CNRS-LEGOS contribution to SANGOMA

Pierre De Mey
and collaborators

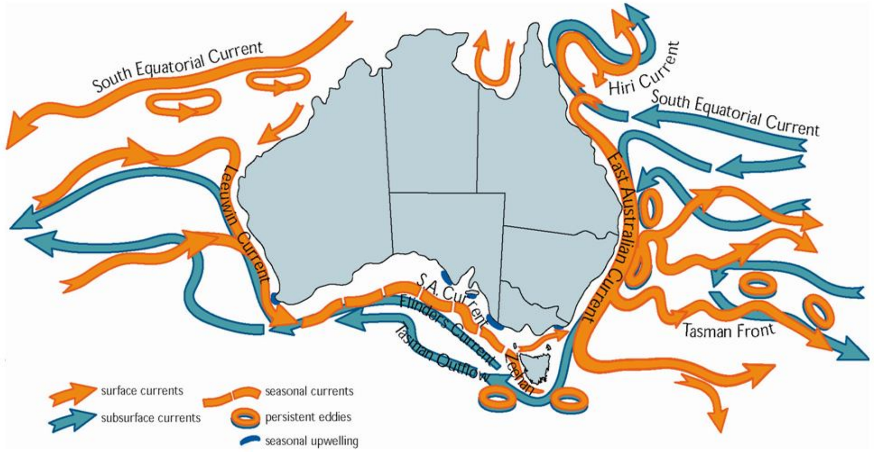
SANGOMA kick-off meeting
November 24-25, U. Liège, Belgium

Outline

- Context: coastal ocean, Bay of Biscay
- SEQUOIA assimilation toolbox & BELUGA EnKF
- RMSpectrum: stochastic array performance assessment
- What we plan to do in SANGOMA
- Surprise ending!

Coastal ocean context

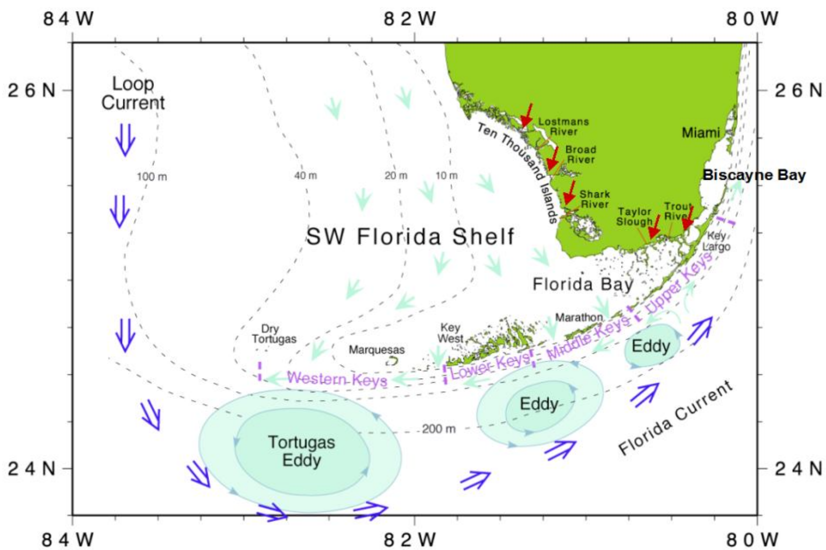
- Coastal-ocean and open-ocean problems are intertwined, errors are multiscale
- Non-: -homogeneous, -isotropic, -linear
- Spectrum of LF (meso, slope currents, plumes, some wave types) and HF (free surface, tides, atmospheric response, upwellings, cascading, other wave types) processes, error dynamics needed



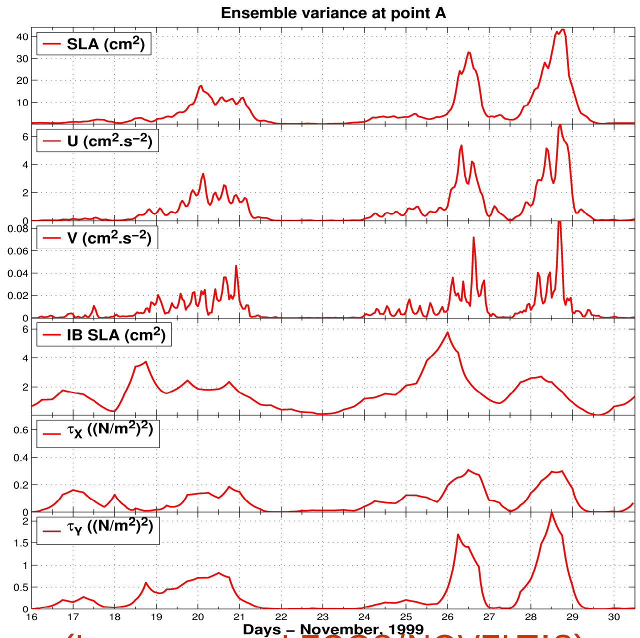
(CSIRO)

Oregon CTZ
SST
24 Aug 2003

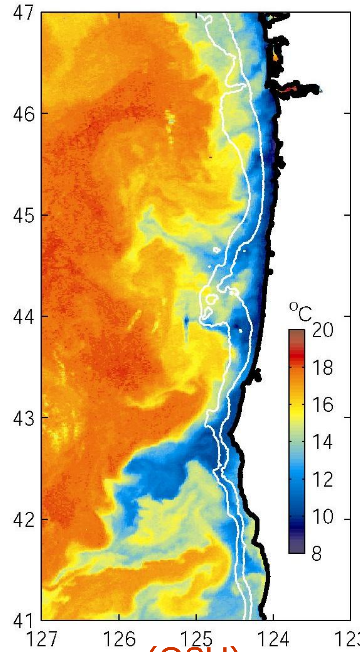
Ensemble variance(t) 11/1999
English Channel



(Kourafalou, RSMAS)



