TESTING THE PERFORMANCE OF OBSERVATIONAL SYSTEMS WITH ENSEMBLE METHODS

P. De Mey (1), N. Ayoub (1), J. Lamouroux (2), M. Le Hénaff (3)
(1) LEGOS, Toulouse, France, pierre.de-mey@legos.obs-mip.fr
(2) NOVELTIS, Toulouse, France
(3) RSMAS/University of Miami, Miami, USA

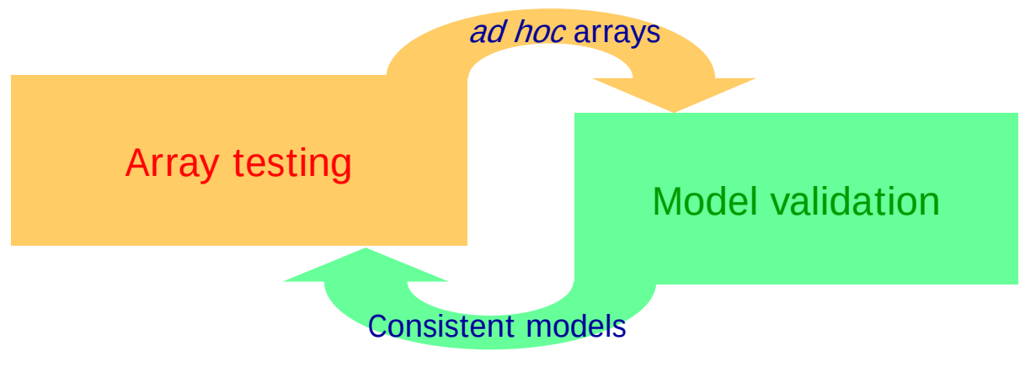
Abstract – In this poster, we address the question of how the performance of multisensor, space- and marine-based observational systems at adding value on top of pre-existing knowledge and of prior state estimates can be characterized. This is explored on a theoretical point of view, with the objective of helping the design of observational systems, and helping sponsors make decisions about them. This study is part of the "Multisensor Impact assessment in Coastal and Shelf Seas" (MICSS) OST project selected by CNES and NASA in 2008 (PI: P. De Mey; Co-Is: N. Ayoub, F. Birol, J. Lamouroux and F. Lyard).

1. Array testing with estimation methods

Preliminary comment 1: Data assimilation provides a useful theoretical framework for array design

- We know our models are wrong, but observations are not truth!
  - The transfer function of an instrument can critically depend on environmental factors (e.g. particle traps) or electronics
  - The time response of an instrument can contain transients which can be impossible to identify as such (e.g. altimeter range trackers)
  - Measurements do not always observe the same processes as the model (= representation errors)
  - Measurement errors are often not well known (they contain drifts, biases, saturation, and their models are imperfect)
  - Observation operators are sometimes complex engineering models by themselves (this is the case of most remote sensing data)
  - Observation operators can be imperfect (e.g. in their handling of subgridscale processes)
  - Etc.
- In matching imperfect models with imperfect observations, data assimilation provides a useful theoretical framework
  - Framework useful beyond assimilation proper – e.g. for array design

Preliminary comment 2: Array testing and model consistency testing are interdependent

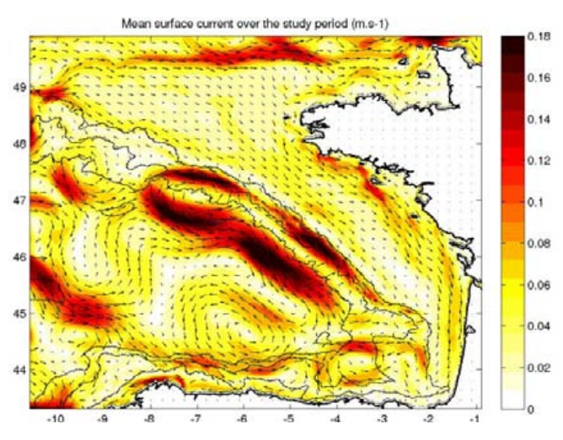


- One reasonable criterion for array design is to ensure a fair detection of model errors, for model validation & assimilation
- In turn, models which are meant to benefit from those observations must be realistic and provide estimates consistent with observations
- In theory, it would make sense to develop both components together

