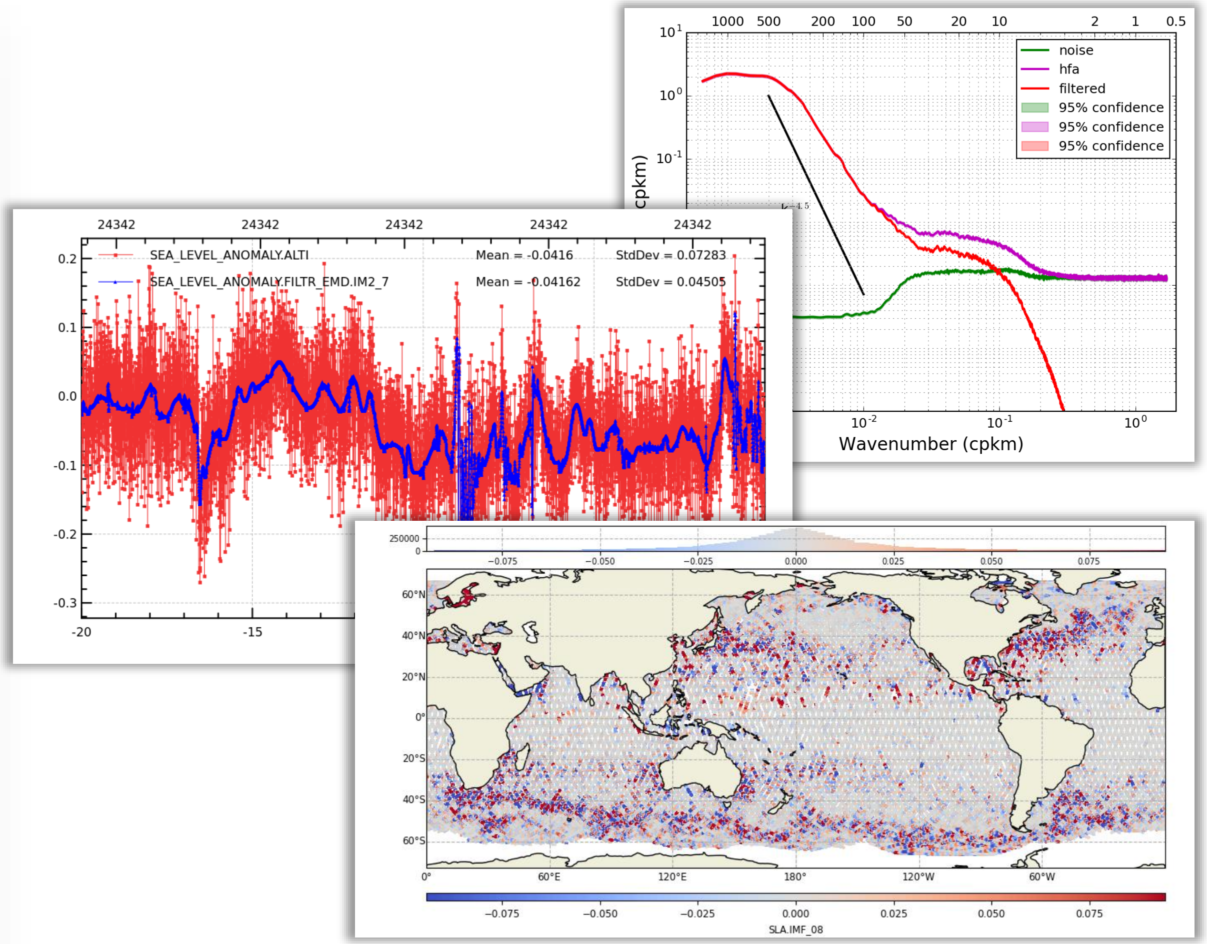
# EMD filtering applied to LRM 20 Hz sea level anomaly observations

F. Nencioli[1], M.-I. Pujol[1], N. Picot[2], G. Dibarboure[2] and Y. Quilfen[3]

[1] Collecte Localisation Satellites, [2] CNES, [3] Ifremer

2022 OSTST, 2022-11-01



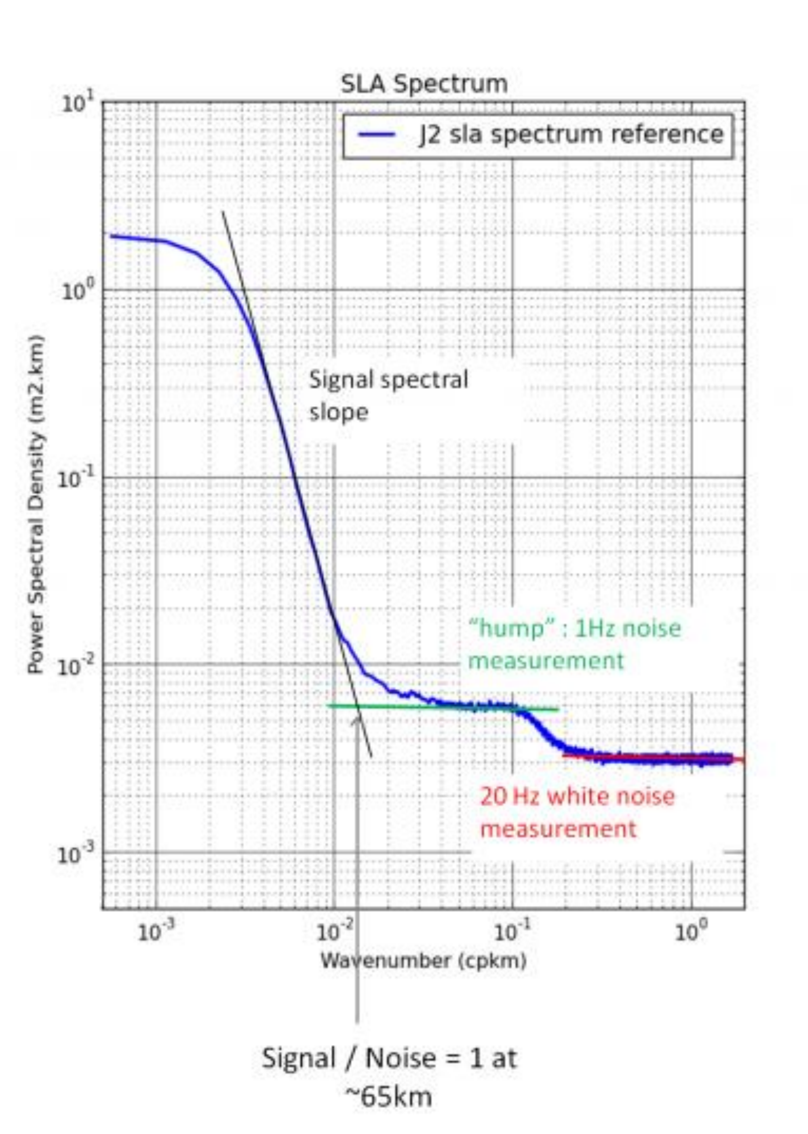


CLS
COLLECTE LOCALISATION SATELLITES
cnes
Ifremer

fnencioli@groupcls.com

# Context

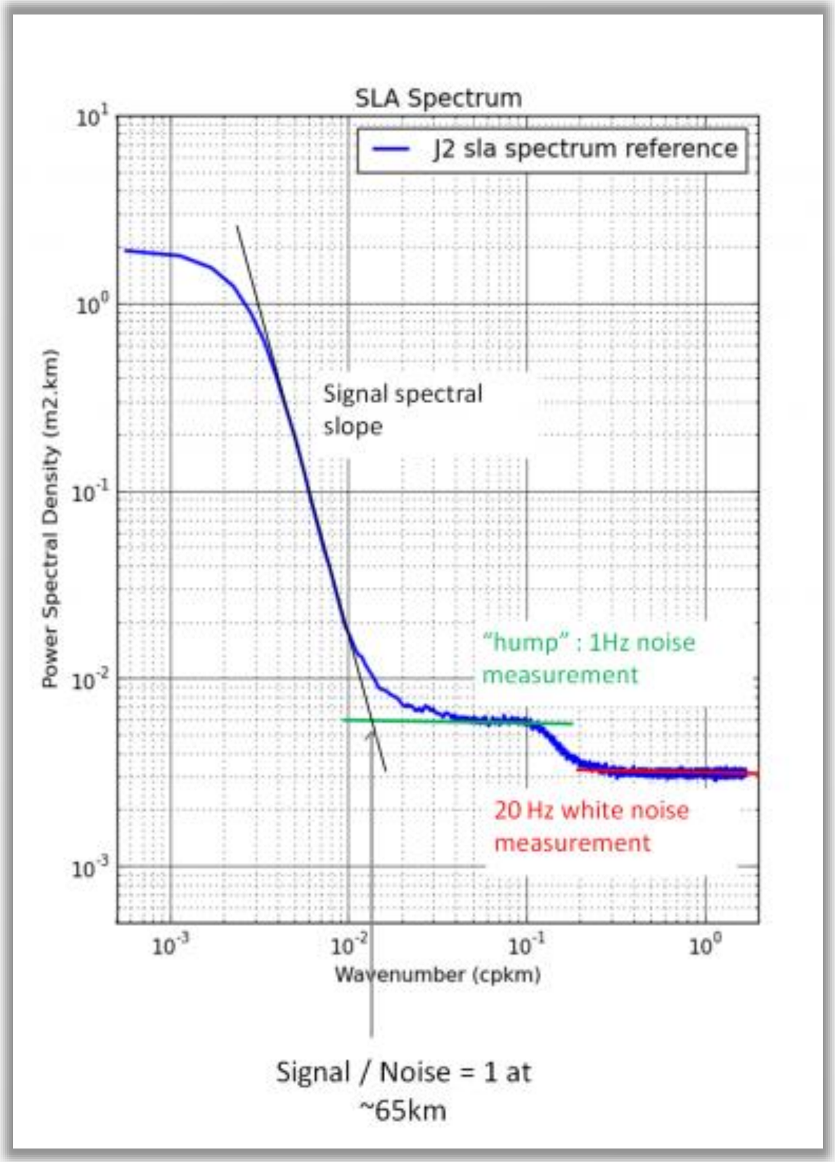
- 20 Hz along-track LRM altimetry observations collected at a spatial resolution of 300 m
- Possibility to investigate processes at spatial scales < 100 km
- Limitations due to “hump artefact” in the sea surface height (SSH) spectra at scales between 3 to 100 km
- Artefact due to to inhomogeneities in backscatter strength within the LRM footprint
- Induced retracking errors which are smoothed along the satellite track
- Spectral hump can already be mitigated by:
  - ➢ more restrictive editing algorithm
  - ➢ high-frequency specific corrections, such as the high-frequency adjustment (HFA)



# Objectives

- The the Empirical Mode Decomposition filter (EMD) is a novel filtering method
- Specifically designed for the analysis of non-stationary and non-linear signals

➢ Assess the effectiveness of the EMD filter in mitigating the impact of the hump



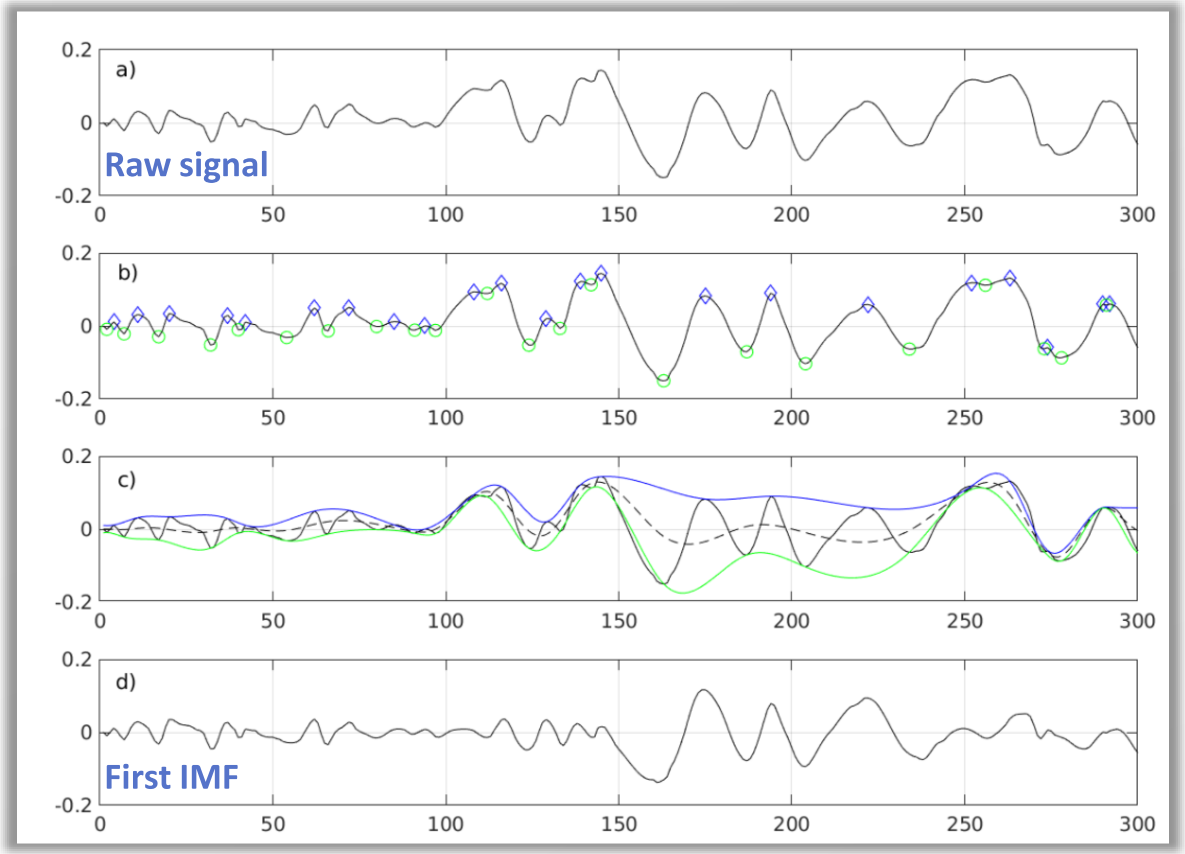

## Data

20Hz Jason-3 observations from cycle 20

From 23 Aug to 2 September 2016

Along-track Geophysical Data Record (GDR) product

# Empirical Mode Decomposition (EMD): method overview

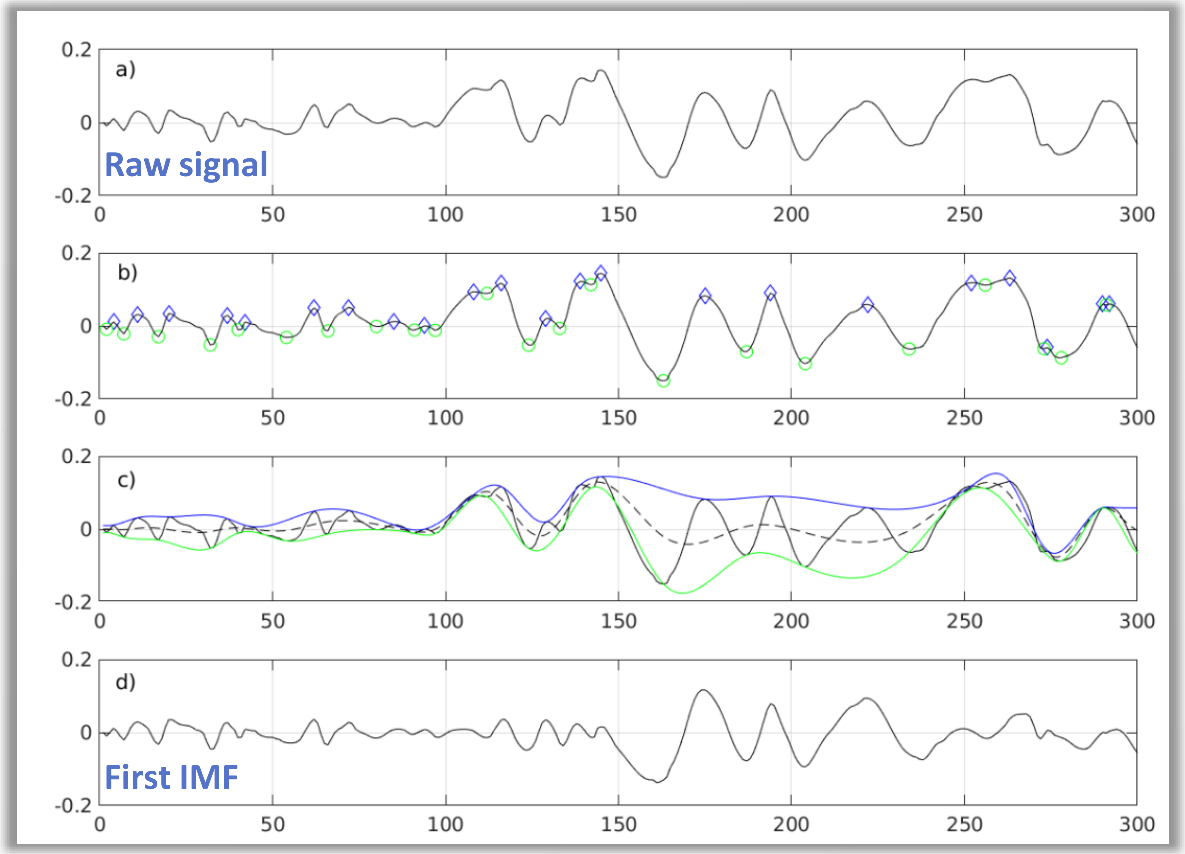



**Algorithmic method (not based on mathematical theory)**

1. Identify local minima and maxima of the input signal
2. Define upper and lower envelops through interpolation (spline) of local extrema
3. Compute the average of the two envelopes (dashed line)
4. Subtract this mean from the raw signal
5. Repeat 1 to 4 for a fixed number of iterations to identify the Intrinsic Mode Function (IMF)