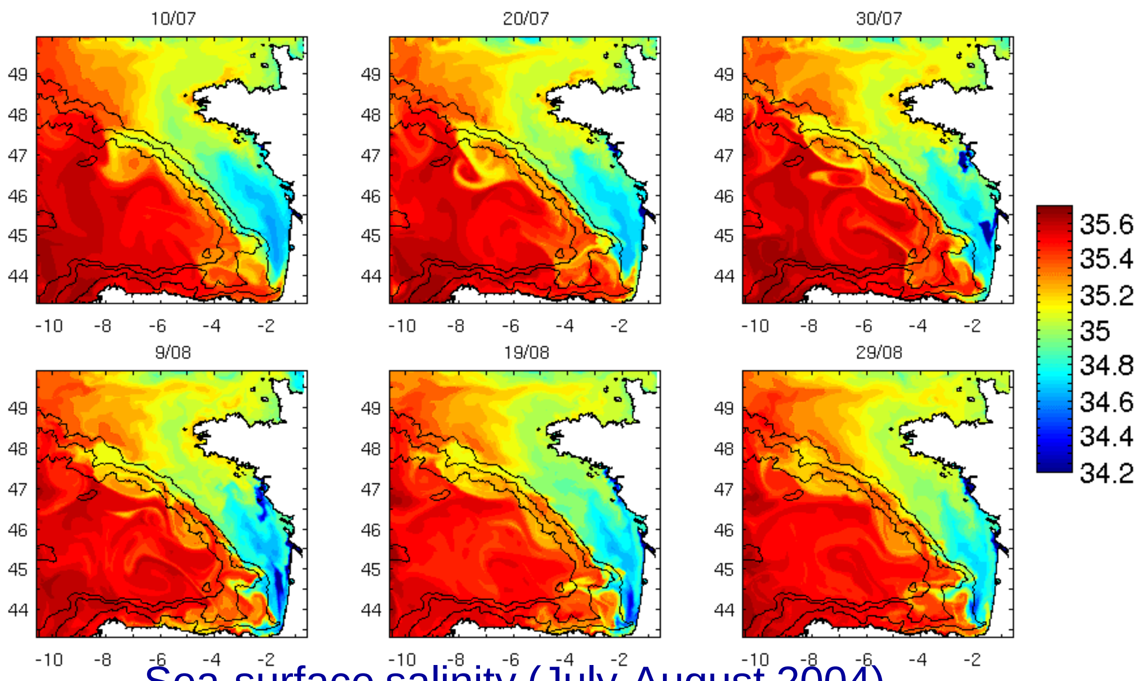
(Lamouroux, LEGOS/NOVELTIS)



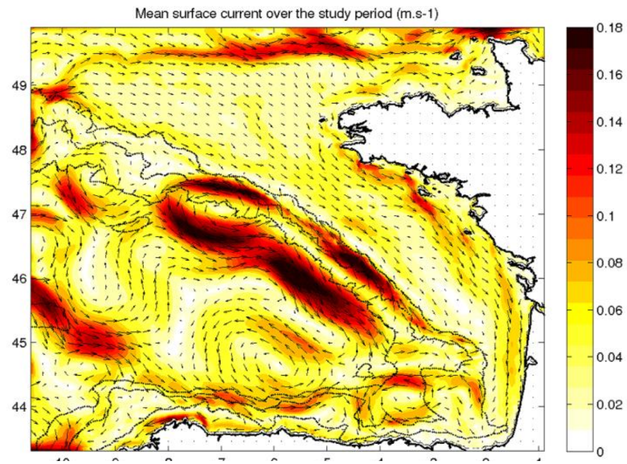
(OSU)

Bay of Biscay (BoB) configuration

- SIROCCO 3DFD, 3-km horizontal resolution, free surface, sigma-step vertical scheme (41 levels max), major river runoff, tidal friction
- Downscaled from MERCATOR PSY2v3 (1/12°) + tides
- Cyclonic slope circulation, anticyclonic recirculation
- Mesoscale activity above abyssal plain



Sea-surface salinity (July-August 2004)



BoB mean surface currents (July-August 2004) & bathymetry

SEQUOIA assimilation toolbox (1/2)

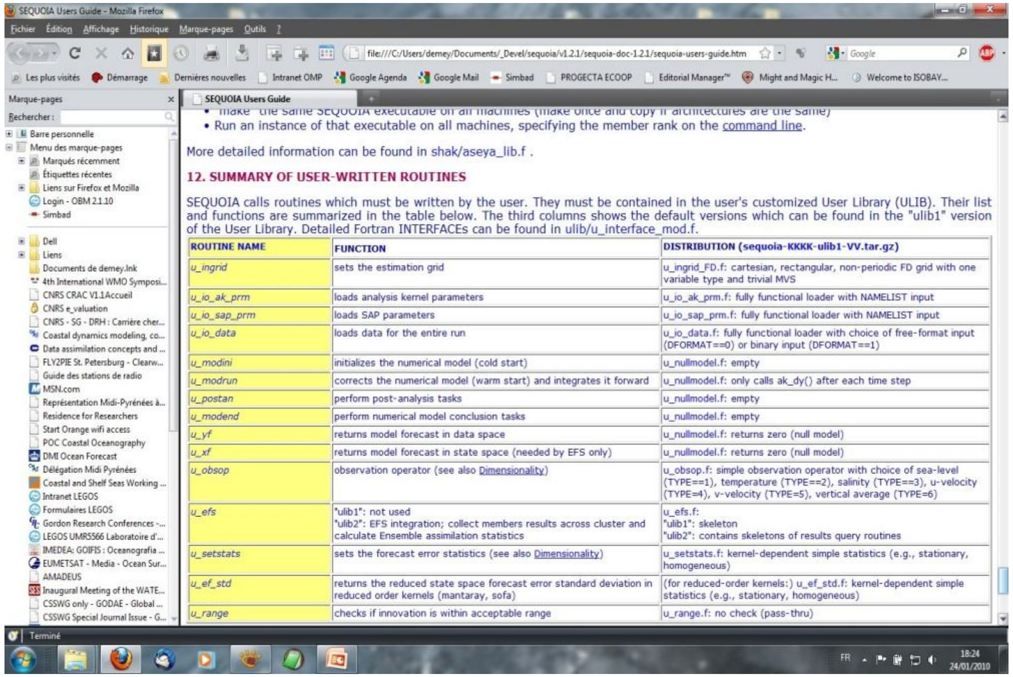
- Modular research code, f95, LGPL, http://sirocco.omp.obs-mip.fr
- Libraries
  - Core libs: common lib, analysis kernel
  - User-space libs: normalised model interfaces, I/Os (no netCDF but normalised data model), caribou post-processing unit
- Common lib:
  - Unstructured horizontal grid, does FD, FE
  - Generic treatment of vertical dimension, does z-, s-levels, etc.
  - Does ensemble DA on multicore or cluster with MPI
  - K-d tree search, interpolation, etc.
- Algebra/algorithms: choice of analysis kernels
  - Cheap, OI-type methods:
    - SOFA kernel: EnROOI-1D *(Fernandez et al., pers.comm., 2008; + usages précédents de SOFA)*
    - MANTARAY kernel: EnROOI-3D (static SEEK) *(Lamouroux et al., OSM2006; Jordà et al., 2010)*
  - Non-linear error dynamics:
    - BELUGA kernel: data-space 4-D EnKF *(Mourre et al., 2004, 2006; Le Hénaff, Lamouroux, Ayoub, 2009, 2011)*

SEQUOIA assimilation toolbox (2/2)

- Data in BELUGA
  - Data-space estimates provided by model interface → no need for explicit observation operators (in standard uses), can assimilate almost whatever can be formed by model – current limit: time averages
  - Observational errors can be correlated in space, time, across variables (representativeness errors usually are)
  - K-d tree search
  - Suboptimality options
    - Localisation
    - Thinning
- Model interface
  - The model (analysis) grid is in user-space library
  - SEQUOIA controls the model via normalised calls: u_modini(), u_modrun(), u_modend(), u_yf()
  - The model makes SEQUOIA calls: ak_dxa() to get the analysis increments, ak_dy() to store the innovation at the current time step

SEQUOIA “kit”