2. Modelling configuration

2.1. Bay of Biscay (BoB) configurations

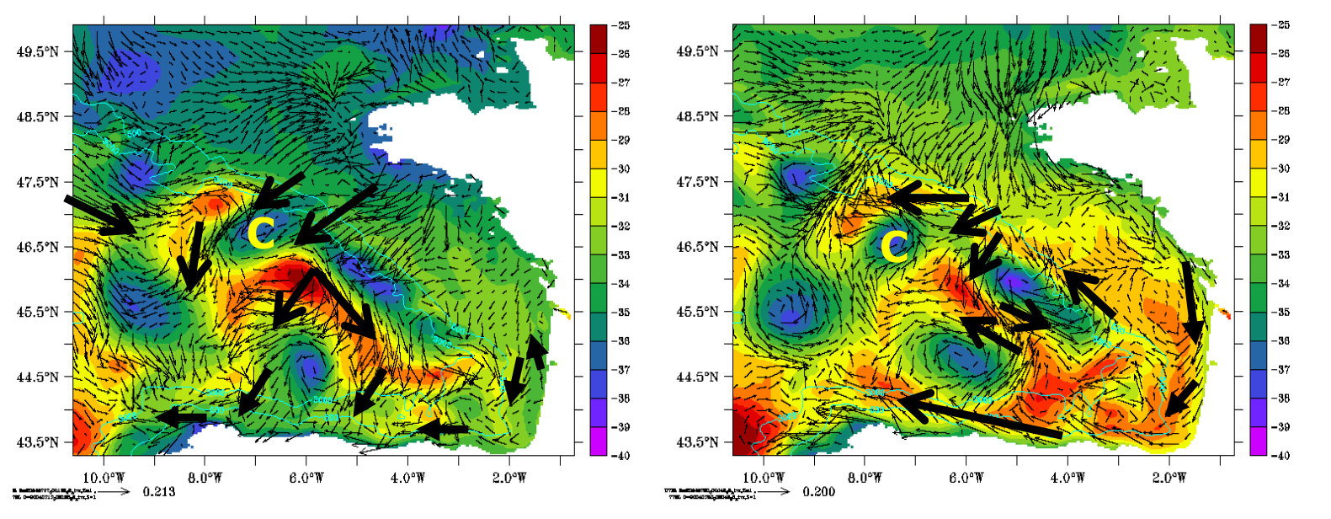
- SYMPHONIE 3DFD, 3-km horizontal resolution, free surface, sigma-step vertical scheme (41 levels max), major river runoff, tidal friction (2004) or tides (2008)
- 3-hourly ALADIN wind (+ atm. surface pressure in 2008)
- Open boundaries, downscaled from MERCATOR PSY2v3 (1/12°)
- Dominant circulation features:
  - Cyclonic slope circulation, anticyclonic recirculation
  - Mesoscale activity above abyssal plain
  - Coastal upwellings (e.g. Galicia)
  - HF processes (shelf/shelf break)



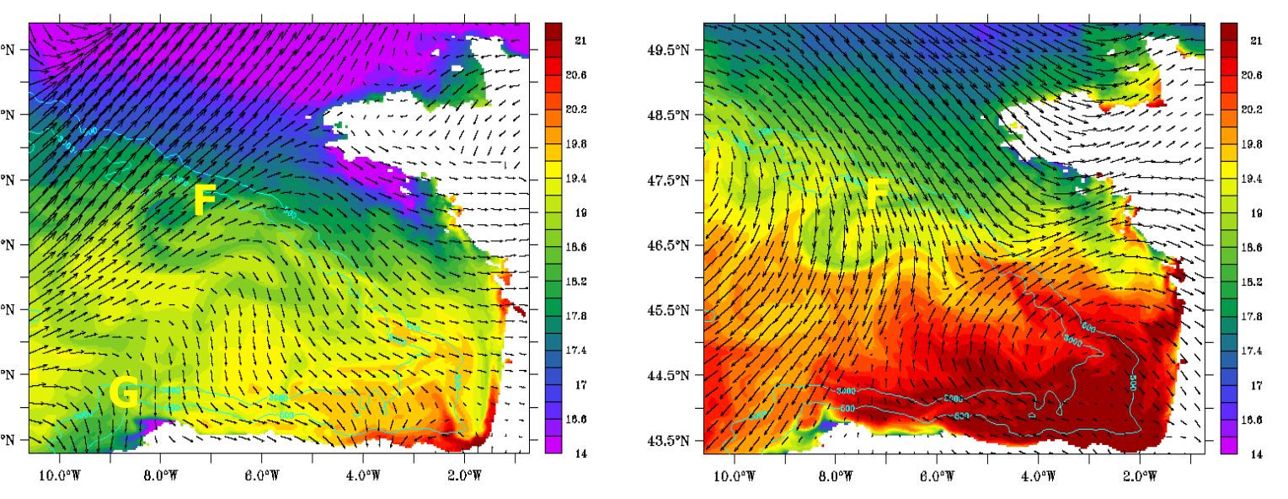
BoB mean surface currents in July-August 2004 & bathymetry