Modulations by highest frequencies isolated as AM/FM function

# Empirical Mode Decomposition (EMD): method overview




Lots of parameters/options to be tuned by the user

**Algorithmic method (not based on mathematical theory)**

1. Identify local minima and maxima of the input signal
2. Define upper and lower envelops through interpolation (spline) of local extrema
3. Compute the average of the two envelopes (dashed line)
4. Subtract this mean from the raw signal
5. Repeat 1 to 4 for a fixed number of iterations to identify the first Intrinsic Mode Function (IMF)

Modulations by highest frequencies isolated as AM/FM function

- Repeat the process to identify successive IMFs at progressively lower frequencies
- Process halted when no more local minima/maxima are found
- Signal reconstructed using significant portion of IMFs

# EMD: Noise Energy and Denoising

➢ EMD denoising based on **identifying IMF1 noise energy** and **propagating it** to other IMFs

- The noise (***$n_1$***) is identified from the first IMF based on **wavelet denoising analysis**
- Noise time series is used to compute the noise Energy of the first IMF (E1) using the equation

$$E_1 = \left(\frac{median|n_1|}{0.6745}\right)^2$$

- *E1* is then propagated to the *n-th* IMF using the relation

$$E_n = \frac{E_1}{0.719} 2.01^{-n}$$

- *En* is used to define the noise threshold *Tn* to denoise the IMF n according to the equation

$$T_n = A\sqrt{E_n * 2logN}$$

- IMF peaks (max/min) with absolute values below the threshold T(n) are considered noise and removed from the signal reconstruction
- The signal is reconstructed using the equation

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

  - ❑ IMF<M1 are not used to reconstruct the signal (only noise, very low SNR)
  - ❑ M1<=IMF<=M2 are denoised
  - ❑ IMF>M2 are used without denoising (very large SNR)

- EMD denoising repeated several times **(20 iterations) shuffling the initial noise** timeseries (**CIIT method**)

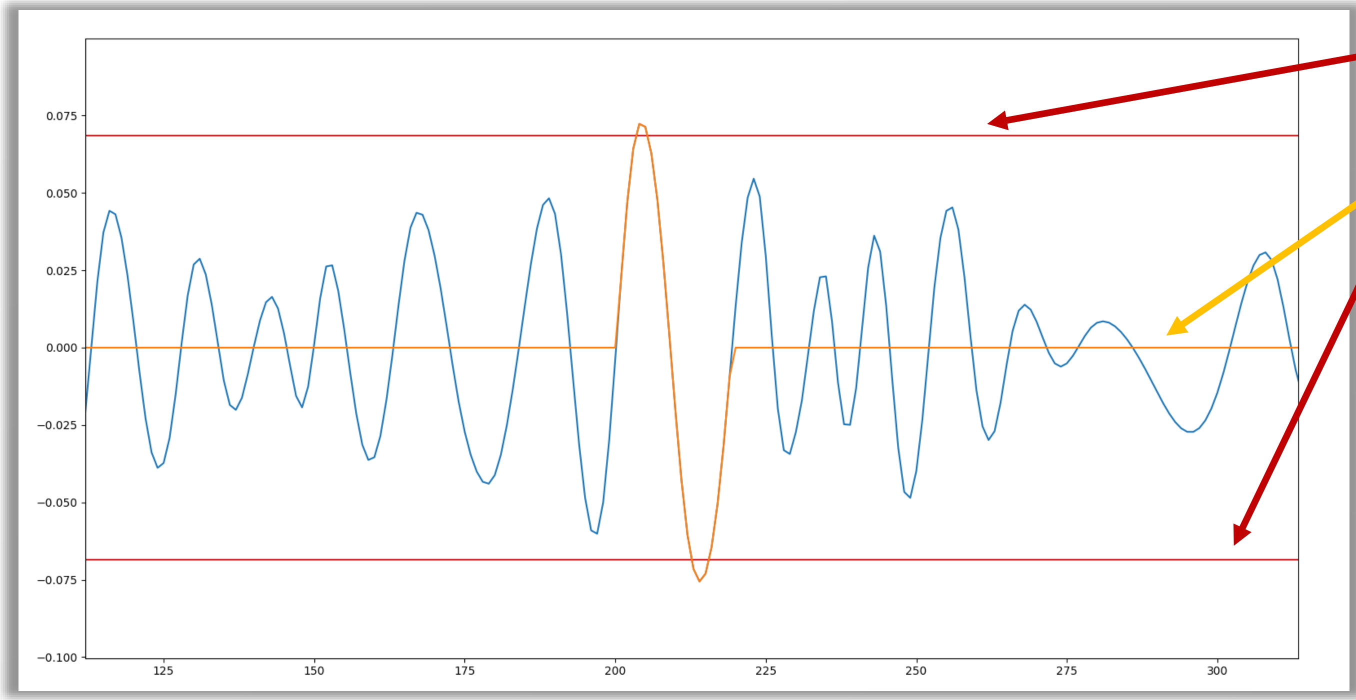
➢ Initial parameter used

$A = 0.7 \quad 2.5 < T_n < 3 \quad M_1 = 1 \quad M_2 = 3$

# EMD: Noise Energy and Denoising

## About peak selection

- ❑ For each IMF retained ony peaks above noise threshold
- ❑ Each IMF consists of flat portion + various peaks




**Noise levels**

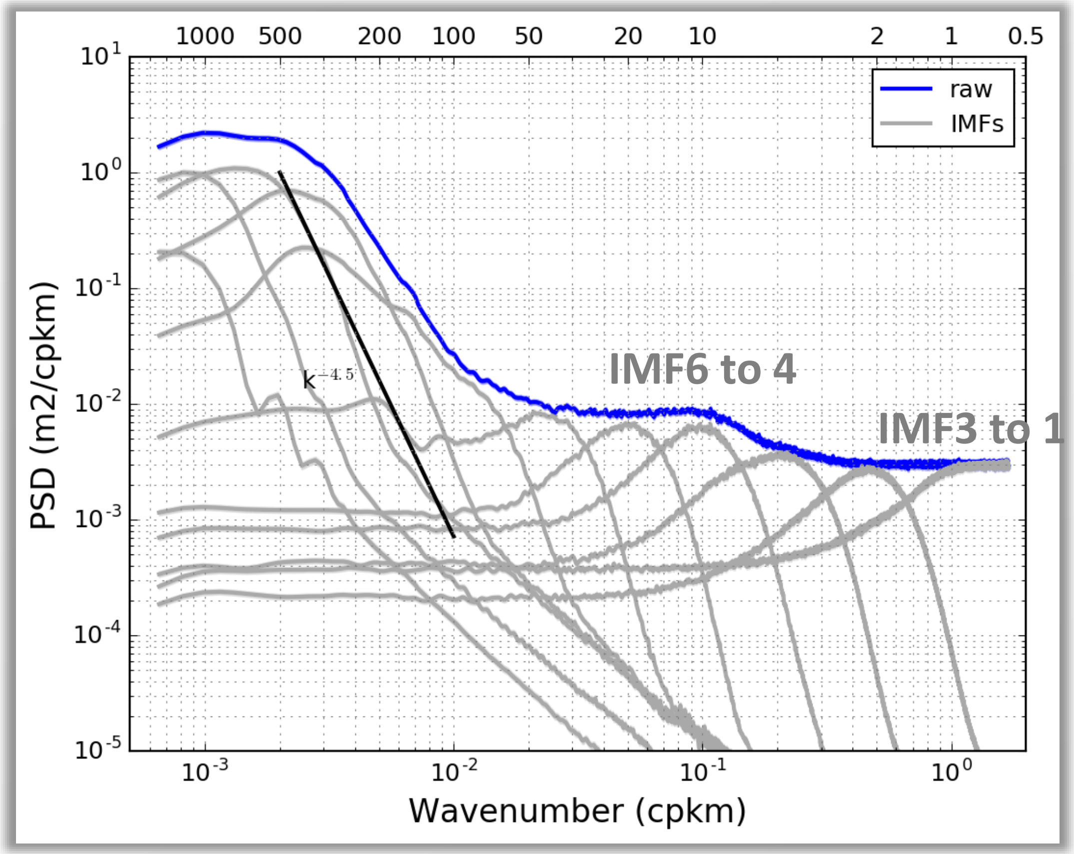
**Retained IMF (single peak)**

**<u>Even more options to be defined by users :</u>**

- ➢ Currently, if peak above threshold full peak is added to reconstructed signal (**HARD thresholding** option)
- ➢ Only signal above thershold could be added (**SOFT thresholding** option: not tested yet!!!)

# IMFs from Jason-3 Cycle 20 observations

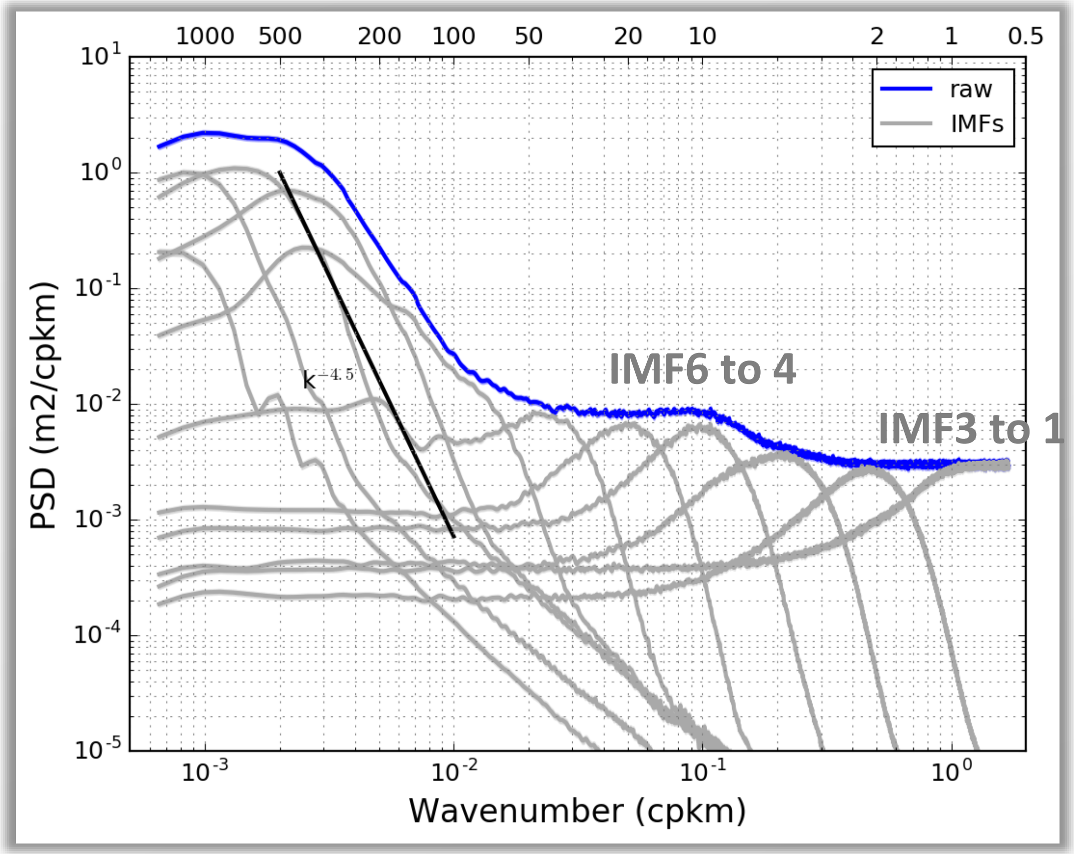
- IMF1 to 3 seem likely pure noise
- IMF4 to 6 within the hump portion of spectra




(Spectra computed on ~1500 km long segments)

# IMFs from Jason-3 Cycle 20 observations

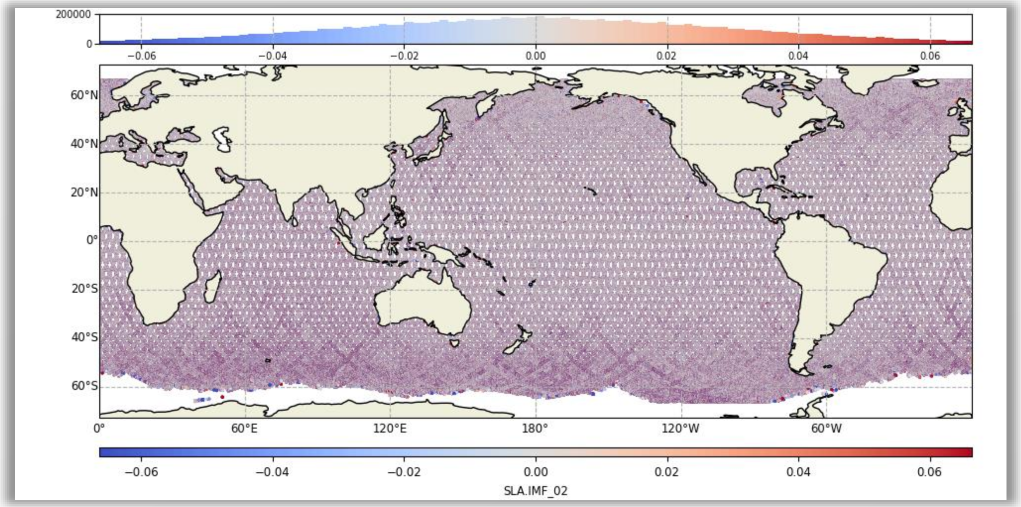
- IMF1 to 3 seem likely pure noise
- IMF4 to 6 within the hump portion of spectra



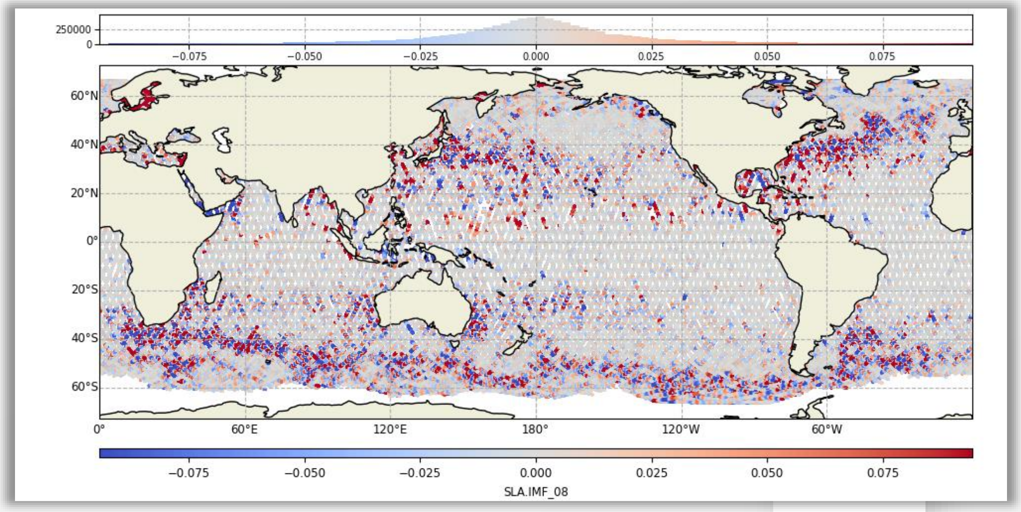

(Spectra computed on ~1500 km long segments)

