates the denoising process efficiency

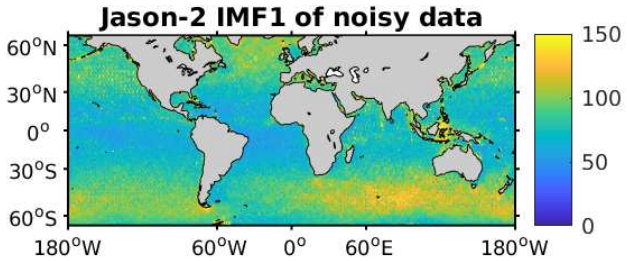


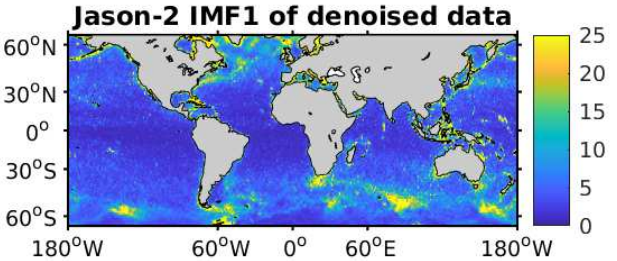


IMF1 mean square amplitude (cm^2) of Hs noisy (left panels) and denoised (right panels) measurements, period 2014 / 2016.
From Quilfen and Chapron (2019)









Future work

- ✓ Although the EMD denoising is self-adaptive to the actual noise, a constant is used to weight the prescribed thresholds, that has been initially set as a result of numerical simulations and subjectively from visual inspection. It can be tuned to smooth more or less the data, or to adjust other metrics linked to the noise and uncertainty distributions

- ✓ The EMD denoising should be still more efficient as post-processing of data processed with more recent re-tracking algorithms (than MLE4) if they offer:
  - a lower instrumental and processing noise
  - more stationary noise statistics (less Hs-dependent)
  - a better processing / identification of outliers (slicks, rainfall, land and ice contamination)

- ✓ The EMD denoising will be tested for high frequency altimeter measurements, e.g. to process the CFOSAT altimeter 5-Hz measurements that already benefit from the adaptive re-tracking