2.2. Typical features of the variable surface circulation (2004)

Sea-Level Anomaly (SLA) and Residual Surface Currents



Sea-Surface Temperature (SST) and Wind



July 17, 2004 July 30

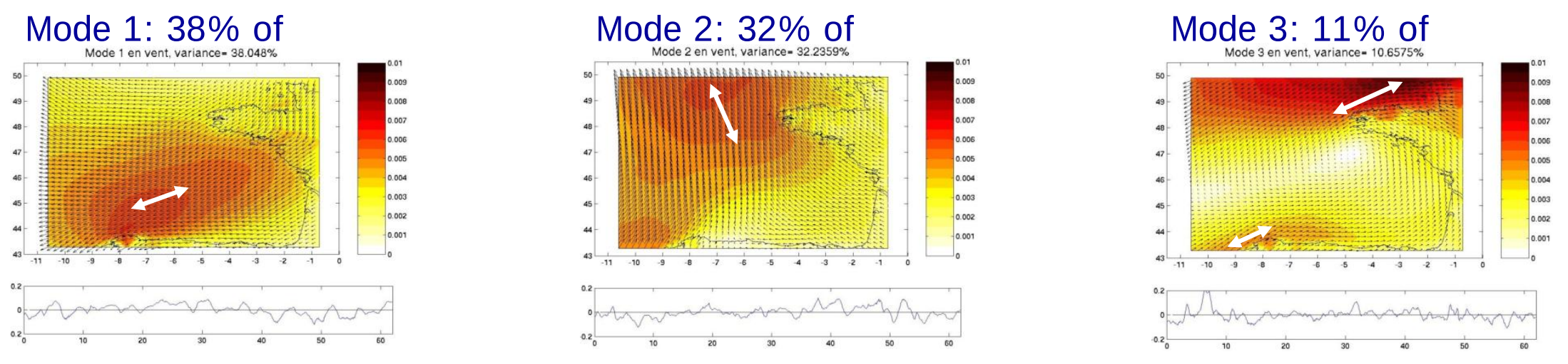
2.3. Ocean Ensemble generation

Assumptions on state error sources
Wind stress + pressure
Bathymetry
River runoff
Turbulence (meso, mixing)
Large-scale circulation
Initial/boundary conditions

→Perturbation strategy
→Ensemble generation

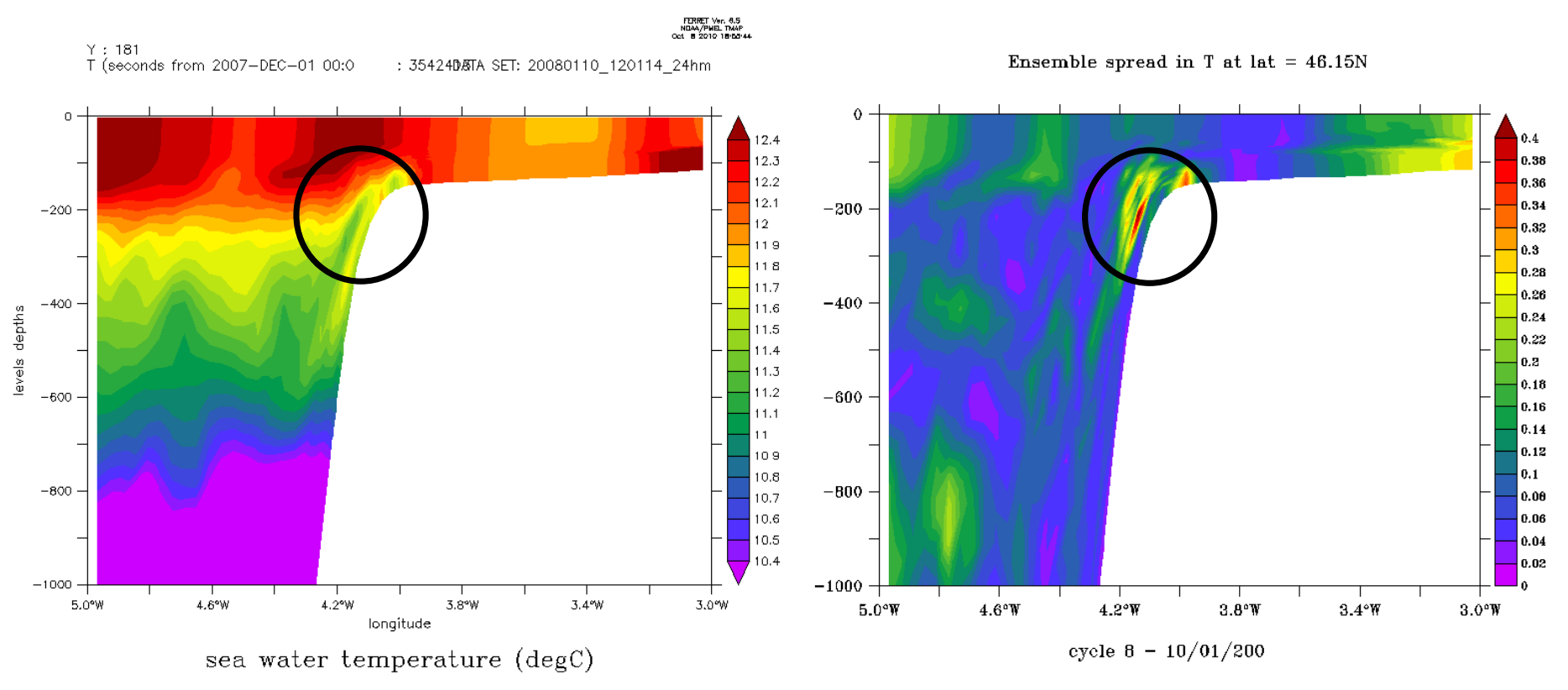
Generation example with wind velocity errors (2004):