**Along-track SLA for a given IMF**

**IMF2**
Uniform variability => Random noise to be removed



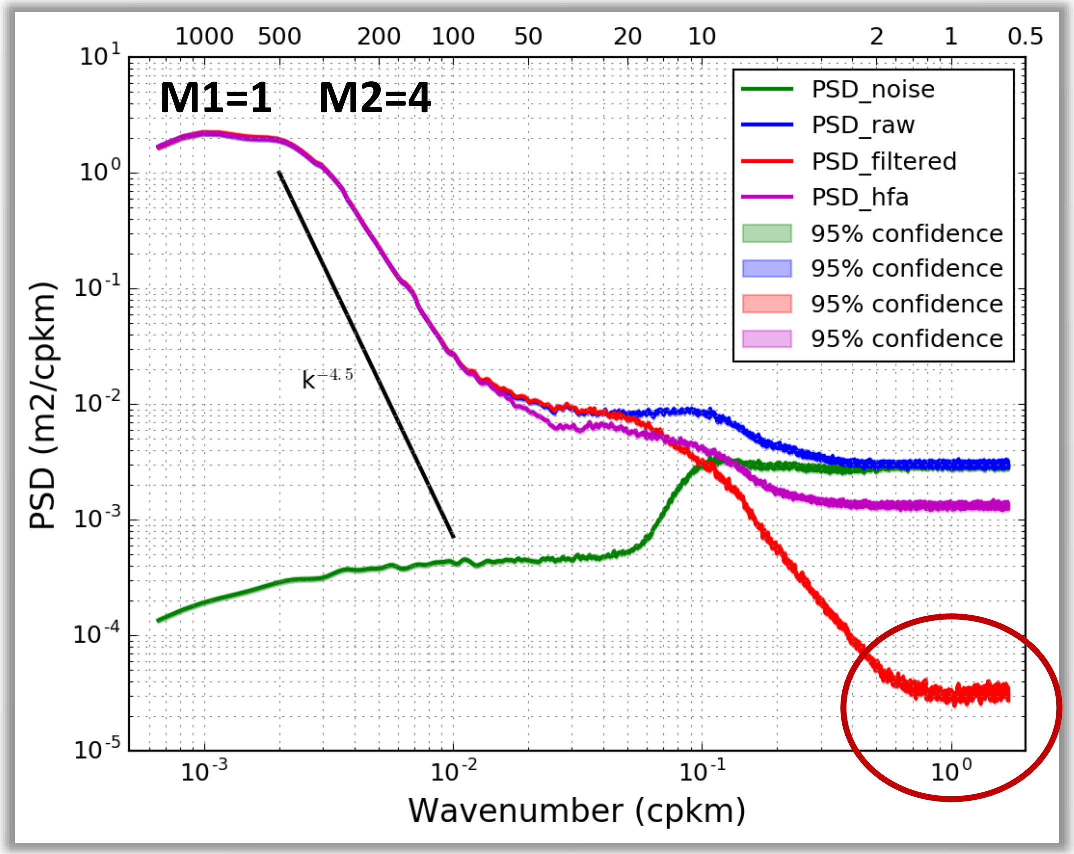

**IMF8**
Variability associated with main ocean current systems => physical signal to be retained



# Signal reconstruction: Number of denoised IMFs (M2)

## Data & Method

- J3 20Hz observations, Cycle 20
- CIIT method (denoised IMF1+ noise shuffle)



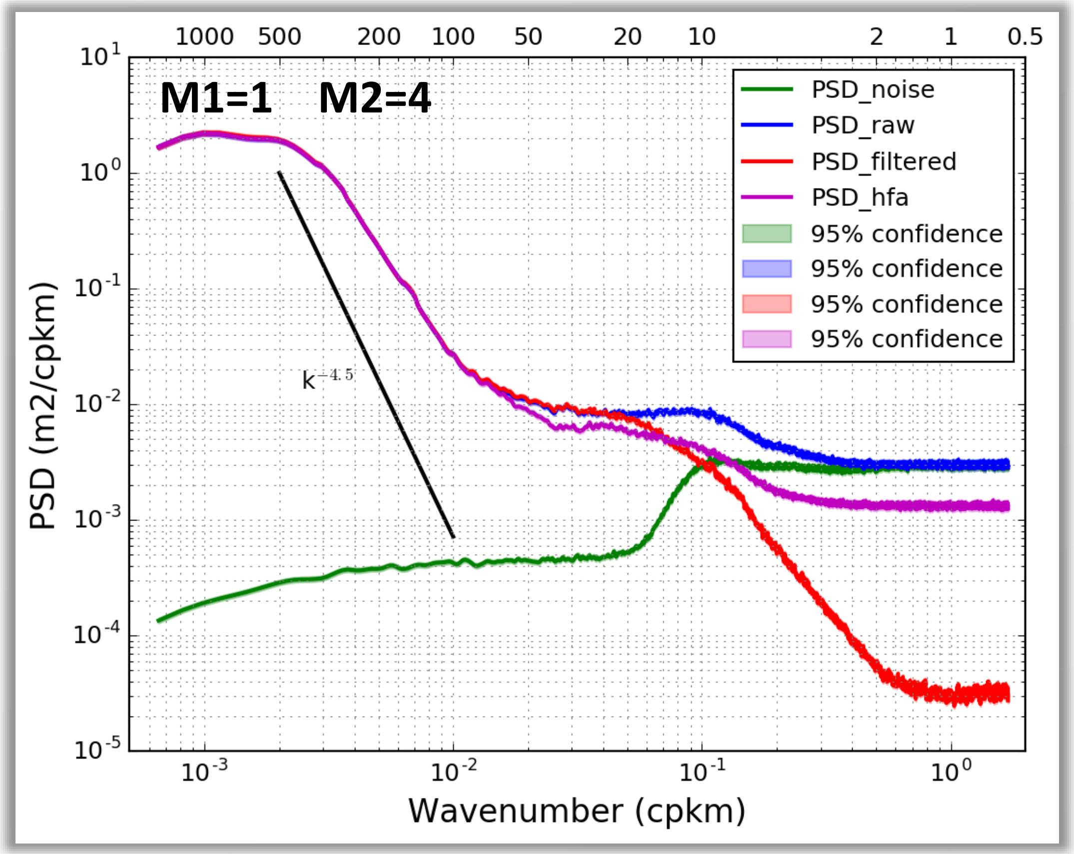

Please disregard this portion of the spectra

Due to a bug (now fixed!!!) in data preprocessing

# Signal reconstruction: Number of denoised IMFs (M2)

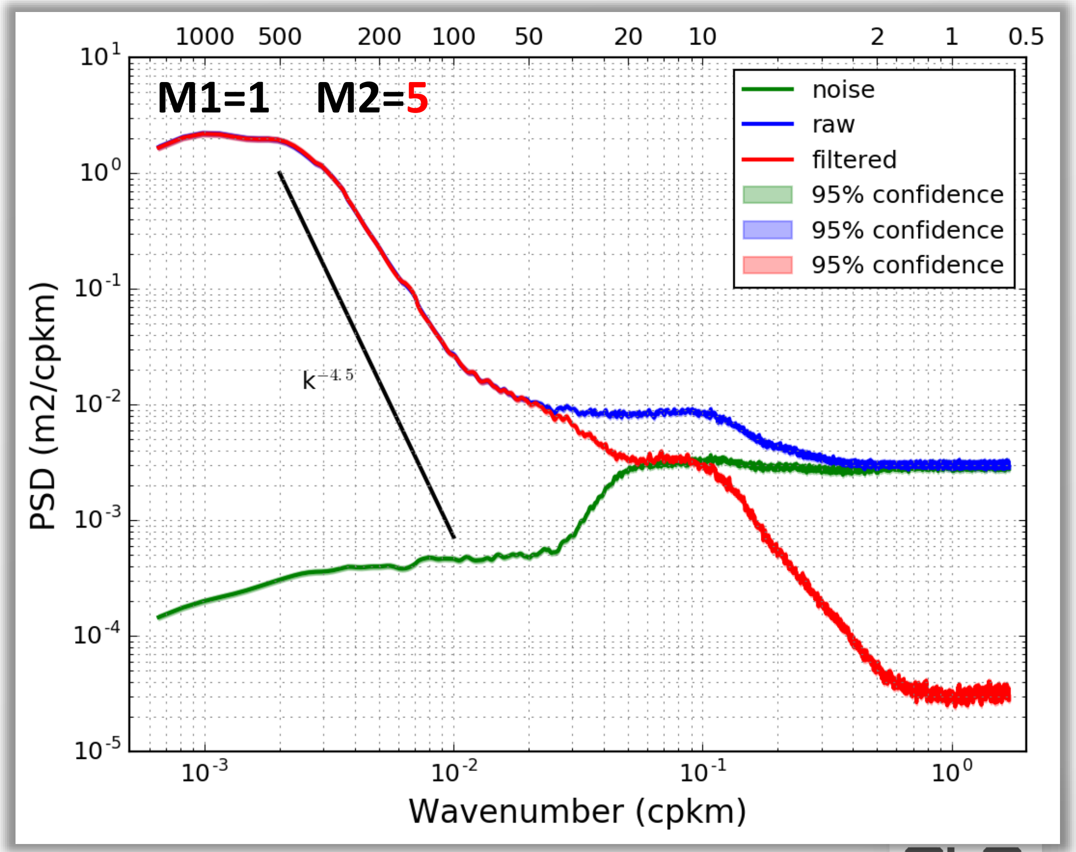
## Data & Method

- J3 20Hz observations, Cycle 20
- CIIT method (denoised IMF1+ noise shuffle)




## Increasing number of denoised IMF

- Flat portion of noise spectra shifted further towards longer wavelengths (~50 km)

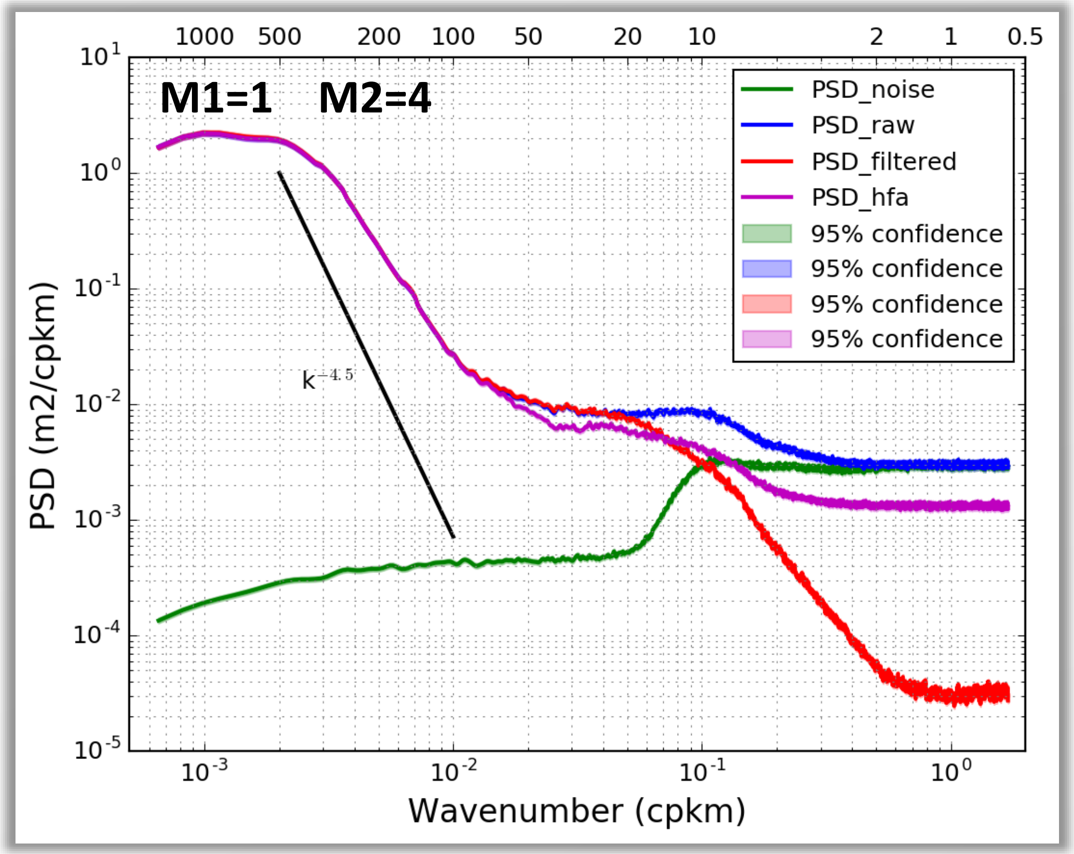



CLS
COLLECTE LOCALISATION SATELLITES

# Signal reconstruction: Number of denoised IMFs (M2)

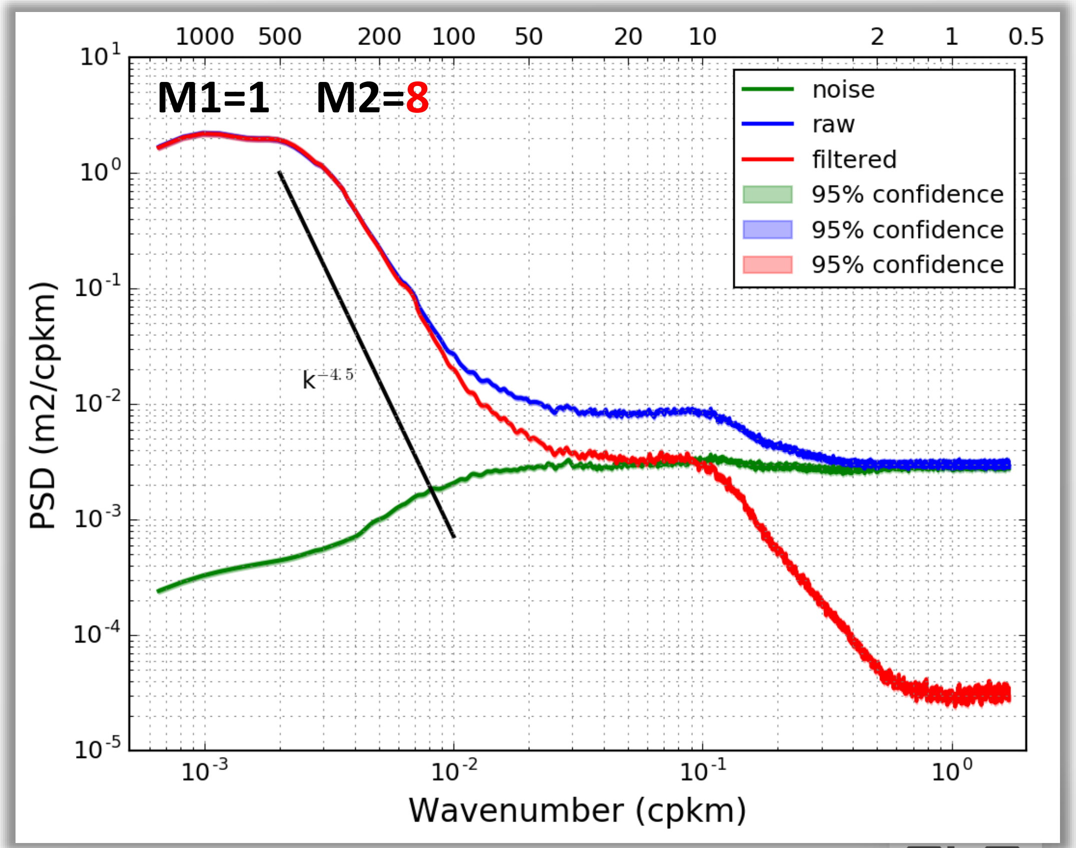
## Data & Method

- J3 20Hz observations, Cycle 20
- CIIT method (denoised IMF1+ noise shuffle)