- Code + Makefile
- Hypertext (HTML) Users Guide
- Manual, still incomplete (algorithms)
- Scripts: Beowulf, SLURM, PBS
- MPI-based & non-MPI versions (upon request)
- Test cases

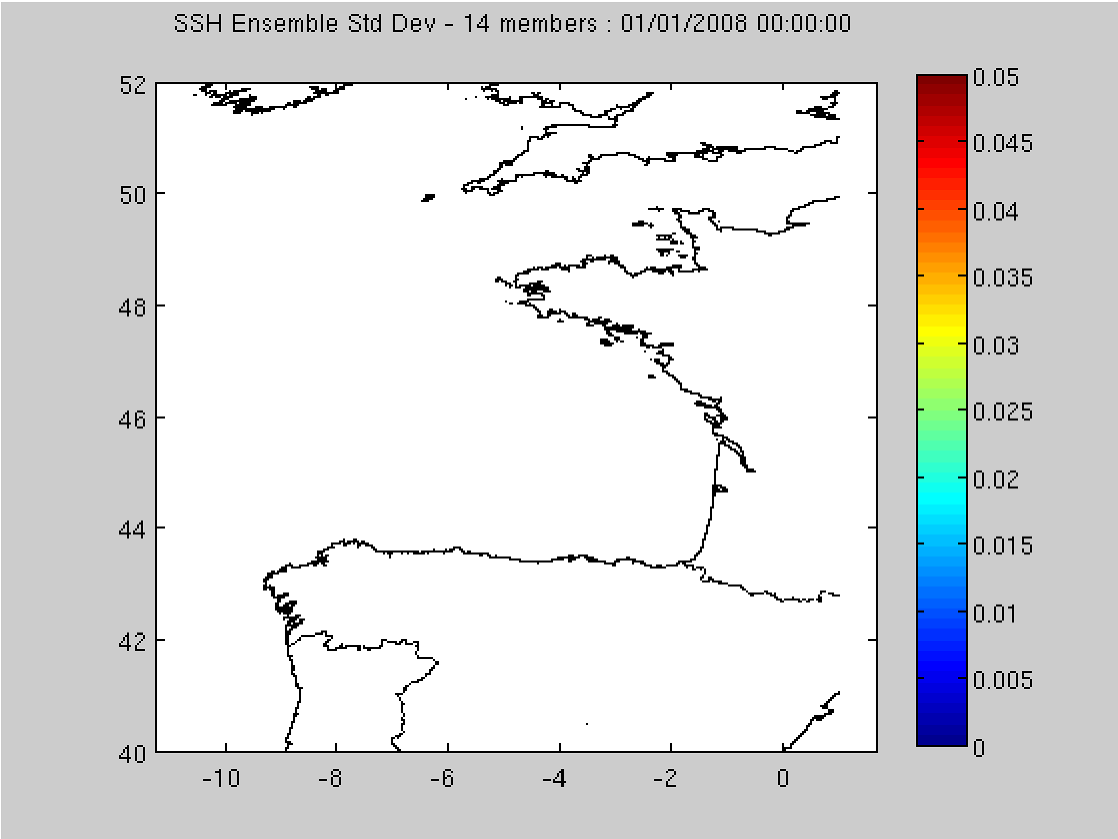


Ensemble generation

- Assumptions on state error sources
  - Wind stress + pressure
  - Bathymetry
  - River runoff
  - Turbulence (meso, mixing)
  - Large-scale circulation (incl. “mean”)
  - Initial/boundary conditions
- Perturbation strategy varies

Ensemble variance(t) in SLA (January 1-31, 2008)

- Error-subspace response to wind errors: mesoscale turbulence, Kelvin/shelf waves, water pile-up on shelves (including English channel, aligned with dominant winds axis)



BoB (Ayoub, Lamouroux et al., 2010)

Community assessment of model error in the BoB: ensemble-based error estimates 1/2

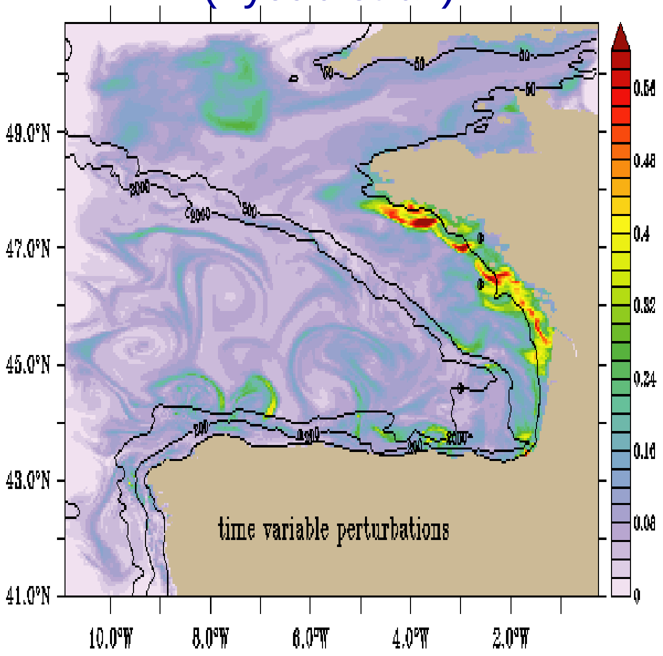
- Examine the response of several regional models to various stochastic perturbations in winter 2007-2008, analyse similarities and differences
- Interpret ensemble spread as error estimates
- Mid-project meeting held at L'Houmeau, France, Sept. 22-23, 2011

Participants	Model configurations	Assimilation
LEGOS / NOVELTIS (De Mey, Ayoub, Lamouroux, Lyard)	• SYMPHONIE 3km BoB + Celtic sea • MERCATOR obc + tides	• BELUGA (AEnKF) • ARM • Data: ALT, SST, ++
SHOM (Baraille, Hoang, Morel)	• HYCOM 1.8km BoB • MERCATOR obc + tides	• Reduced-order scheme based on AF and Schur vectors • Data: ALT, SST
PREVIMER / ACTIMAR (Dumas, Lecornu, Cranéguy, Charria)	• « MANGA »: MARS3D 4km BoB + Celtic Sea + English Channel • MERCATOR obc + tides	• EnKF (NERSC) • ARM (coll. NOVELTIS+LEGOS) • Data: SST, ++
LEGI (Brasseur, Brankart)	• HYCOM, several configs. 1/3°-1/15° incl. BoB	• SEEK, ensemble method, Truncated Gaussian filter • Data: profiles, ++
MERCATOR Océan / CLS / LEGOS (Testut, Benkiran, Quattrocchi, Léger)	• NEATL12 (FACADE) v2 • BISCAY12 • GLORYS1V1 obc + FES2004	• SAM-2 (anomaly-based SEEK) • Stoch. Mod. → EnKF • Data: ALT, SST, profiles