- Generate samples of surface atmospheric variables by randomly combining 10 bivariate ($\mathbf{U}_w$) variability EOFs (Auclair *et al.*, 2003)
- One set of Gaussian random coefficients every 5 days
- Integrate ocean members, providing samples of oceanic and atmospheric surface variables



The Ensemble is meant to describe:

- Uncertainties regarding some processes in the model (in response to perturbations)
- Uncertainties associated with modelling errors or inadequacies of the numerical schemes → e.g. instabilities linked to tracer inversion near the shelf break in this version of SYMPHONIE



3. Ensemble-based Representer Analysis

3.1. A simple problem

$\mathbf{x}$ augmented state vector $(n,1)$ over time interval of interest

(let me insist on the fact that this is an augmented state vector – everything that will be shown in this talk includes time as well as space in the definition of observations and prior state estimate)

$\mathbf{y}^o$ observations $(p,1)$ verifying $\mathbf{y}^o = H(\mathbf{x}^t) + \varepsilon$, with:

$H(\,)$ observation operator (not necessarily linear, but use linearized version)
$\varepsilon \in N(0,\mathbf{R})$

Q: how can we characterize the performance of an array $(H, \mathbf{R})$?

Assume we have a prior state estimate of $\mathbf{x}$ and associated error statistics (if not, any observational array will bring valuable information proportionately to its cost):

$\mathbf{x}^f = \mathbf{x}^t + \eta$, with:

$\eta \in N(0,\mathbf{P}^f)$

3.2. What information does the array bring in?

Incremental information brought in by the observations (on top of prior):

Innovation vector $\mathbf{d} \equiv \mathbf{y}^o - \mathbf{y}^g = \mathbf{y}^o - H(\mathbf{x}^f) \approx \varepsilon - \mathbf{H}\eta$

The 2nd-order statistics of $\mathbf{d}$ can be used to characterize the amount of discrepancy brought in by the observational array (on top of prior):

$\langle \mathbf{d}\mathbf{d}^T \rangle = \mathbf{R} + \mathbf{H}\mathbf{P}^f\mathbf{H}^T$, with:

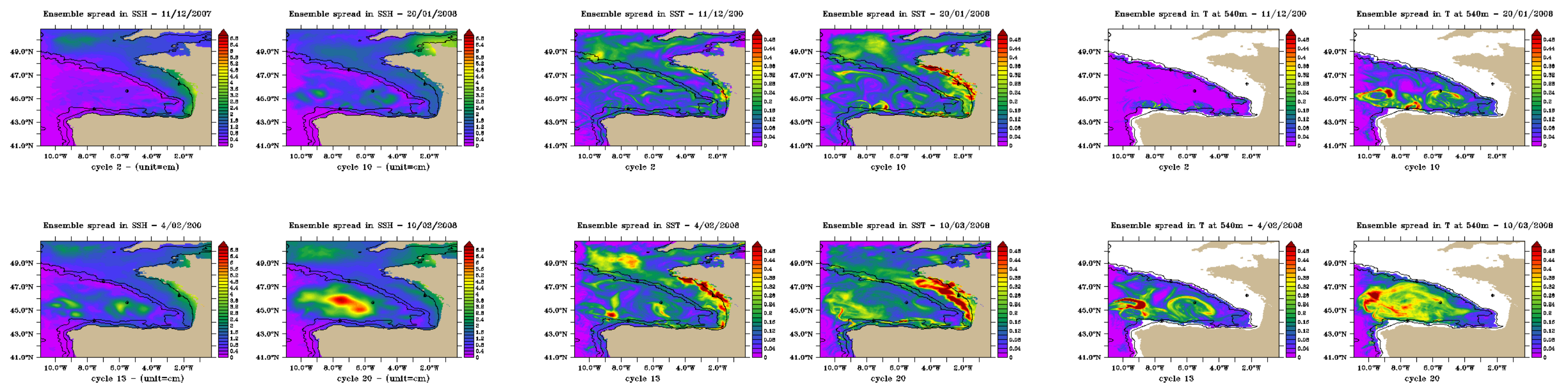
$\mathbf{H}\mathbf{P}^f\mathbf{H}^T$ Representer matrix : prior state error covariance in observational space
$\mathbf{P}^f\mathbf{H}^T$ Matrix of representers : provide extrapolation from observational array