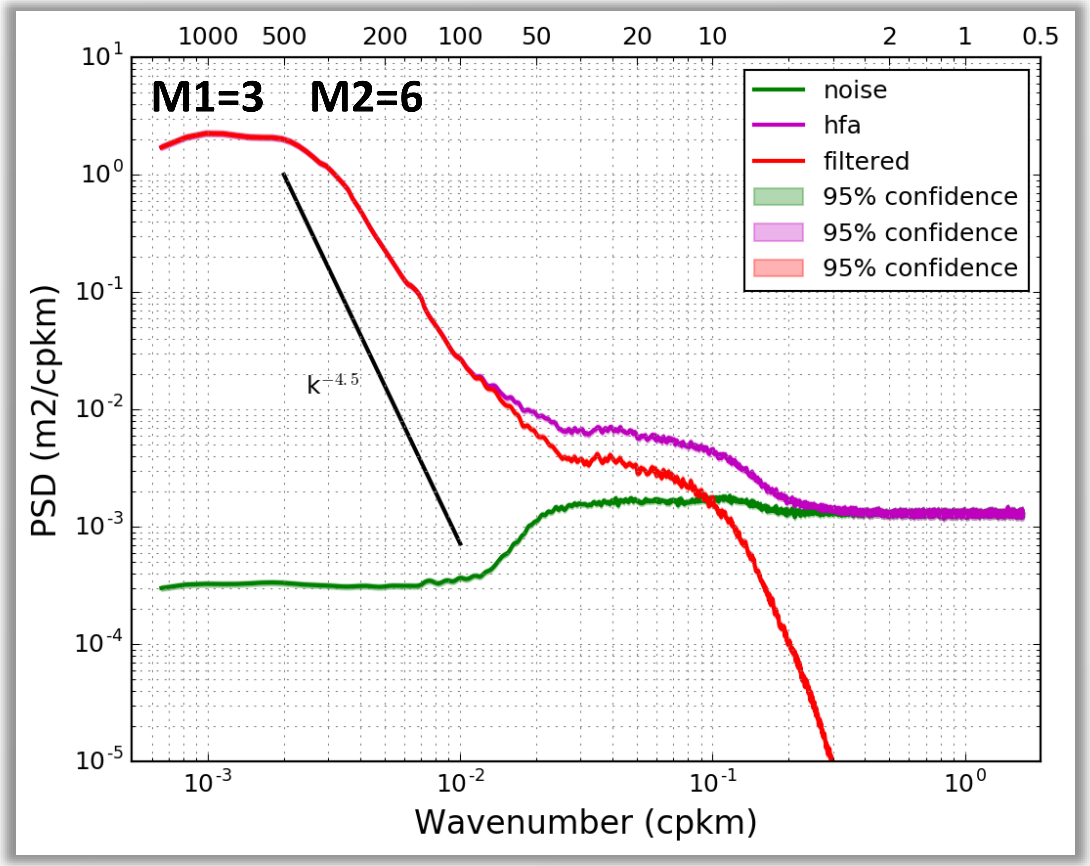

## Increasing number of denoised IMF

- Hump between 100 and 50 km persists, but spectral energy down to noise level



# EMD First conclusions

- EMD ineffective at removing spectral hump
- **To be used in synergy with HFA**
- **First IMF all noise** => no wavelet decomposition **(IIT method instead of CIIT)**
- Signal reconstructed only using signal starting from IMF 3
- IMFs 3 to 6 denoised




Bug fixed!!!

**Parameters MUST BE defined by the user!!!**

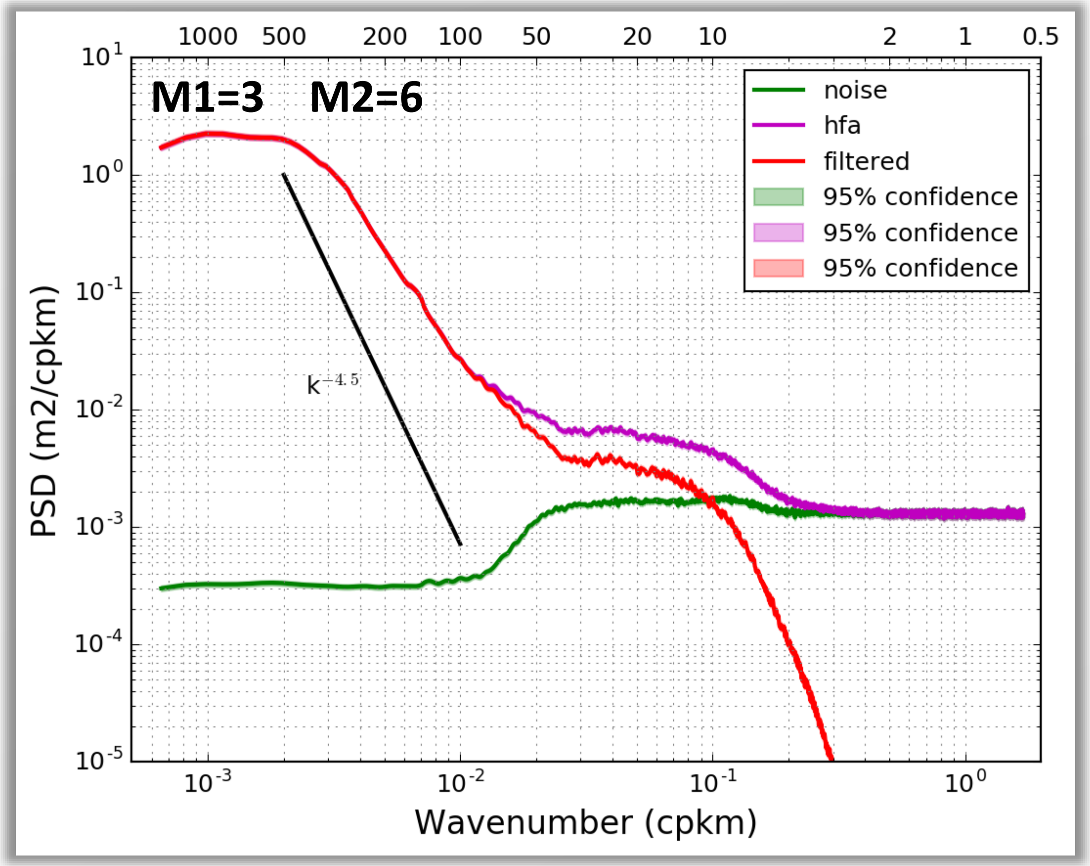
# EMD First conclusions

- ❑ EMD ineffective at removing spectral hump
- ❑ **To be used in synergy with HFA**
- ❑ **First IMF all noise** => no wavelet decomposition **(IIT method instead of CIIT)**
- ❑ Signal reconstructed only using signal starting from IMF 3
- ❑ IMFs 3 to 6 denoised

- Following the noise Energy relation…
- …up to IMF5 still significant noise contribution

| | | |
|---|---|---|
| IMF1 | | 60% total noise |
| IMF2 | 34.4% IMF1 noise | 20% total noise |
| IMF3 | 17.1% IMF1 noise | 10% total noise |
| IMF4 | 8.5% IMF1 noise | 5% total noise |
| IMF5 | 4.2% IMF1 noise | 2.5% total noise |

$$E_n = \frac{E_1}{0.719} 2.01^{-n}$$

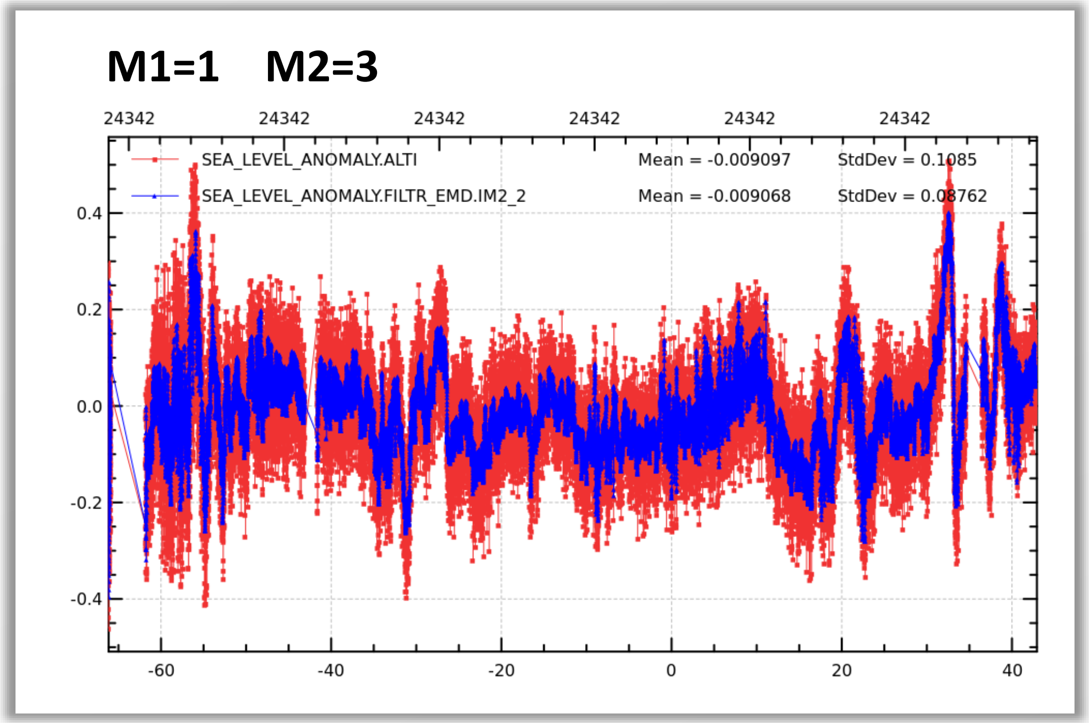



**Parameters MUST BE defined by the user!!!**

# Reconstructed signal: Number of denoised IMFs (M2)

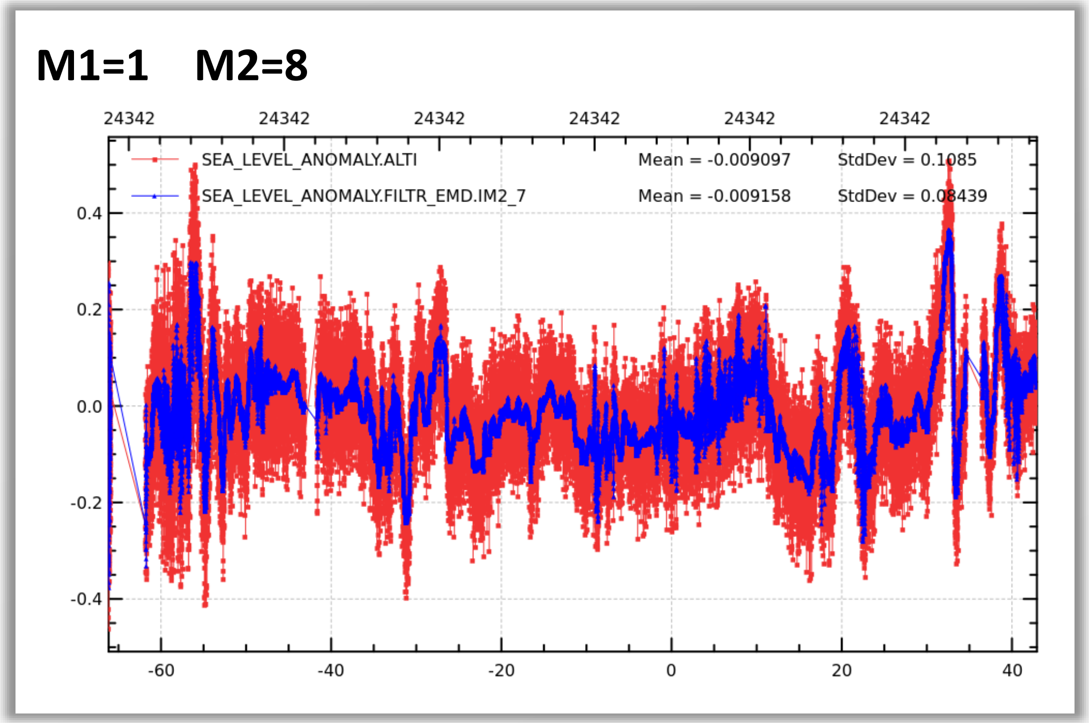
## Data & Method

- J3 20Hz observations, Cycle 20 Track 10
- CIIT method (denoised IMF1+ noise shuffle)




## Larger number of denoised IMFs

- Decreased high frequency content (as expected)

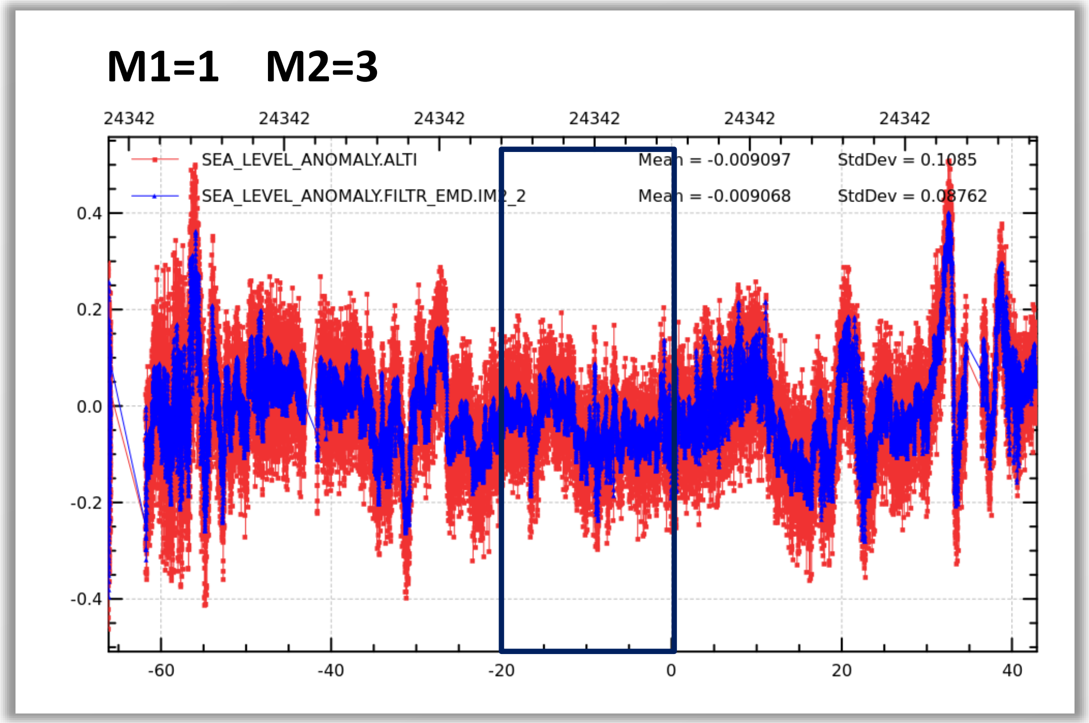



CLS

# Reconstructed signal : Number of denoised IMFs (M2)

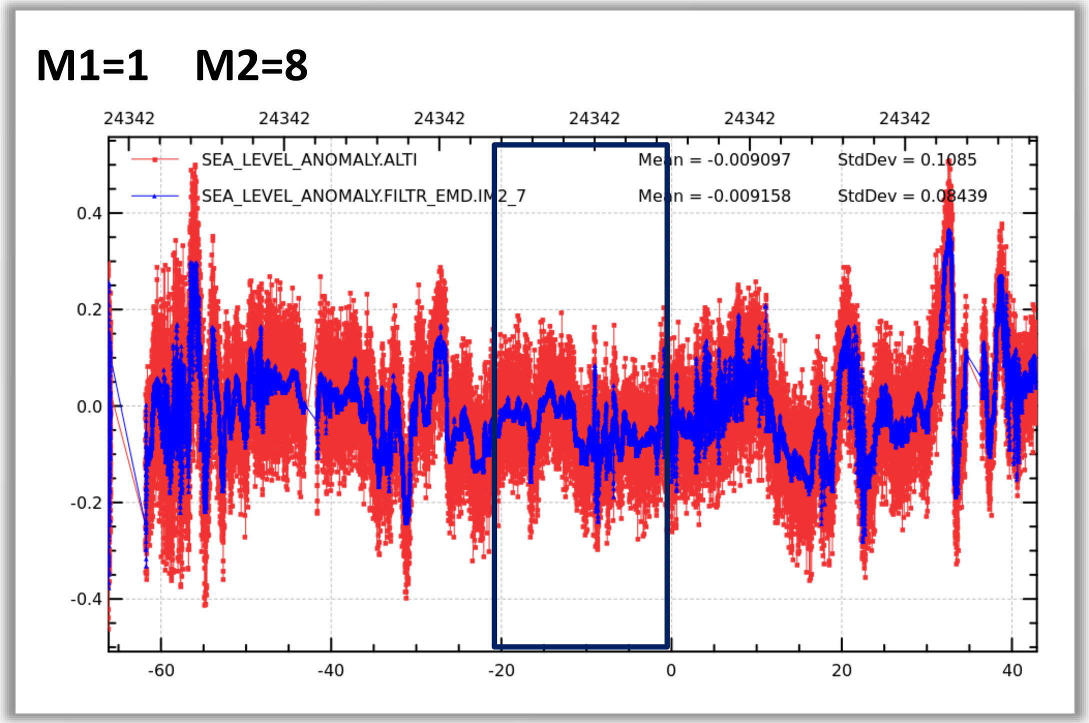
## Data & Method

- J3 20Hz observations, Cycle 20 Track 10
- CIIT method (denoised IMF1+ noise shuffle)