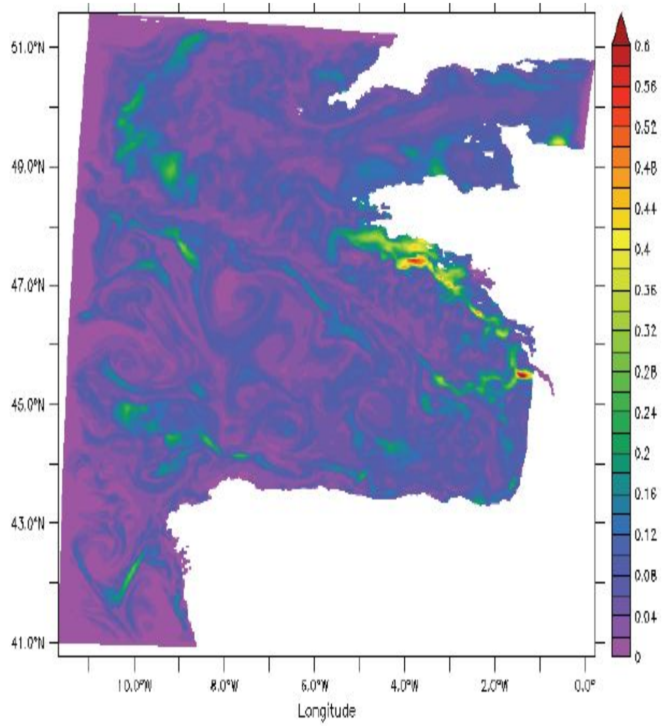
Community assessment of model error in the BoB: ensemble-based error estimates 2/2

SST Ensemble stdev(°C) in response to wind uncertainties

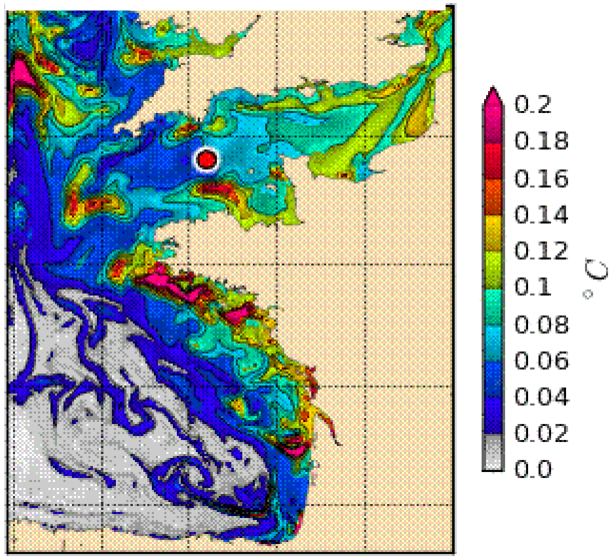
SIROCCO 3DFD/BISCAY 21/01/2008 (Ayoub et al.)



NEMO/BISCAY 21/01/2008 (Quattrocchi et al.)



MANGA 15/01/2008 (Heyraud et al.)

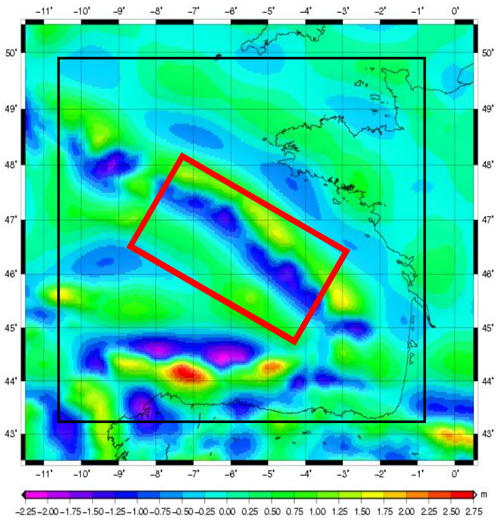


- Specific response on the shelf (intense, faster, small-scale patches)
- Specific response over the abyssal plain (weaker, slower, filament-like)
- (replace “response” by “errors” above, given the specific error source)
- Use error proxies to guide data collection in coastal regions

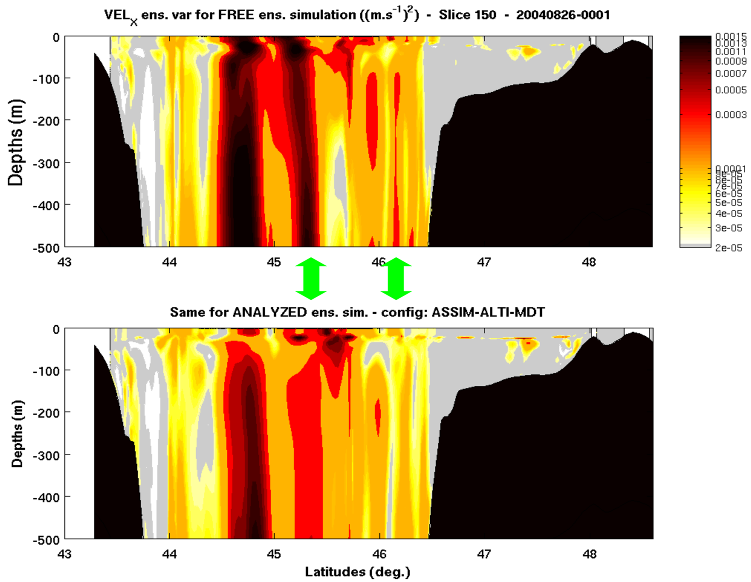
Posterior ensemble variance: Impact of assimilating the GOCE geoid

- The signal associated with the slope current is small and at the limit of the ability of GOCE to detect + the geoid omission errors over the slope are large
  - Use guess of slope current, e.g. model forecast
- Simultaneous assimilation of simulated SLA and GOCE geoid estimate with EnKF
- Reduction of posterior ensemble variance in coastal current system

GOCE omission error estimate



Velocity ensemble variance along N-S section, BoB



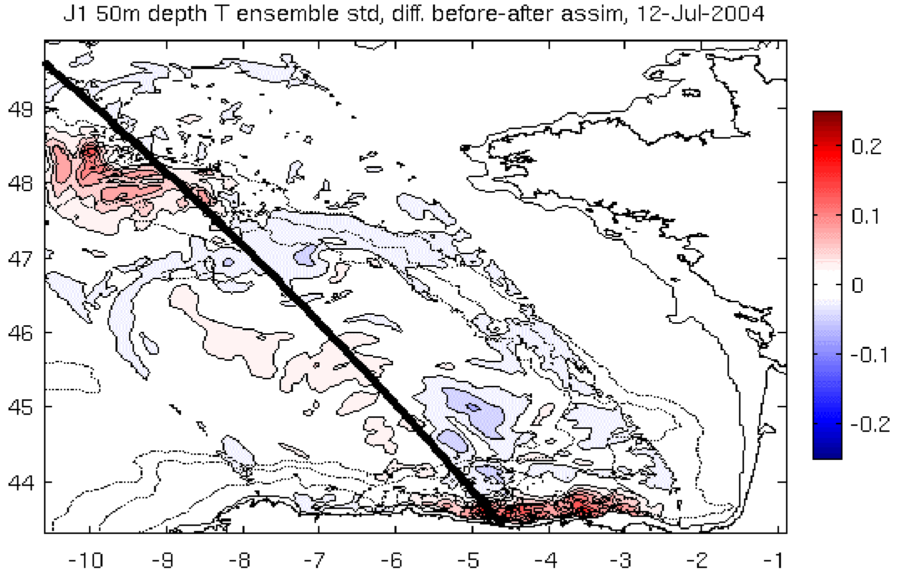
(Lamouroux and De Mey, 2009)