➔ Representers contain information on how observations are able to detect prior state error, and constrain an "optimal" solution through extrapolation:

- Extrapolation in space and time
- Extrapolation across variables (in particular the unobserved ones: multivariate character)

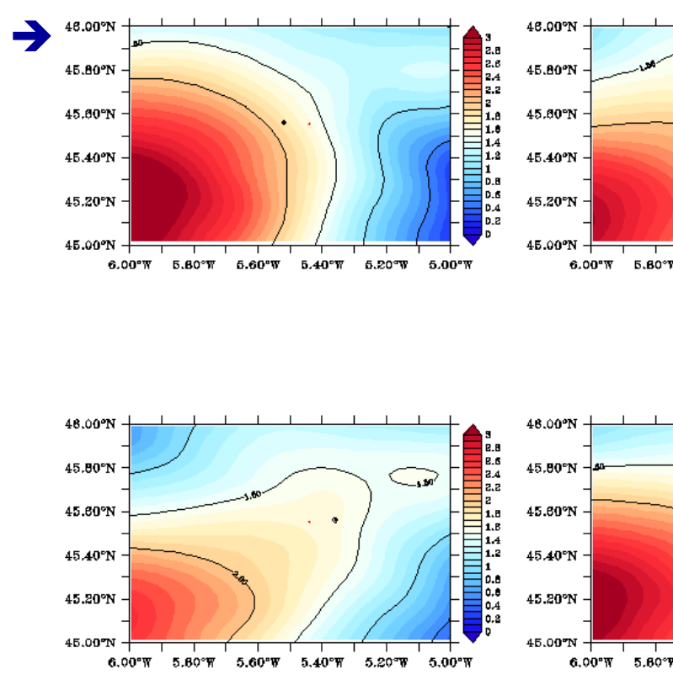
3.3. Ensemble spread as a function of time: SSH, SST, T540

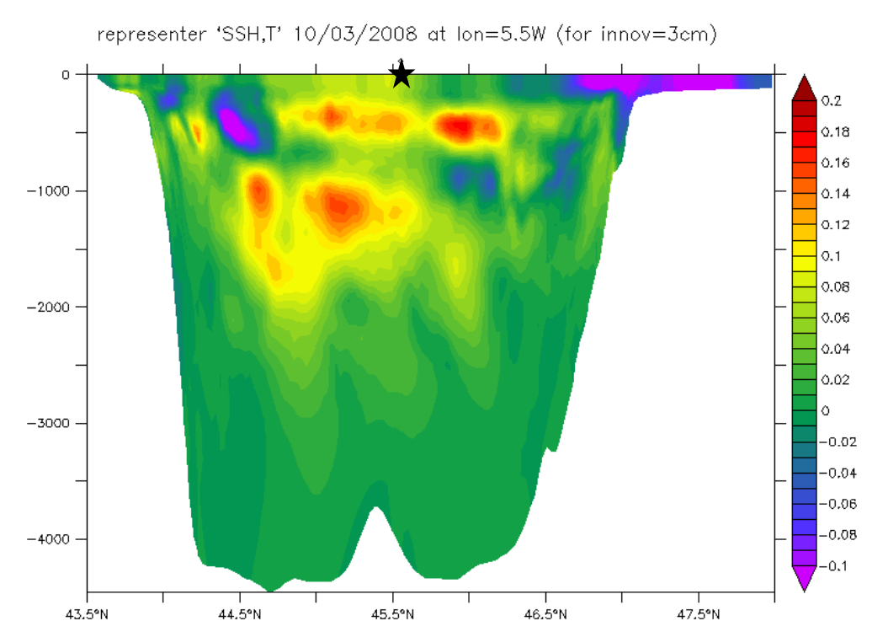
- Wind velocity errors
- Structures slowly fill up above the abyssal plain, in particular sprouting from the North Iberian shelf
- The response on the shelf is more quickly established and more time dependent