## Larger number of denoised IMFs

- Decreased high frequency content (as expected)

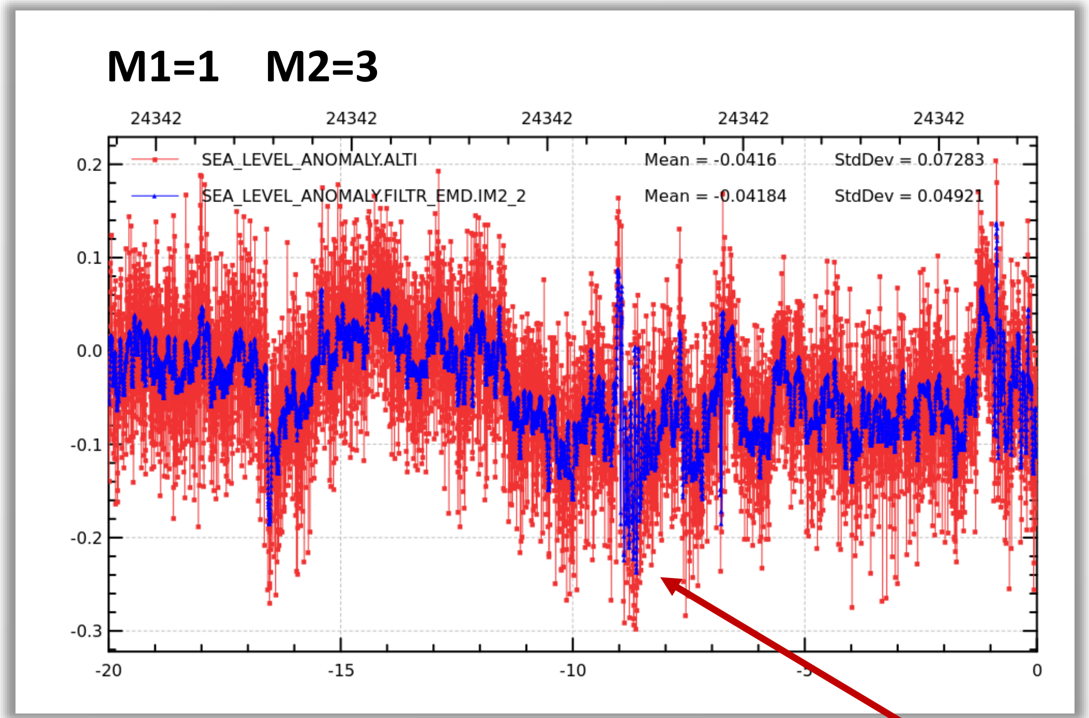



Next slide: Zoom -20 to 0

# Reconstructed signal : Number of denoised IMFs (M2)

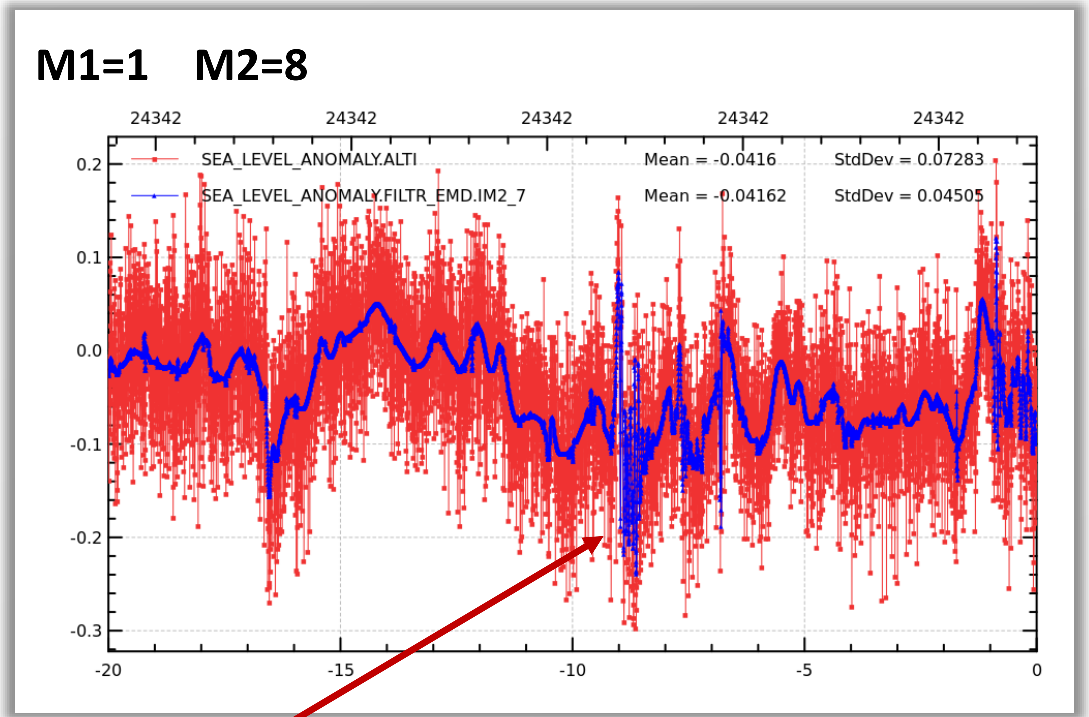
## Data & Method

- J3 20Hz observations, Cycle 20 Track 10
- CIIT method (denoised IMF1+ noise shuffle)




## Larger number of denoised IMFs

- Decreased high frequency content (as expected)




Persisting region of large high-frequency oscillations

CLS
COLLECTE LOCALISATION SATELLITES

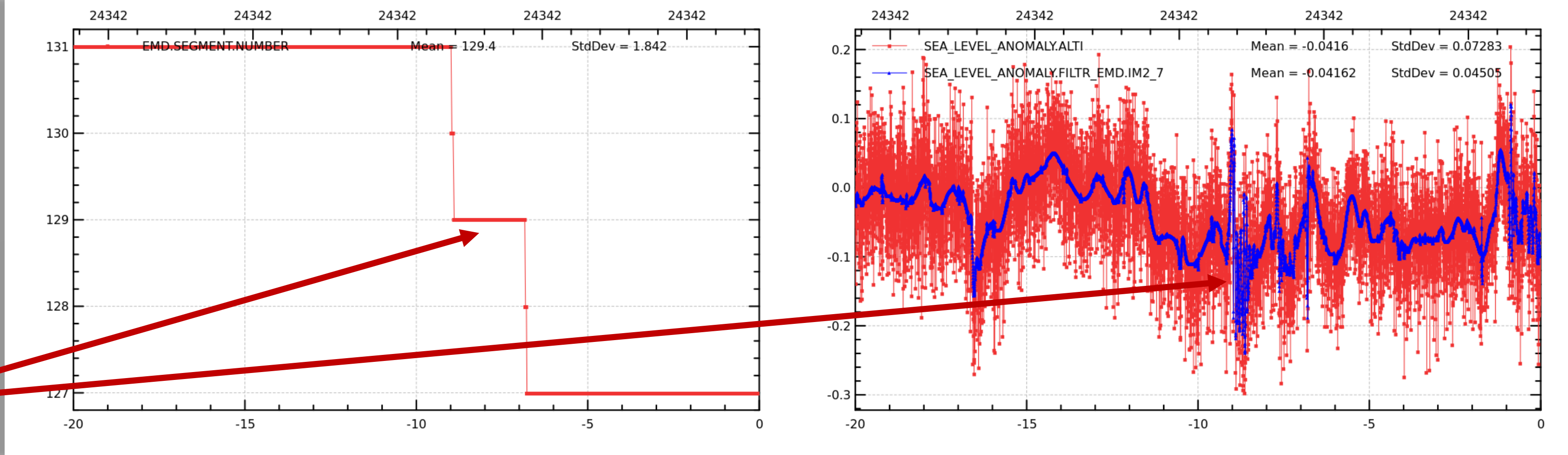
# Reconstructed signal : Track Segmentation

**EMD requires continuous data record**

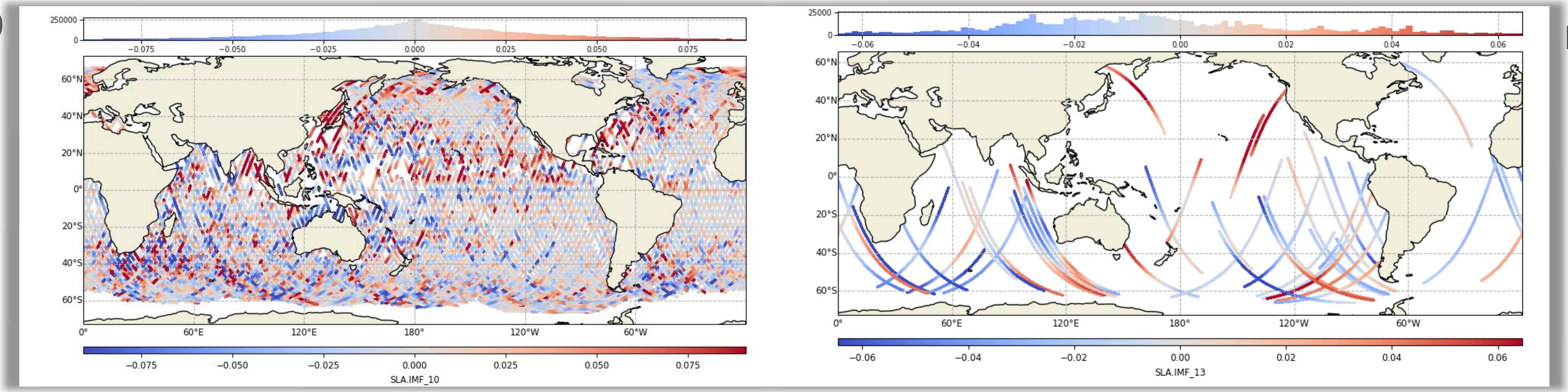
**Original segmentation**:

- Per day
- No filling

- ❑ Few segments with long IMFs (>IMF10)
- ❑ Discontinuities due to short segments (**different noise levels**)




**IMF10**




**IMF13**

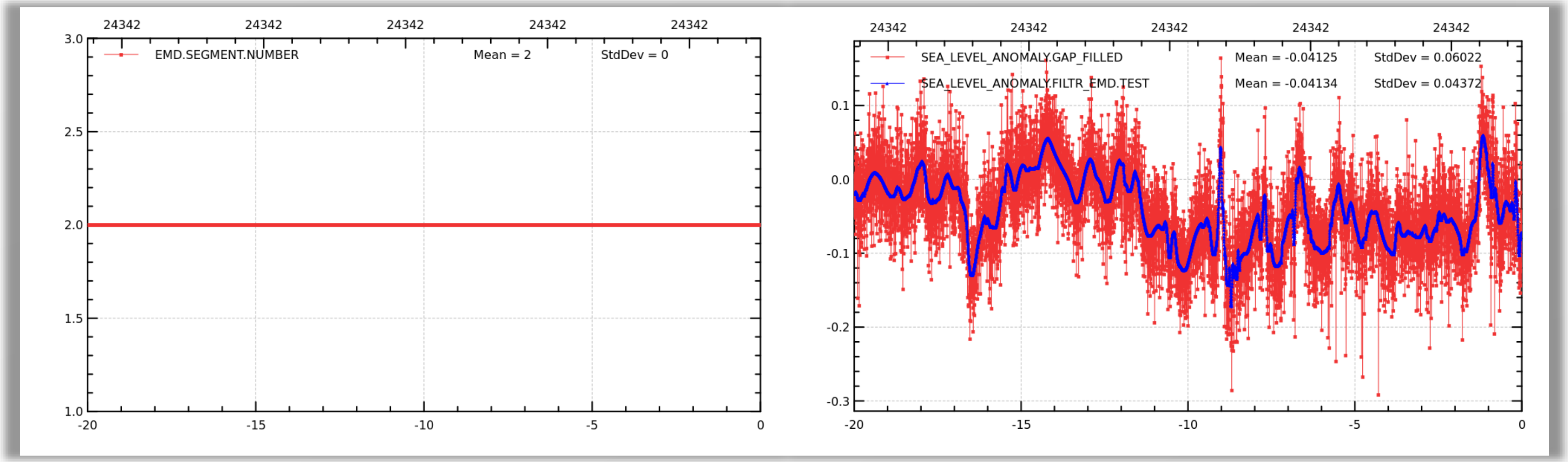
# Reconstructed signal : Track Segmentation

**EMD requires continuous data record**

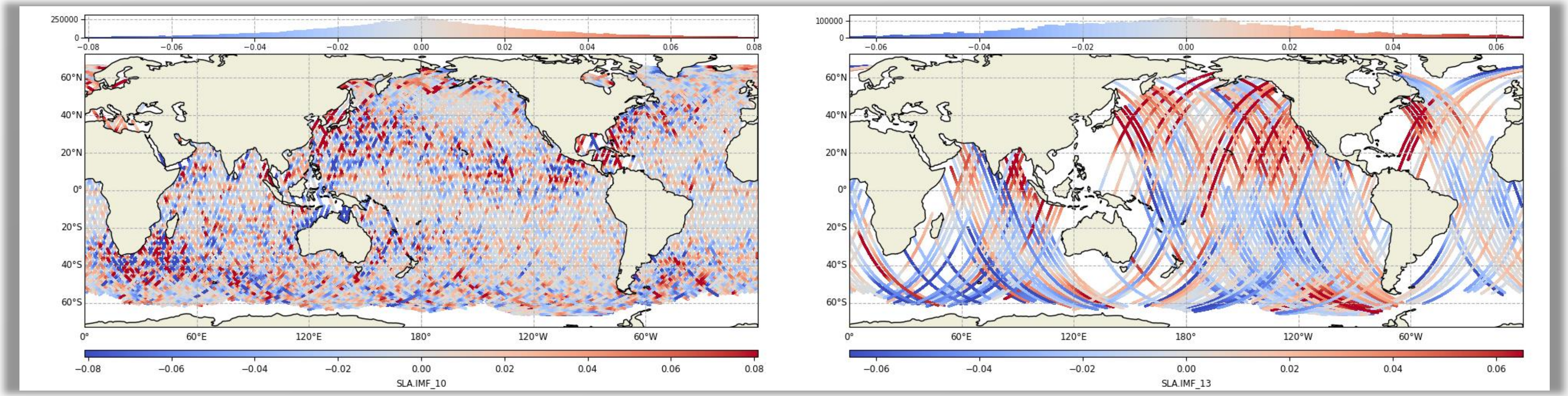
**New segmentation**

- Per track
- Gap filling

- ❑ More segments with long IMFs (>IMF10)
- ❑ Same noise level along the full track (no segments)