Posterior ensemble variance:
Impact of assimilating wide-swath vs. nadir altimeter data

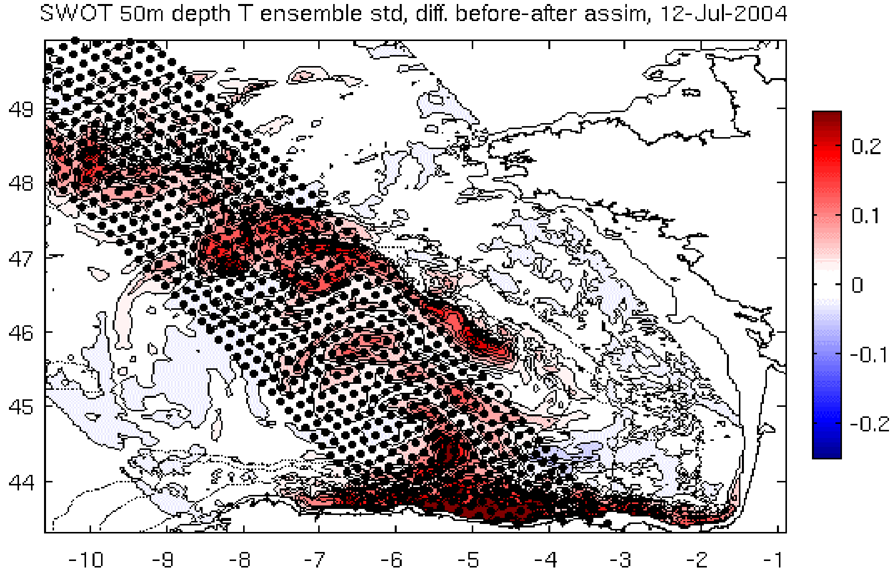
EnKF, ocean-only state vector

Reduction of BoB 50-m temperature ensemble variance at analysis time (°C rms)
in presence of wind forcing errors

Jason-1



SWOT on Jason-type orbit



(Le Hénaff & De Mey, 2009)

RM spectrum analysis for array performance assessment

Le Hénaff & De Mey, 2009 propose a criterion for array design, characterizing how many d.o.f.s of guess (forecast) error an array can detect.

Scaled representer matrix $\mathbf{F}$:

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

$\boldsymbol{\mu}$ *array modes*
$\boldsymbol{\sigma}$ *spectrum of SRM* ←→ spectrum of $\mathbf{I}$

Modal representers: $\boldsymbol{\rho}_{\mu} = \mathbf{P}^{g}\mathbf{H}^{T}\mathbf{R}^{-1/2}\boldsymbol{\mu}$

Stochastic implementation:

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

$$\hat{\mathbf{F}} = \frac{1}{m-1}(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{A}^{f})(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{A}^{f})^{T} = \mathbf{S}\mathbf{S}^{T}$$

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

with

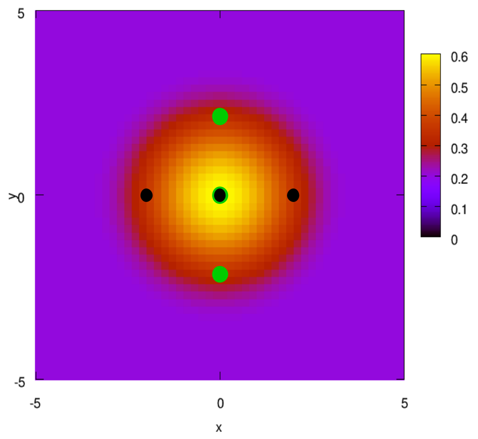
$\mathbf{A}^{f}$ *forecast ensemble anomalies*

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

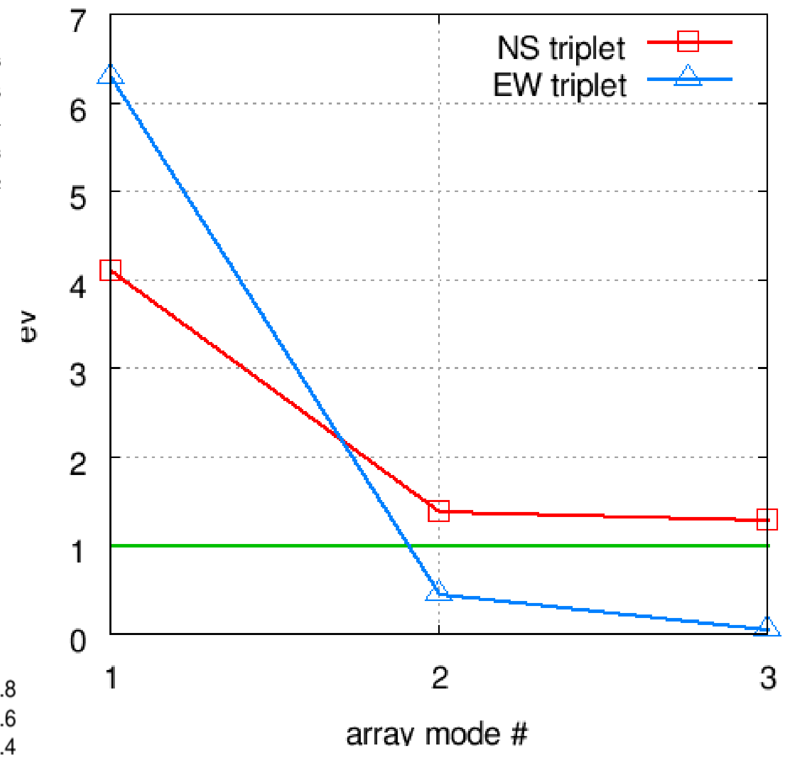
- Two levels of analysis
  - Just count eigenvalues above 1
  - Explore array modes & modal representers
- Ensemble samples
  - Stochastic modelling
  - EnKF (online analysis)
- State error sources
  - Wind stress
  - Wind stress + pressure
  - Bathymetry
  - River runoff
  - Turbulence (meso, mixing)
  - Large-scale circulation
  - Initial/boundary conditions