3.4. Representers of SSH above the abyssal plain

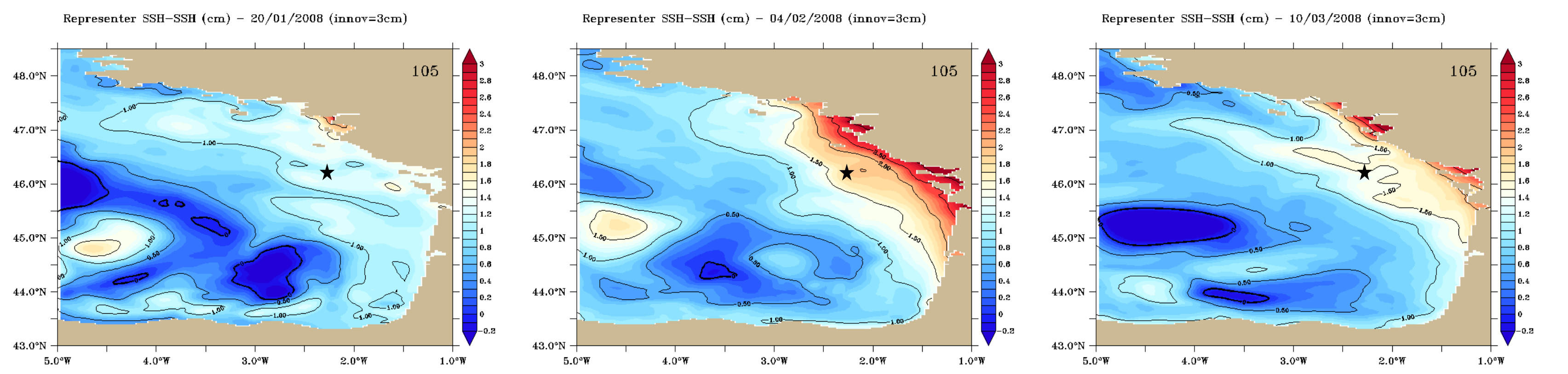
- Potential impact on subsurface variables – both on the main thermocline, and on the thermostat-like depth range around 500m depth (analysis in progress)
- Local SSH-SSH representers of SSH measurements 6km apart show that the potential impact of high-resolution altimetry (akin to what SWOT would provide) contains:
  - The high-resolution information contained in the signal itself (not shown here, obviously)
  - The spatial variability of the influence functions (shown on right panel) ➔





3.5. Representers of SSH on top of the South Armorican shelf

- Mostly a shelf-wide response (correlations with abyssal plain variables are probably artefacts)
- High temporal variability



4. Array performance and design: RMSpectrum analysis

4.1. A qualitative/intuitive criterion of array performance

The 2nd-order statistics of innovation $\mathbf{d}$ can be used to characterize the amount of discrepancy brought in by the observational array on top of the prior state estimate:

$\langle \mathbf{d}\mathbf{d}^T \rangle = \mathbf{R} + \mathbf{H}\mathbf{P}^f\mathbf{H}^T$

Qualitative/intuitive criterion of array performance:

- $\mathbf{R}$ "dominates"
  → most of the discrepancies are attributable to observational error
  → observations are not very useful
- $\mathbf{H}\mathbf{P}^f\mathbf{H}^T$ "dominates"
  → most of the discrepancies are attributable to prior state errors
  → observations can be used to identify and correct prior state errors

4.2. Towards a formal criterion of array performance

Two paths (among others) to formalize the intuitive order relationship…

Bennett's "array modes" (e.g. Bennett et al., 1997): these are orthonormal rotation vectors $\boldsymbol{\beta}$ obtained by diagonalizing the representer matrix:

$\mathbf{H}\mathbf{P}^f\mathbf{H}^T = \boldsymbol{\beta}\boldsymbol{\lambda}\boldsymbol{\beta}^T$

$\boldsymbol{\beta}$: observable degrees of freedom of the physical system for that configuration
$\boldsymbol{\lambda}$: spectrum of RM, to be compared to the diagonal of $\mathbf{R}$ (obs. noise floor)

Le Hénaff & De Mey (Le Hénaff et al., 2009): in the general case of non-homogeneous, non-diagonal $\mathbf{R}$, and observational samples scattered in time, space, and across variables, use spectrum $\boldsymbol{\sigma}$ and array modes $\boldsymbol{\mu}$ of the scaled representer matrix $\boldsymbol{\chi}$:

$\boldsymbol{\chi} = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{P}^f\mathbf{H}^T\mathbf{R}^{-1/2} = \boldsymbol{\mu}\boldsymbol{\sigma}\boldsymbol{\mu}^T$

$\boldsymbol{\mu}$: spectrum of SRM, to be compared to the diagonal of $\mathbf{I}$ (obs. noise floor)

Modal representers $\boldsymbol{\rho}_\mu = \mathbf{P}^f\mathbf{H}^T\mathbf{R}^{-1/2}\boldsymbol{\mu}$ = representers for the array modes