**IMF10**



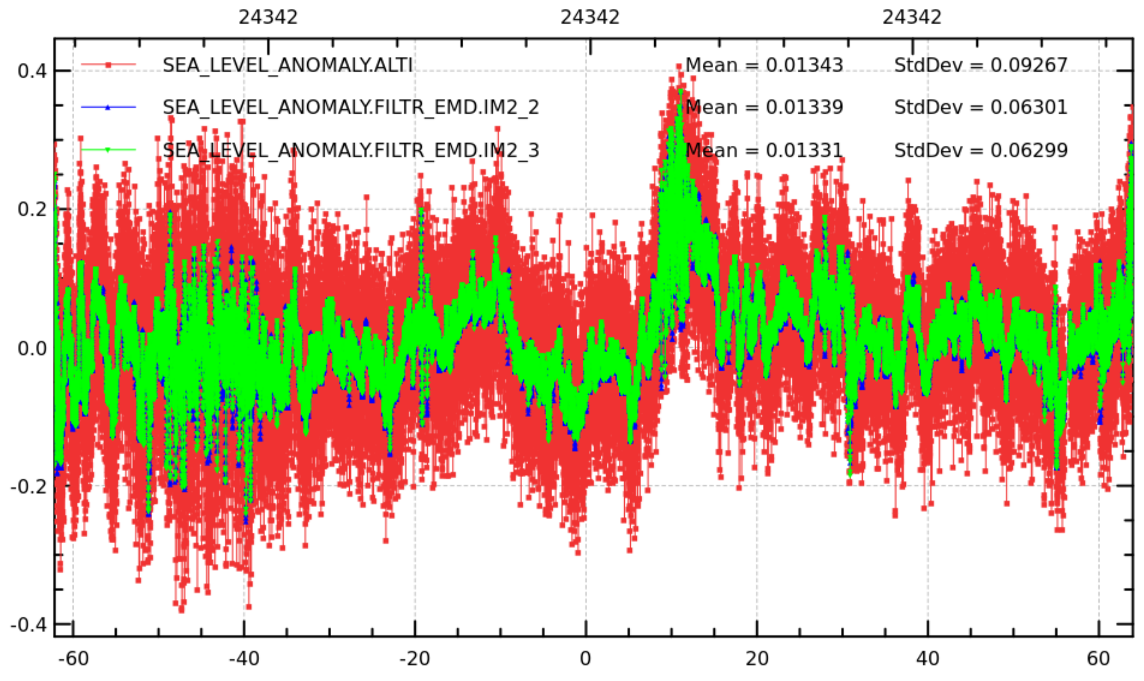

**IMF13**

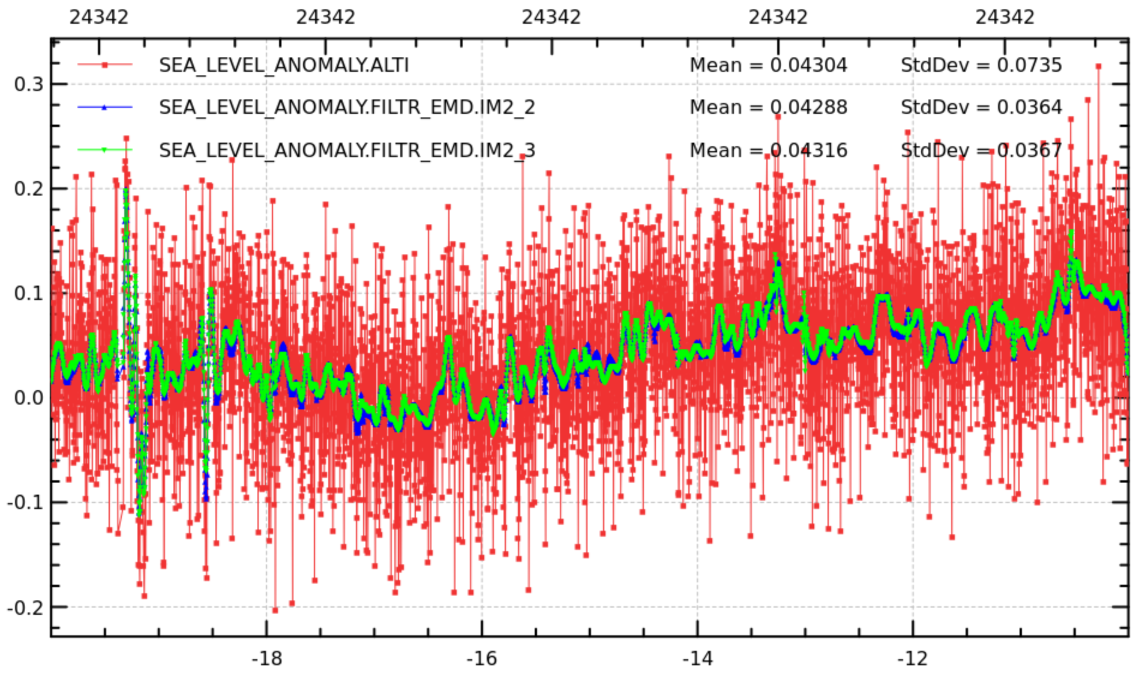
CLS

# Reconstructed signal : Noise Shuffle (ITT method)

- ❑ Filtered signal reconstructed from averaging the results from 20 EMD denoising replica
- ❑ For each replica, the noise time series identified from firsts IMF is randomly shuffled
- ➢ Applying the filter consecutive times with the same parameters returns **slightly different curves** (red and green)

J3 Cycle 020 track 004

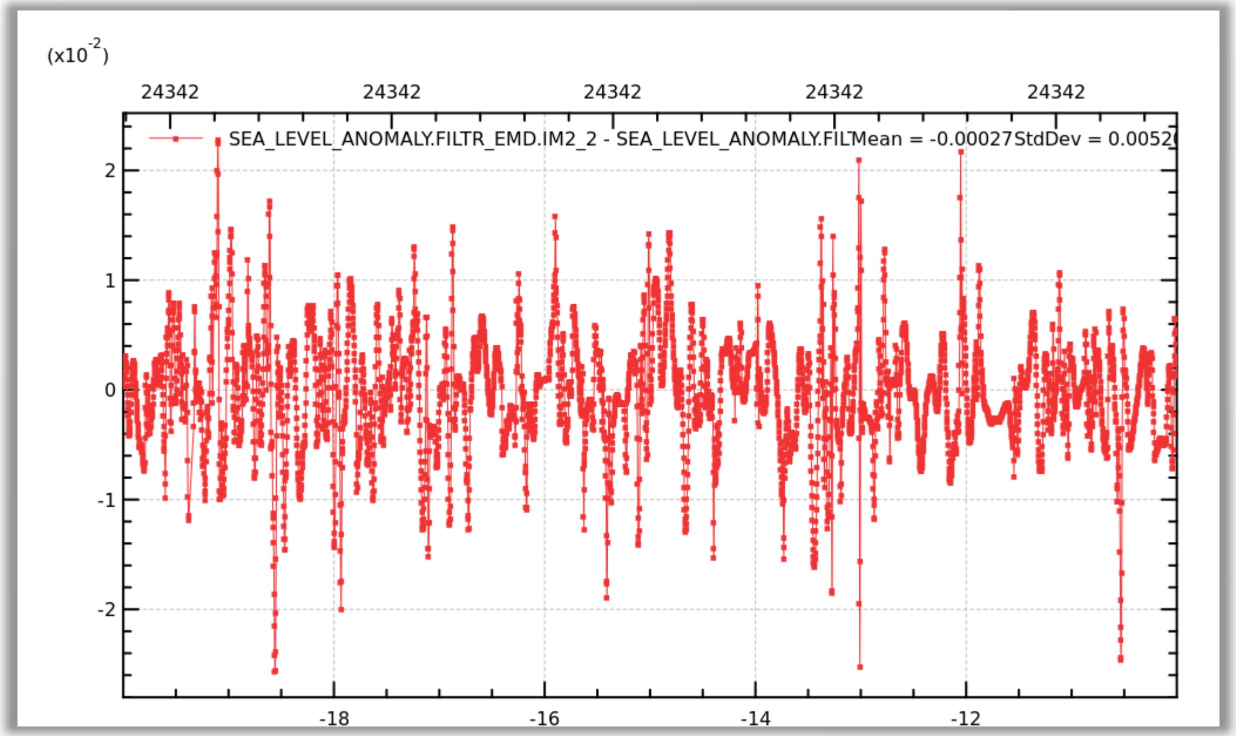


# Reconstructed signal : Noise Shuffle (ITT method)

- ❑ Filtered signal reconstructed from averaging the results from 20 EMD denoising replica
- ❑ For each replica, the noise time series identified from firsts IMF is randomly shuffled
- ➢ The differences (red curve) can be larger than 2 cm (and persist even doubling the number of EMD replicas)

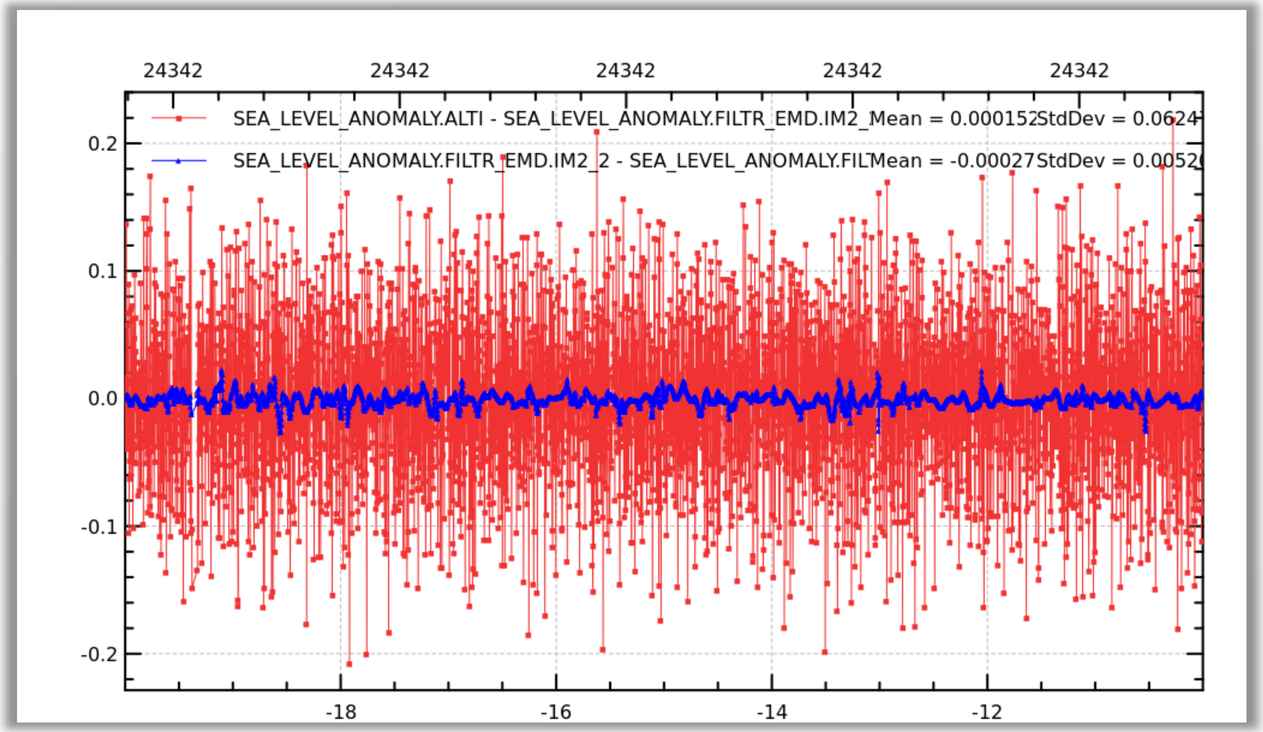
J3 Cycle 020 track 004



# Reconstructed signal : Noise Shuffle (ITT method)

- ❑ Filtered signal reconstructed from averaging the results from 20 EMD denoising replica
- ❑ For each replica, the noise time series identified from firsts IMF is randomly shuffled
- ➢ The difference (**blue** curve) are of the order ~10% of the removed signal (**red** curve)

J3 Cycle 020 track 004



# EMD filtering: the Interval-thresholding (IT) method

- Can noise shuffling be avoided? Yes , **IT method (single EMD replica)**
- More user defined choices on possible EMD methods (not only parameter values):
  1. **Clear iterative interval-thresholding** (CIIT): Several EMD replicas (noise shuffle) + Wavelet denoising of first IMF
  2. **Iterative interval-thresholding** (IIT): Several EMD replicas (noise shuffle) + No denoising of first IMF (all noise)
  3. **Interval-thresholding** (IT): Single EMD replica (no noise shuffle) + no wavelet denoising of first IMF (all noise)

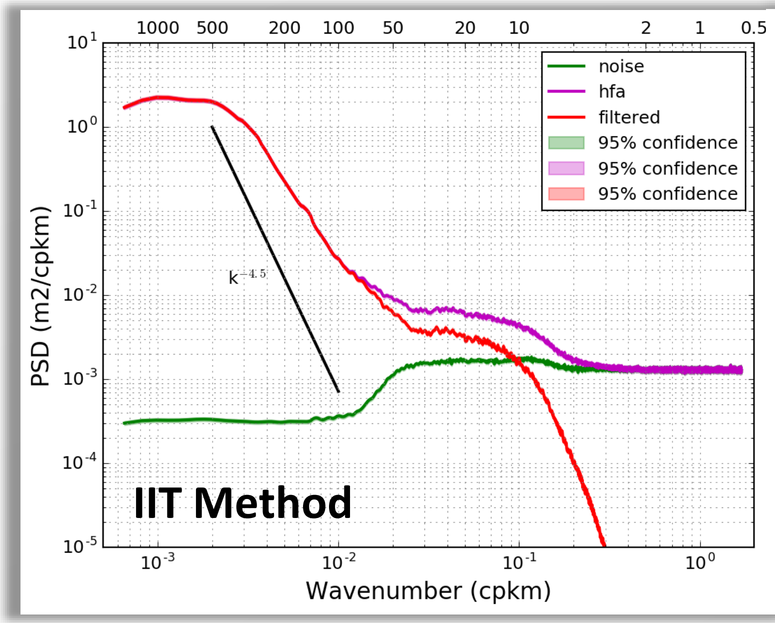
CLS
COLLECTE LOCALISATION SATELLITES

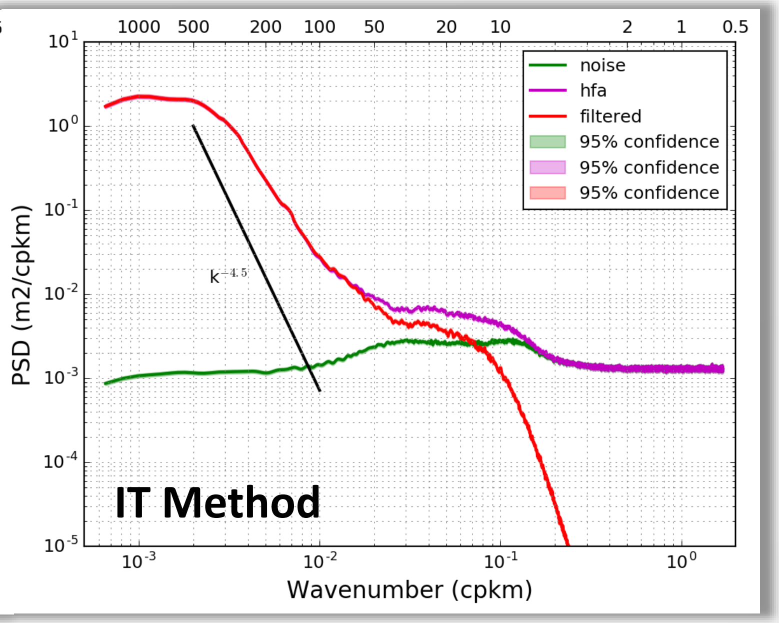
# EMD filtering: the Interval-thresholding (IT) method

- ❑ Can noise shuffling be avoided? Yes , **IT method (single EMD replica)**
- ❑ More user defined choices on possible EMD methods (not only parameter values):
  1. **Clear iterative interval-thresholding** (CIIT): Several EMD replicas (noise shuffle) + Wavelet denoising of first IMF
  2. **Iterative interval-thresholding** (IIT): Several EMD replicas (noise shuffle) + No denoising of first IMF (all noise)
  3. **Interval-thresholding** (IT): Single EMD replica (no noise shuffle) + no wavelet denoising of first IMF (all noise)

- ➢ IIT and IT methods show similar spectra
- ➢ How about the returned filtered signals?