A simple example

Fcst error variance



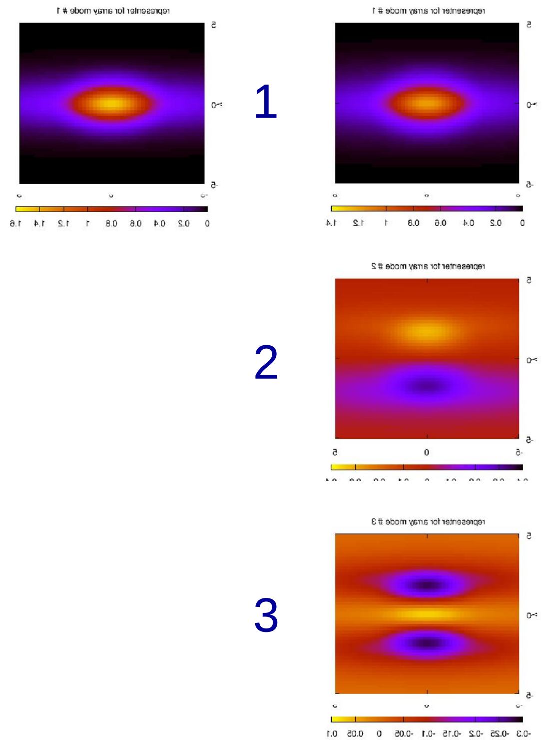
RM Spectra



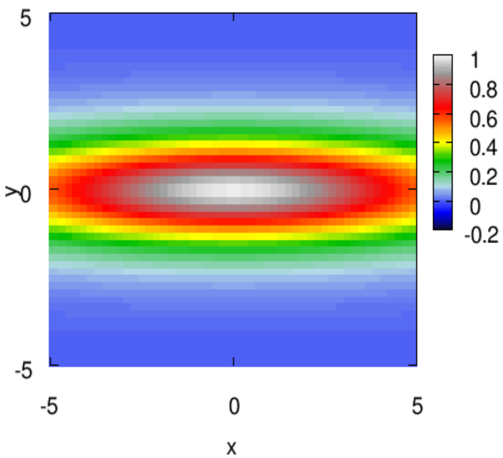
Detectable modal representers

EW triplet

NS triplet



Fcst error corr

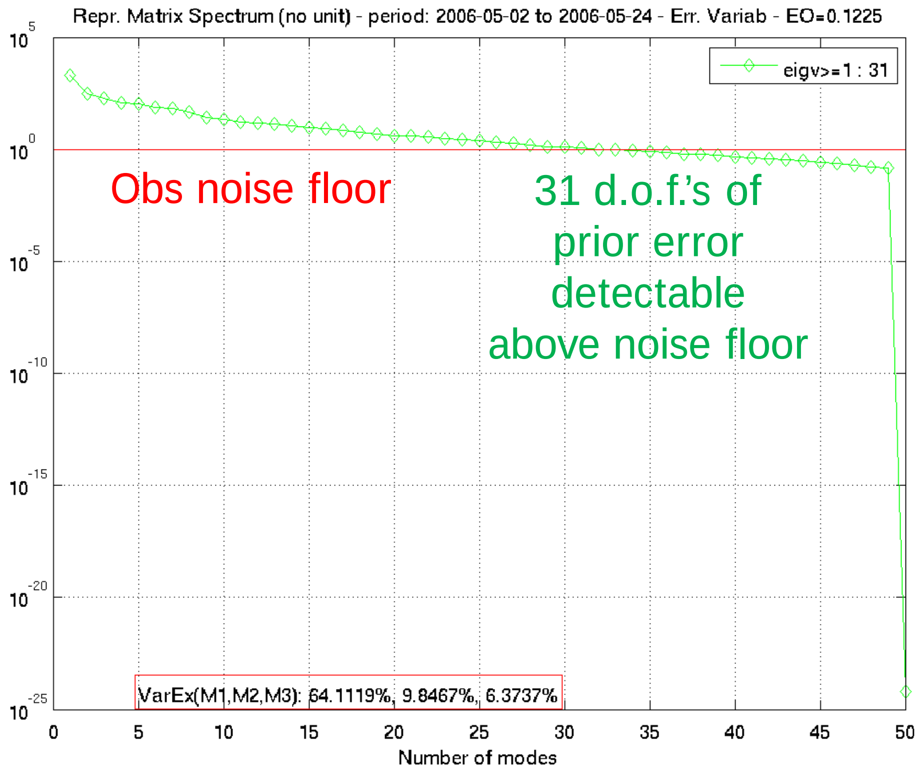


Two observational strategies:

- N-S triplet detects more d.o.f.s (3) amidst observational noise
- E-W triplet more redundant (1 d.o.f.)

RECOPESCA fishing net array (Ifremer)

RM eigenspectrum

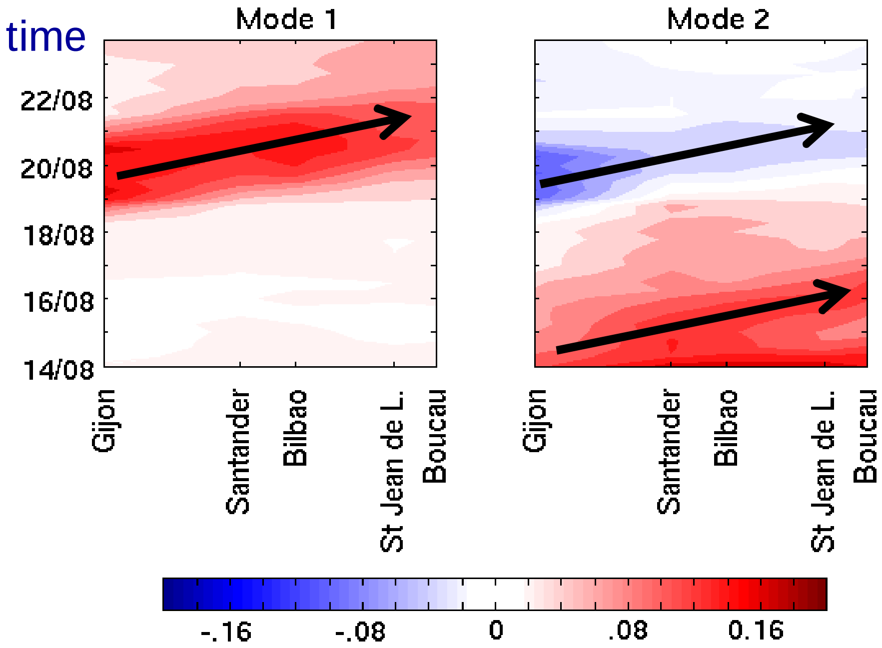


(Lamouroux, Charria et al, 2011)