4.3. Stochastic implementation of RM analysis

$$Matrix of\ samples: \mathbf{A}^f \quad (e.g.\ forecast\ Ensemble\ anomalies)$$

$$\mathbf{S} = \frac{1}{\sqrt{m-1}}\mathbf{R}^{-1/2}\mathbf{H}\mathbf{A}^f \quad (Sakov\ et\ al., 2010)$$

$$\hat{\mathbf{P}}^f = \frac{1}{m-1}\mathbf{A}^f\mathbf{A}^{fT}\ estimate\ of\ \mathbf{P}^f$$

$$\hat{\boldsymbol{\chi}} = \frac{1}{m-1}(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{A}^f)(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{A}^f)^T = \mathbf{S}\mathbf{S}^T\ estimate\ of\ \boldsymbol{\chi}\ matrix$$

$$\langle \mathbf{d}\mathbf{d}^T \rangle = \mathbf{R}^{1/2}(\hat{\mathbf{F}} + \mathbf{I})\mathbf{R}^{1/2} = \mathbf{R}^{1/2}(\mathbf{S}\mathbf{S}^T + \mathbf{I})\mathbf{R}^{1/2}\ basis\ of\ original\ criterion$$

We now have the following stochastic estimates:

$\hat{\boldsymbol{\sigma}}$ = RM spectrum = squares of the singular values of $\mathbf{S}$

$\hat{\boldsymbol{\mu}}$ = Array Modes = singular vectors of $\mathbf{S}$

Modal Representers = $\hat{\boldsymbol{\rho}}_\mu = \frac{1}{\sqrt{m-1}}\mathbf{A}\mathbf{S}^T\hat{\boldsymbol{\mu}}$

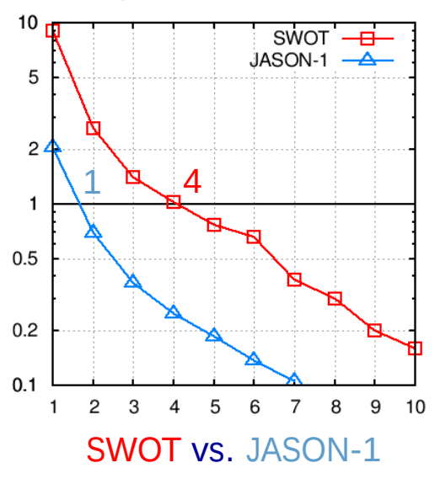
4.4. Stochastic RM spectrum analysis in practice

- RM analysis levels
  - Just count eigenvalues above 1 → useful to convince sponsors & decision-makers
  - Explore array modes & modal representers → scientific analysis
- Origin of ensemble samples
  - From stochastic modelling → array performance results do not depend on assimilation configuration and history (sometimes easier to sell)
  - From an EnKF → online analysis allow to study array performance through regime changes, error estimates are typical of an assimilating system
- Assumptions on prior state error sources
  - Choose to perturb wind stress, surface pressure, bathymetry, river runoff, turbulence (mesoscale, mixing), large-scale circulation, initial/boundary conditions, etc.
  - → Comes back to prioritizing what the array is designed for

4.5. Wide-swath vs. nadir altimeter

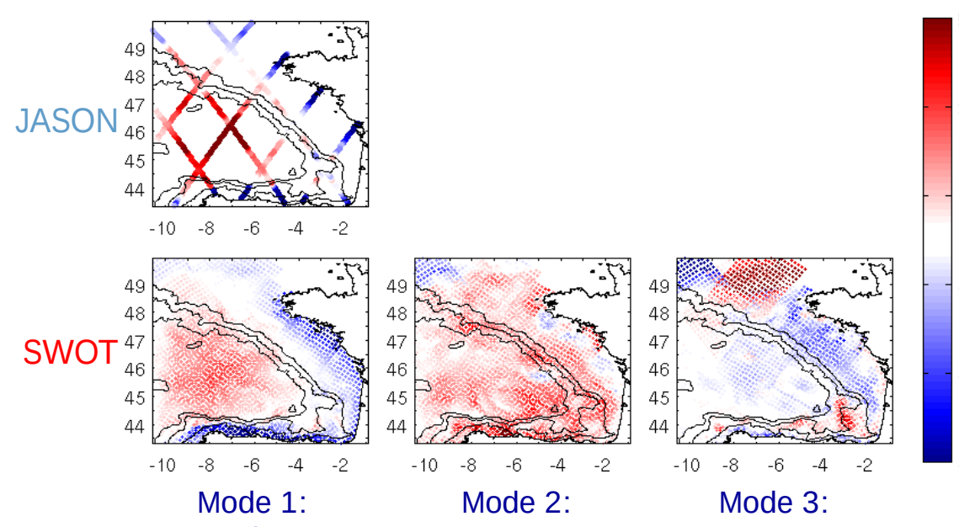
- We compare the performance of the JASON nadir altimeter with SWOT on the JASON orbit
- Only SWOT appears able to usefully detect & constrain coastal mesoscale patterns (array modes 2 and 3) and high-frequency events on the shelf (array mode 3)