(Never stop at the spectra!!!)



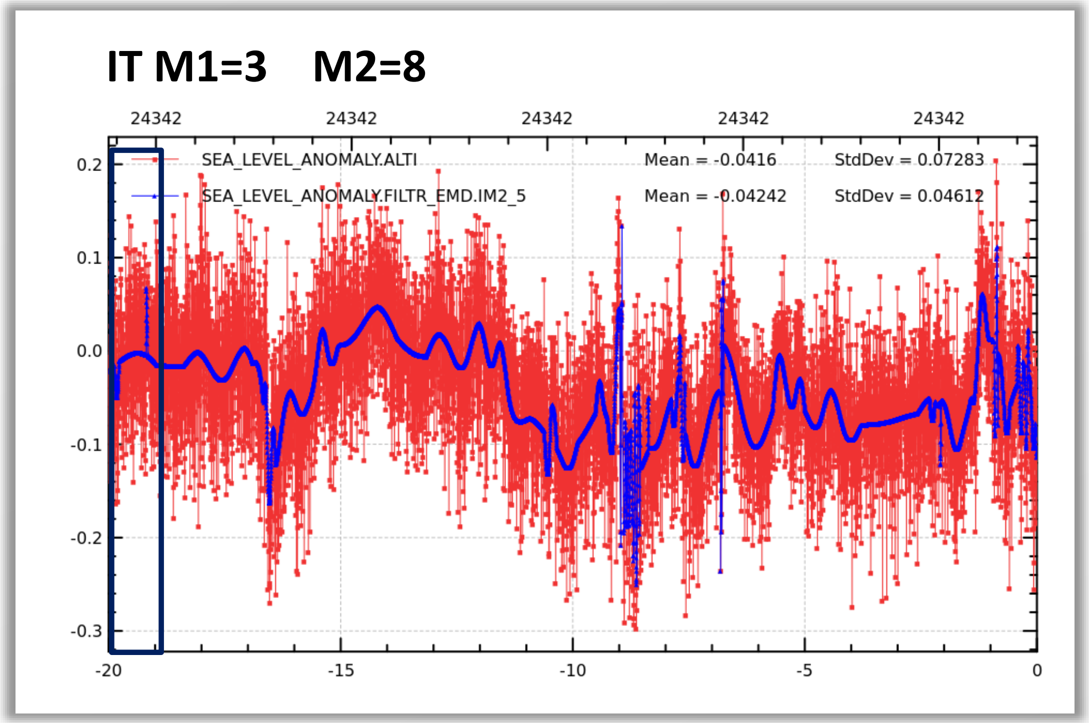




CLS

# Reconstructed signal : IT method

## Data & Method

- J3 20Hz observations, Cycle 20 Track 10
- IT method (no denoised IMF1+ no noise shuffle)

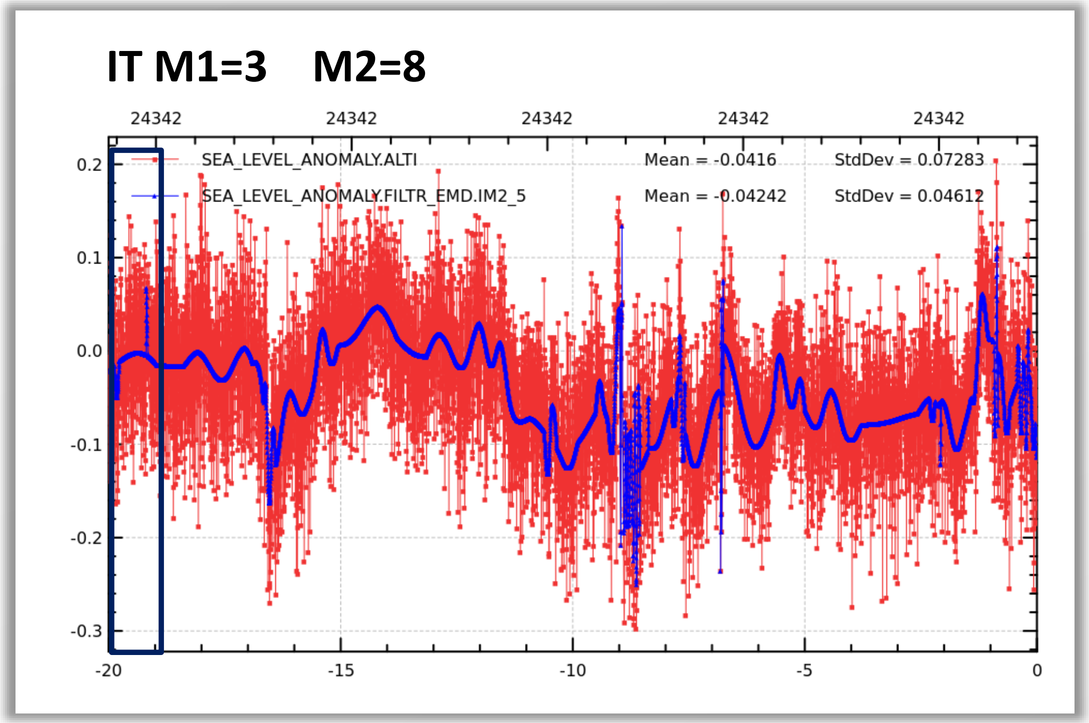



**Zoom -20 to -19 Latitude**

# Reconstructed signal : IT method

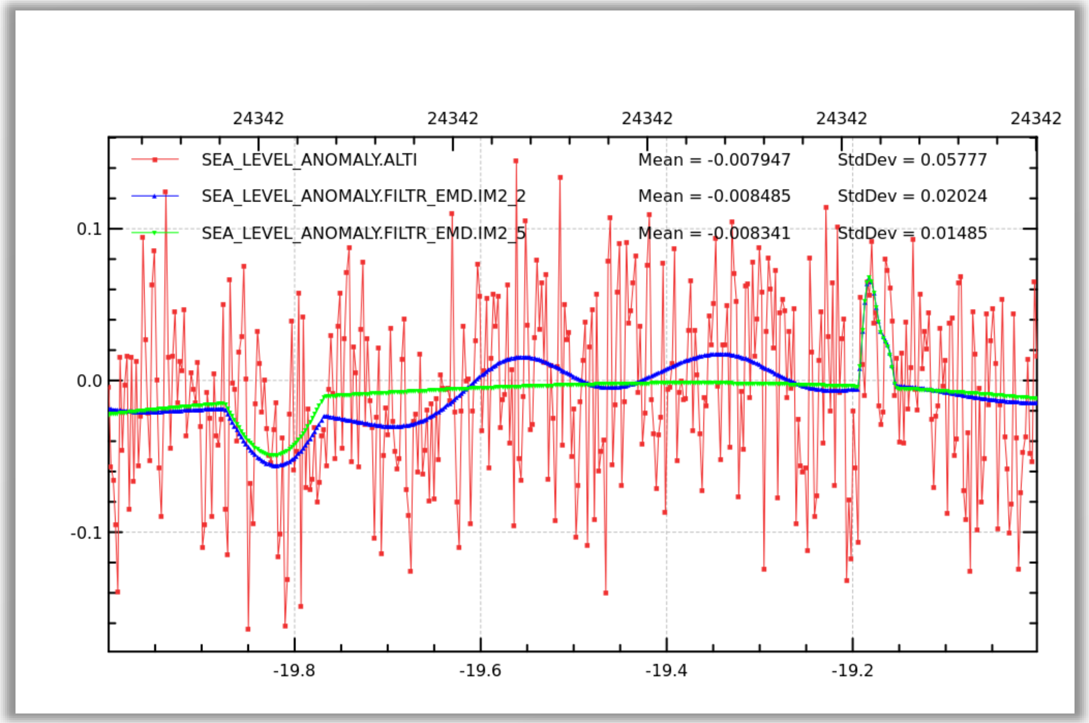
## Data & Method

- J3 20Hz observations, Cycle 20 Track 10
- IT method (no denoised IMF1+ no noise shuffle)




## Results (along-track)

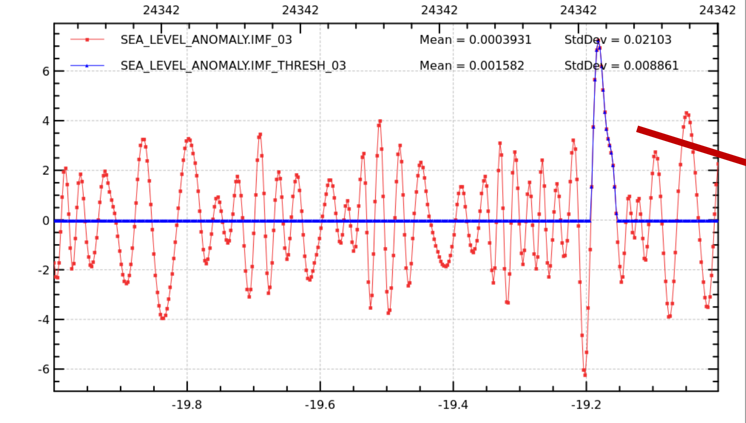
- Sharp angles between smooth curve and peaks
- Peaks associated to retained IMF peaks



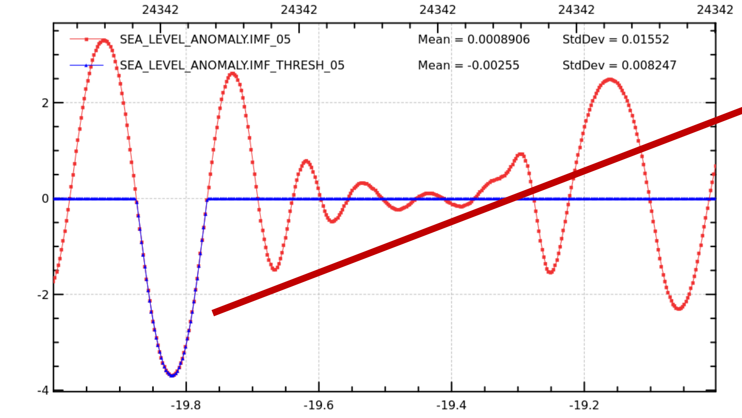
# Reconstructed signal : IT method

- IT method has no replicability issues (no noise shuffle)
- However, issues with smoothness of reconstructed signal
- Problematic for SSH => geostrophic velocities
- With IIT, reconstructing signal from 20 EMD replicas averages out such discontinuities

**IMF3 (red=raw, blue=denoised)**

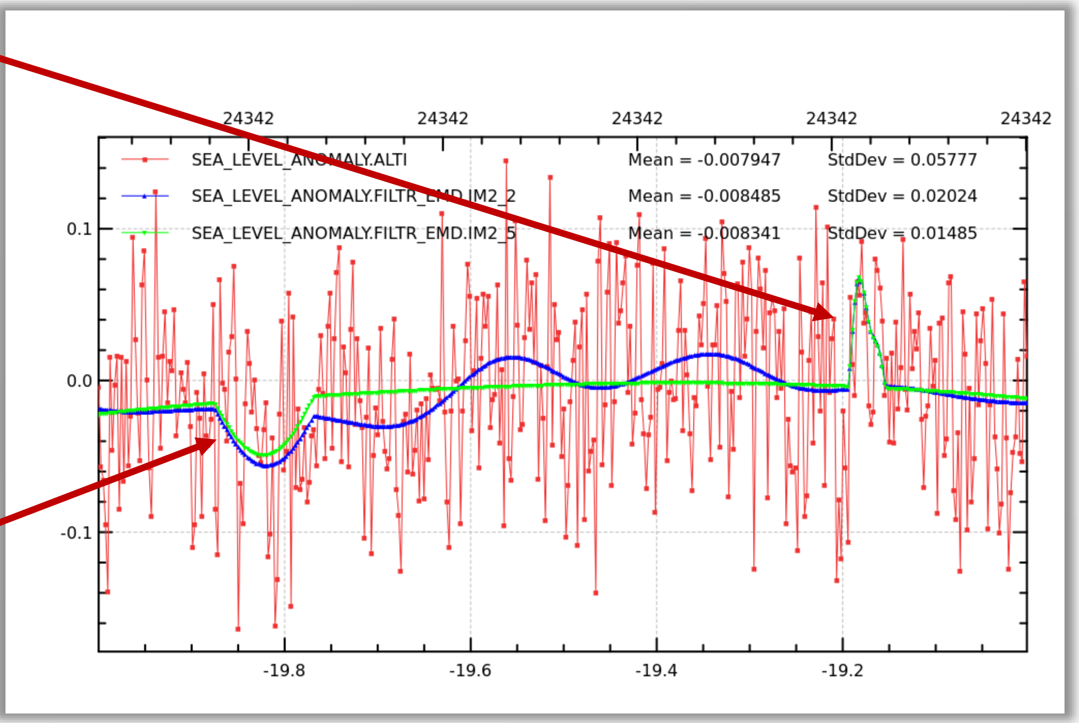



**IMF5 (red=raw, blue=denoised)**



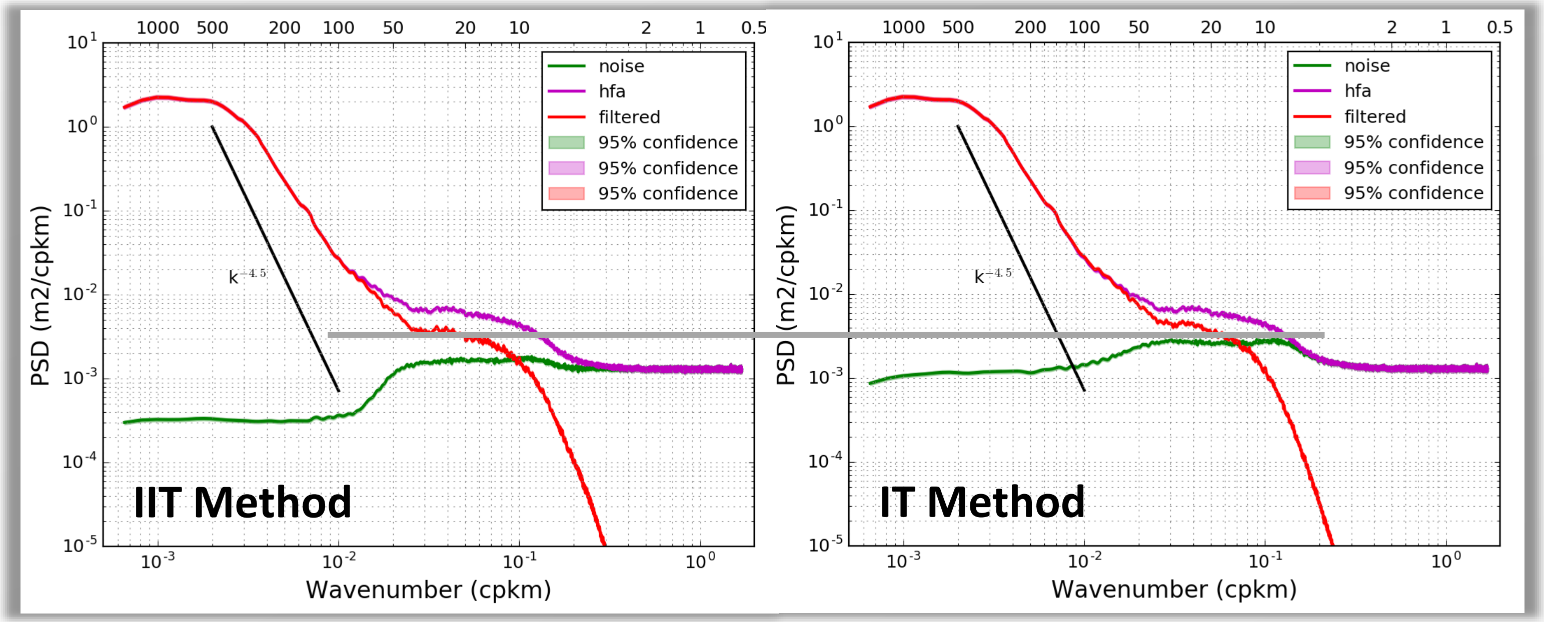

## Results (along-track)

- Sharp angles between smooth curve and peaks
- Peaks associated to retained IMF peaks