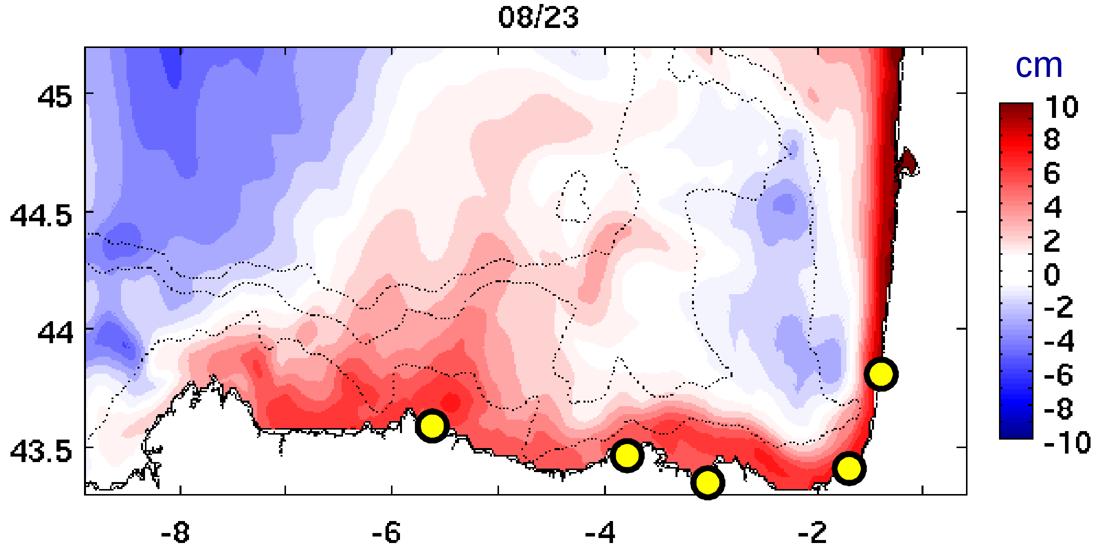
Coastal tide gauges

- BoB experimental configuration, Gaussian wind errors, Summer 2004
- Test existing Spanish tide gauges (hourly data)
- Array modes contain the array’s view of <u>Kelvin waves</u> propagating Eastward in the model’s error subspace, cleanly illustrating that the TG network can detect, and potentially correct, such error processes

Array modes (SLA) with KW propagation



SLA component of mode 1 modal representer

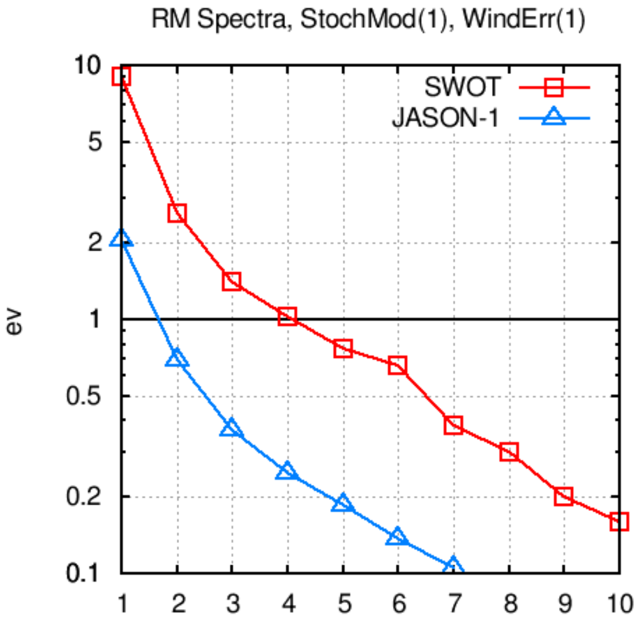


(time slice, August 23, 2004, 0000Z)

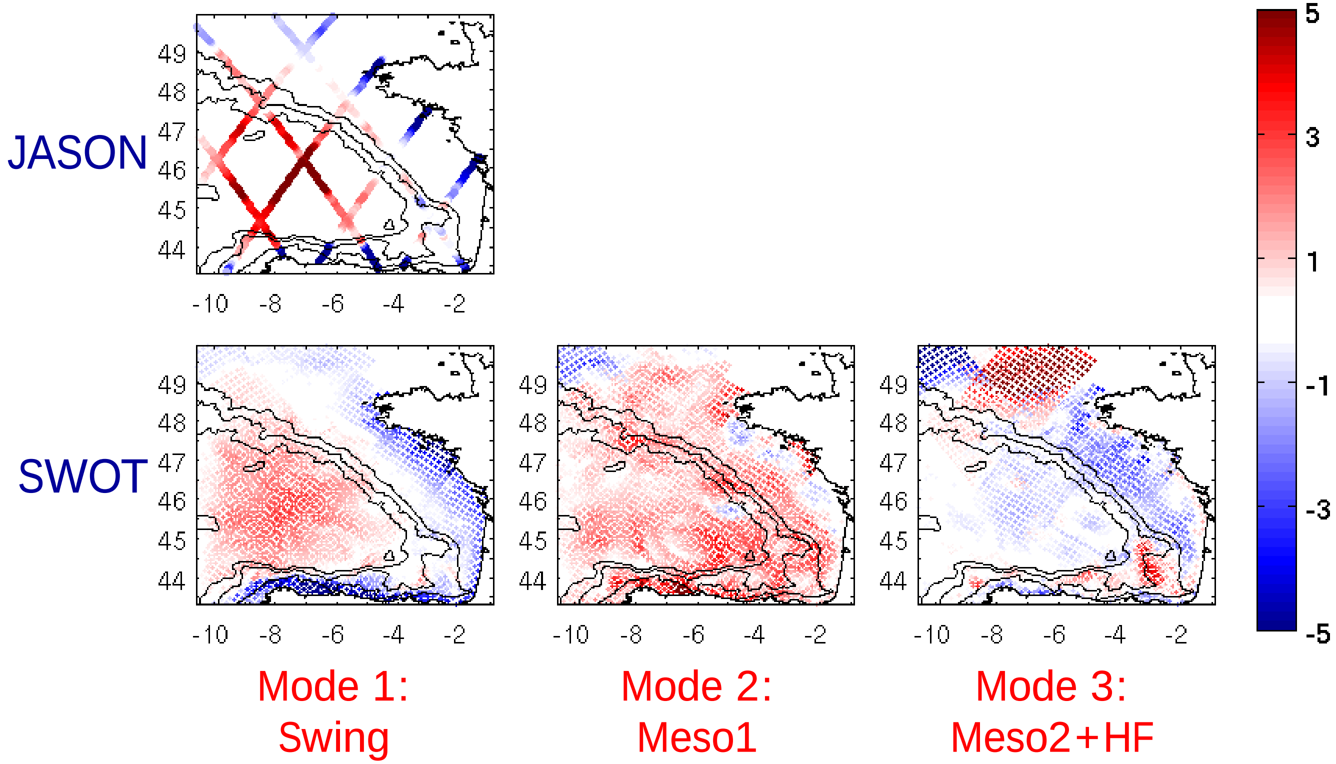
(Le Hénaff & De Mey, 2008)

RM Spectrum analysis: wide-swath vs. nadir altimeter (Le Hénaff & De Mey, ODyn, 2009)