RM spectra



SWOT vs. JASON-1

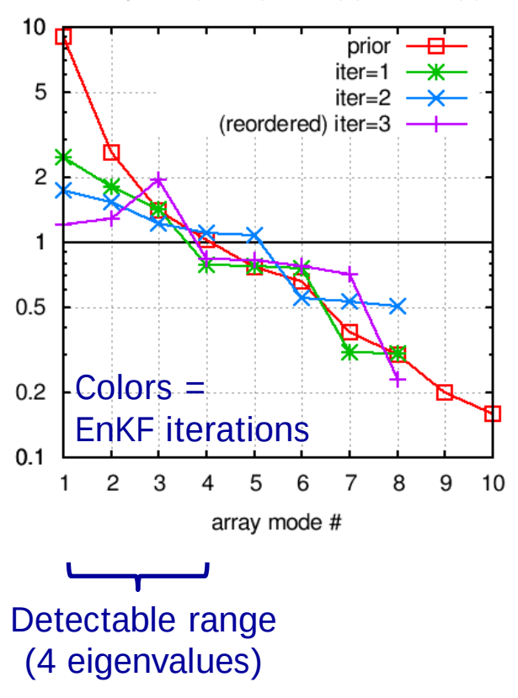
First 3 detectable array modes (SLA)



Mode 1: Swing (common) Mode 2: Meso1 Mode 3: Meso2+HF

4.6. Online RM spectrum analysis with 4-D local EnKF

- Same experimental configuration, but carry out RM analysis online at each 10-day assim cycle (invariant $\mathbf{H}$)
- 4-D local EnKF with BELUGA
- Assimilate simulated SWOT wide-swath altimeter on 10-day orbit for 2 months in summer 2004
- Rank is approximately conserved through assimilation
- Spectra whiten in detectable range
  - Array info is being extracted
  - Mostly large-scale and mesoscale error processes constrained
  - No eigenvalue decrease for high-frequency shelf processes → need for sustained observations of such processes



Detectable range (4 eigenvalues)

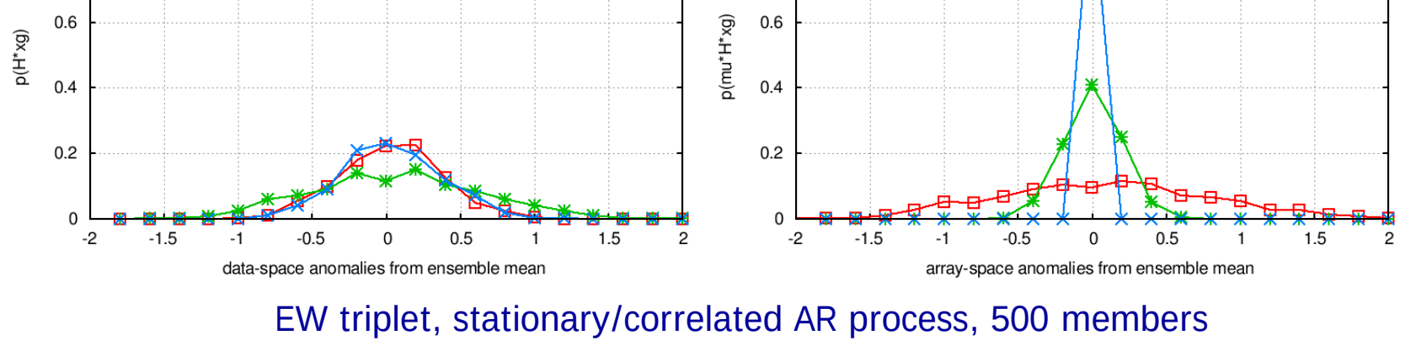
5. Conclusion and outlook

5.1. Conclusions on RM Spectrum analysis

- Stochastic RM analysis is a useful tool for array testing based on the detection of prior errors (Le Hénaff & De Mey, ODyn, 2009)
- Approach provides a recursive way to prioritize array options given a library of ensemble members
- …and, potentially, as a way to hierarchize consistency checking of ensemble forecasts (work in progress), as part of stochastic model testing
- Easily set up online as part of an Ensemble filter, e.g. to study the impact of regime changes on array performance
- Impact analysis can be performed on unobserved variables via modal representers
- Limits
  - Based on detection of forecast errors; controllability checking requires OSSEs
  - Criterion is based on Gaussian pdfs (like most DA schemes)
  - No support yet for systematic errors (biases)
- Complementary with other existing array design approaches: OSSEs, targeted observations, etc.

5.2. Outlook: Can array modes help ensemble consistency analyses? (Work in progress!)

- Problem: check whether probability densities of model forecast and observations are consistent with each other (be it visually, through reliability scores, Bayesian analysis, etc. – not the topic here)
  - Compare pdf's in data space vs. array space
- Low-order array-space forecast pd's have broadest base (by design)
  - Hierarchize ensemble consistency checks from easiest to hardest to pass



EW triplet, stationary/correlated AR process, 500 members