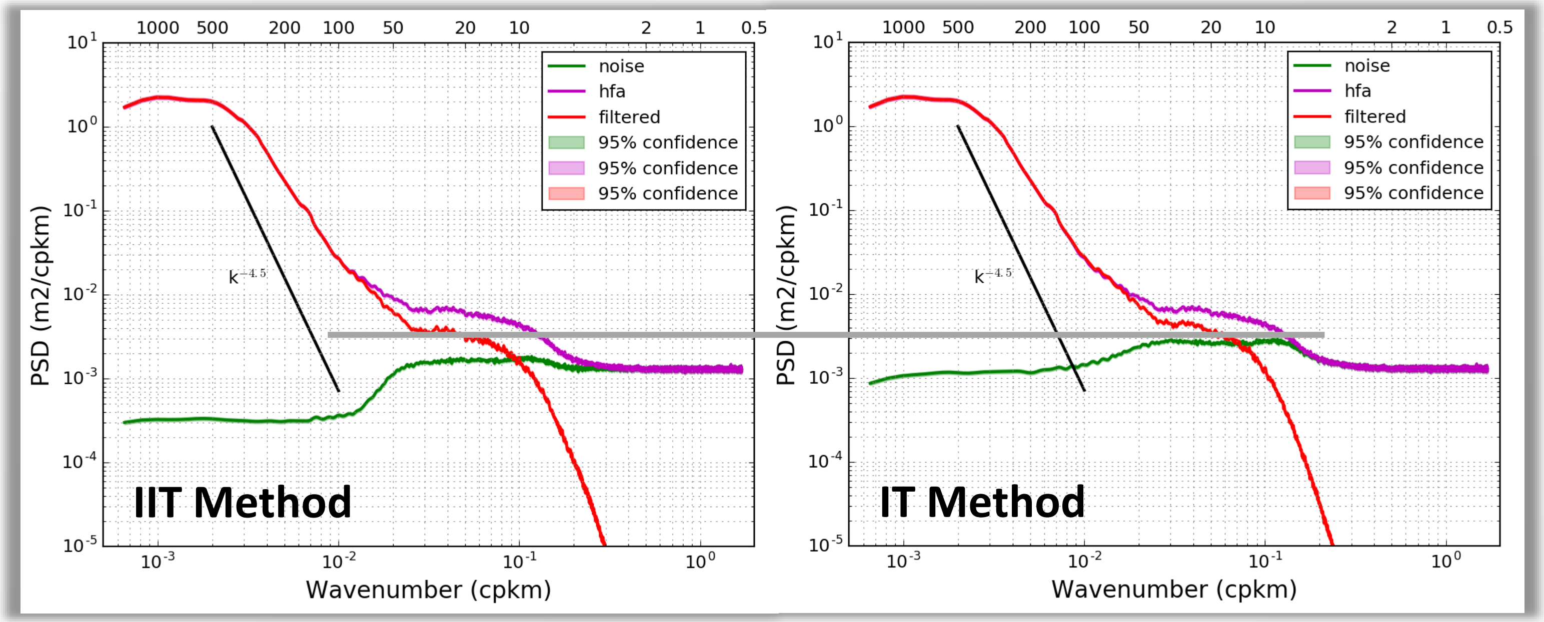
# EMD: Noise/Signal covariance

- Filtered IIT and IT signals similar but not the same
- IIT noise lower than IT in the spectral bump region



# EMD: Noise/Signal covariance

- ❑ Filtered IIT and IT signals similar but not the same
- ❑ IIT noise lower than IT in the spectral bump region




Noise computed as signal – filtered

However, spectra of a sum of signals is different than sum of their spectra if there is covariance between the two signals

- ➢ Noise removed by IIT methods is more correlated to filtered signal than with IT method (Something to do with averaging iterations from noise-shuffling)
- ➢ Part of signal removed?

# Conclusions

## Spectral Hump

- EMD method very good at removing high frequency noise
- Ineffective for mitigating spectral hump if used alone, but good results in synergy with HFA correction

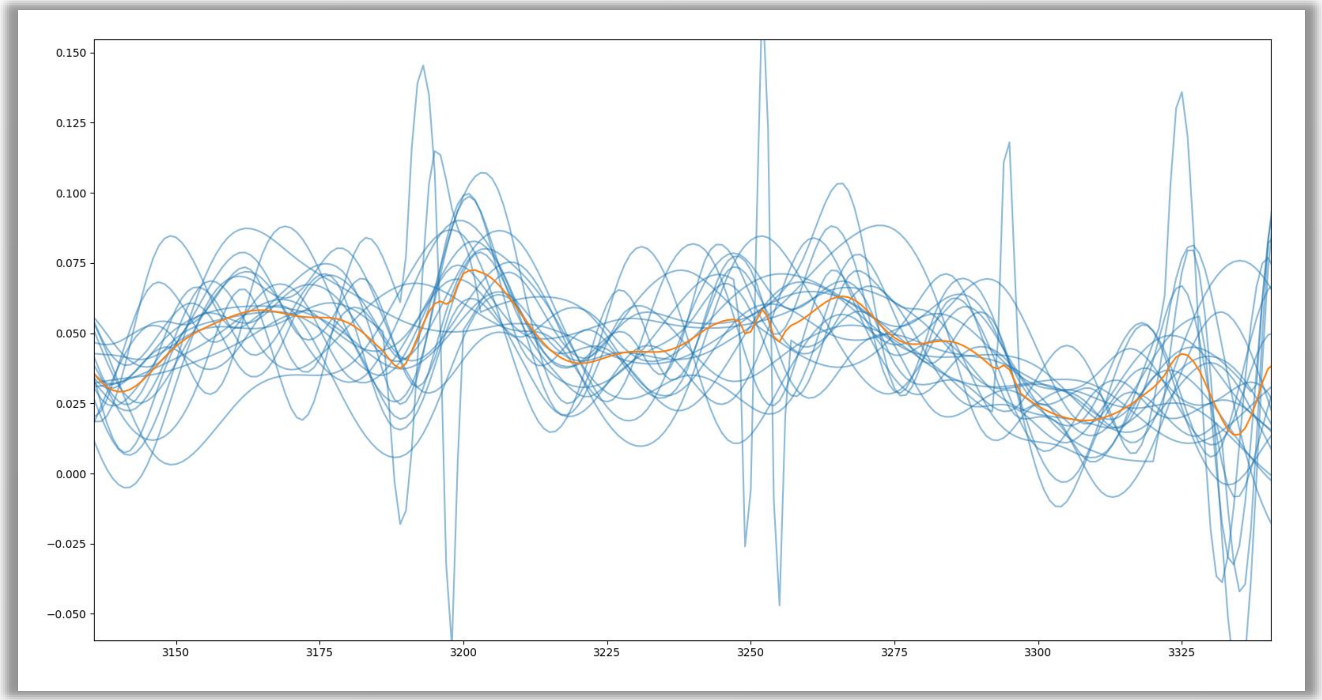
## Method Specifications

- At 20Hz, first IMF can be considered as full noise (no need of wavelet denoising)
- Filter segments of the same track individually might introduce artefacts in the reconstructed signal => fill gaps to filter single tracks as a whole
- IIT method provides a smoother reconstructed signal, but has replicability issues due to noise shuffling (and potential issues with signal/noise covariance to be further explored)
- IT method avoids that, but the reconstructed signal can be characterized by sharp angles where isolated peaks occur
- Lots of paramteres/methods to test and combine => hard to identify the optimal set without a ground truth signal to compare the filter results against

# Replicability issue: effect of noise shuffling

## 20Hz Data

- Example of reconstrcuted signal (orange) from averaging 20 EMD denoised signal from noise shuffling (blue)




- Average signal from 20 EMD répétitions average out the discontinuities
- Smoother denoised curve (orange)
- Nonetheless, some artifact still persist

## ➢ Reccomendations

❑ Noise shuffling is likely needed to obtain a smooth reconstructed signal

❑ To be defined:

- How many iterations are needed to have replicable results?
- How much will a higher number of iterations increase computation time?

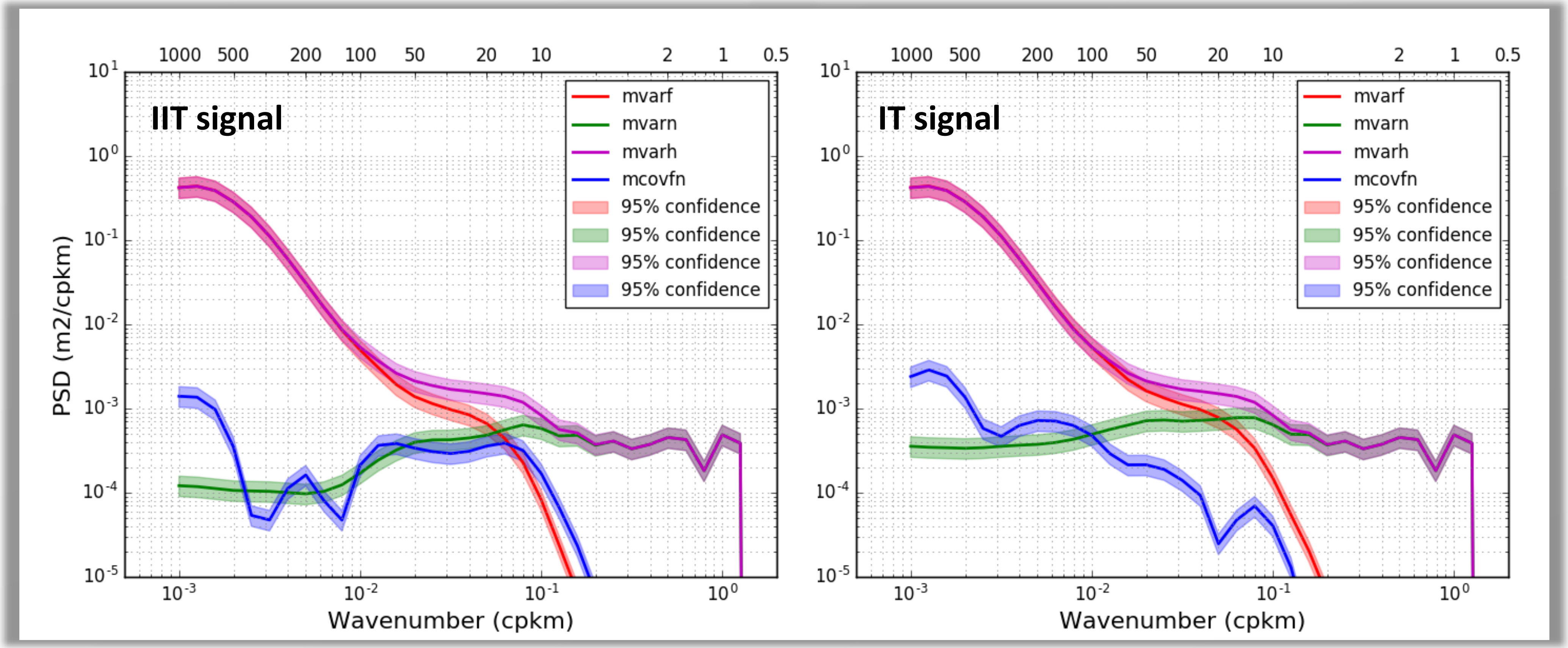
CLS
COLLECTE LOCALISATION SATELLITES

# Spectral analysis: covariance between signals

Noise computed as hfa signal – filtered

However spectra of a sum of signals is different than sum of their spectra if there is covariance between the two signals

hfa, EMD-filtered and noise signal filtered with Lanczos pass-band to compute variance and covariance (**blue**) at varying wavelengths



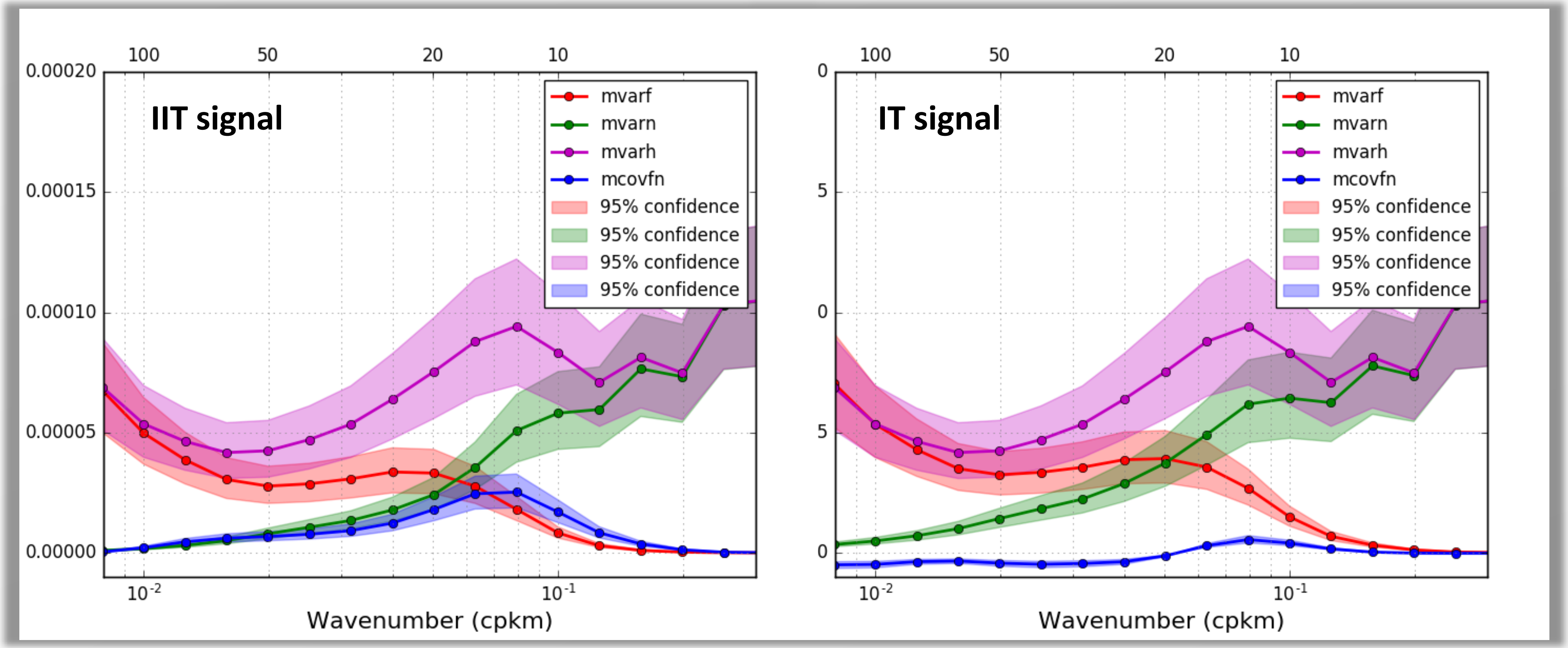

- ❑ Variance and covariance multiplied by wavelength to make it analogous to energy density spectra
- ❑ Since plots are in loglog scale, abs(covariance) was used

# Spectral analysis: covariance between signals

Noise computed as hfa signal – filtered

However spectra of a sum of signals is different than sum of their spectra if there is covariance between the two signals

hfa, EMD-filtered and noise signal filtered with Lanczos pass-band to compute variance and covariance (**blue**) at varying wavelengths




By zooming into the spectral bump region and moving into linear scale:

- Noise removed by IIT methods is more correlated to filtered signal than with IT method
- Something to do with averaging iterations from noise-shuffling