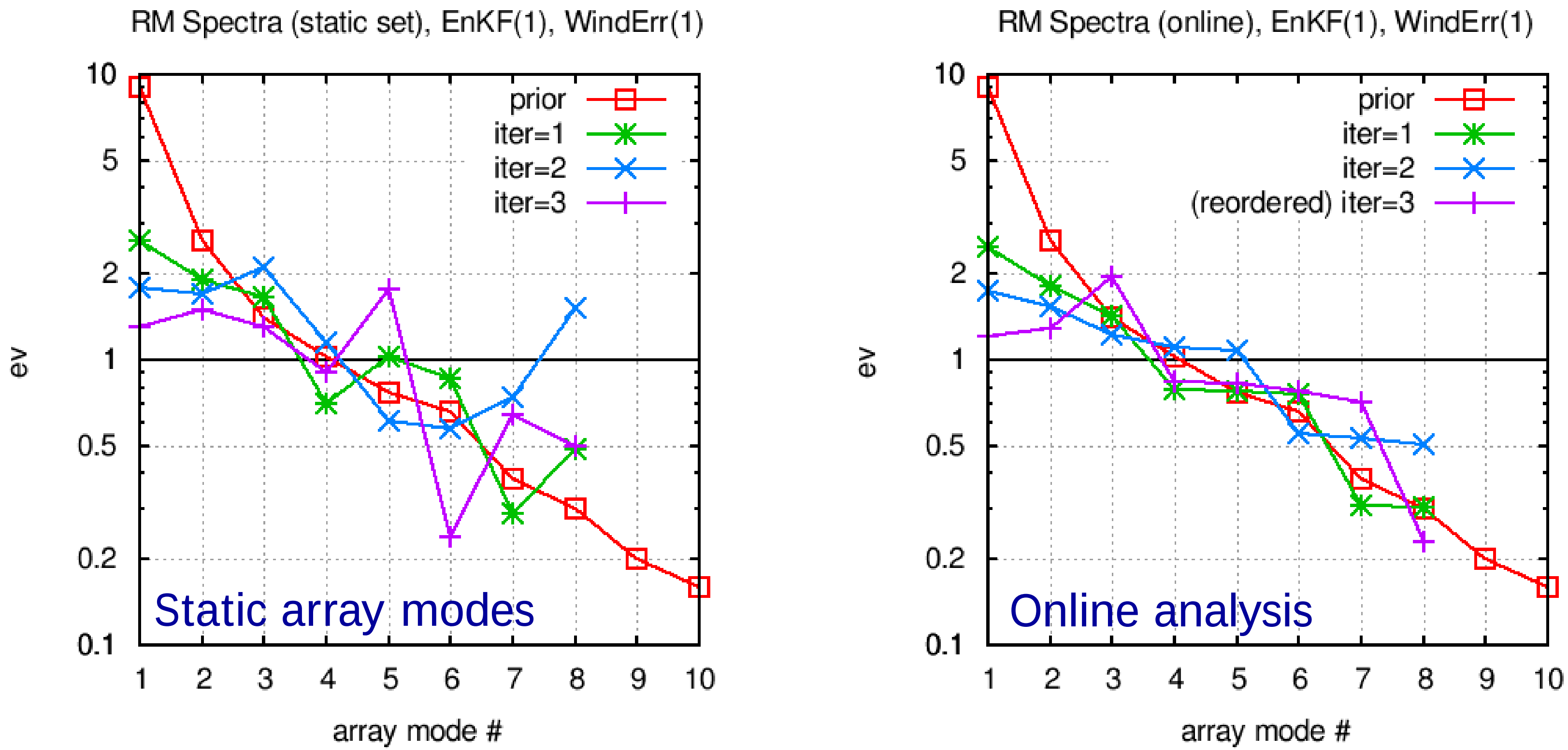
RM spectra



First 3 detectable array modes (h)



RM Spectrum analysis: online analysis with EnKF



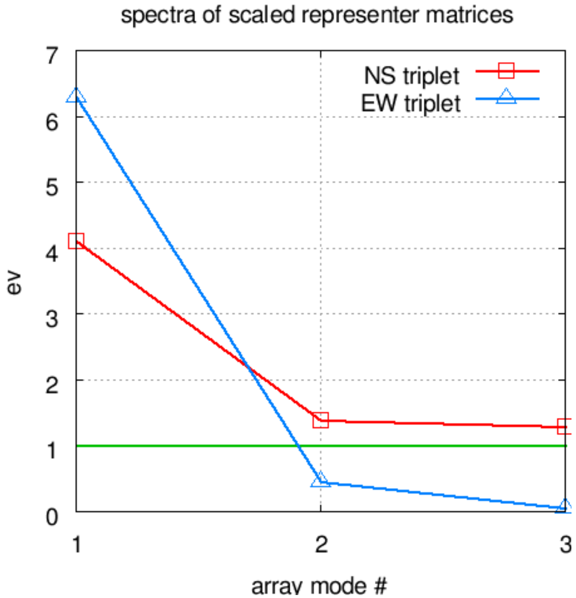
SWOT, Gaussian wind errors, 10-day assim cycles (invariant **H**):

- Spectra whiten in detectable range (array info being extracted)
- Swing & Meso1 array modes evolve slowly throughout regime changes
- Confirm that SWOT alone exhibits useful performance at constraining Swing & Mesoscale error processes, more marginal perf. for HF on shelf

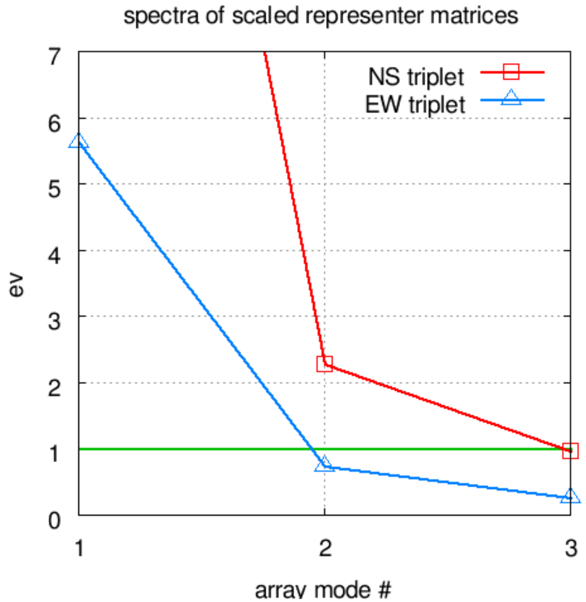
RM Spectrum analysis: correlated observational errors

- Real observational errors are generally correlated (at least through representation errors)
- Triplet example: advantage of NS triplet partially lost
- SWOT example: small impact of along-track correlated roll

Triplet example

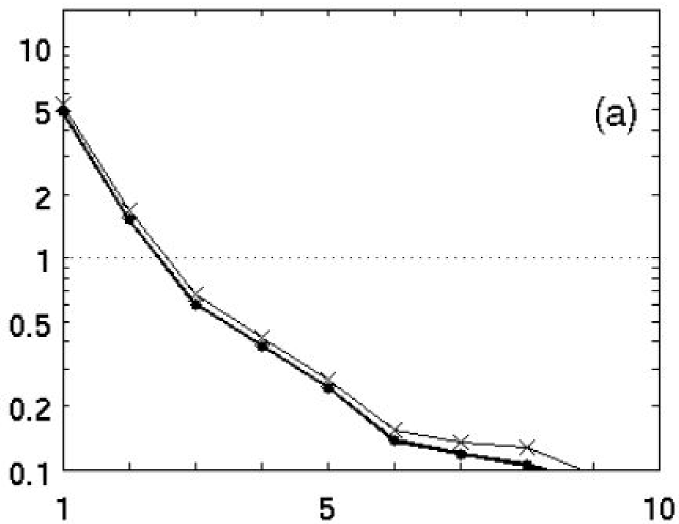


diagonal **R**



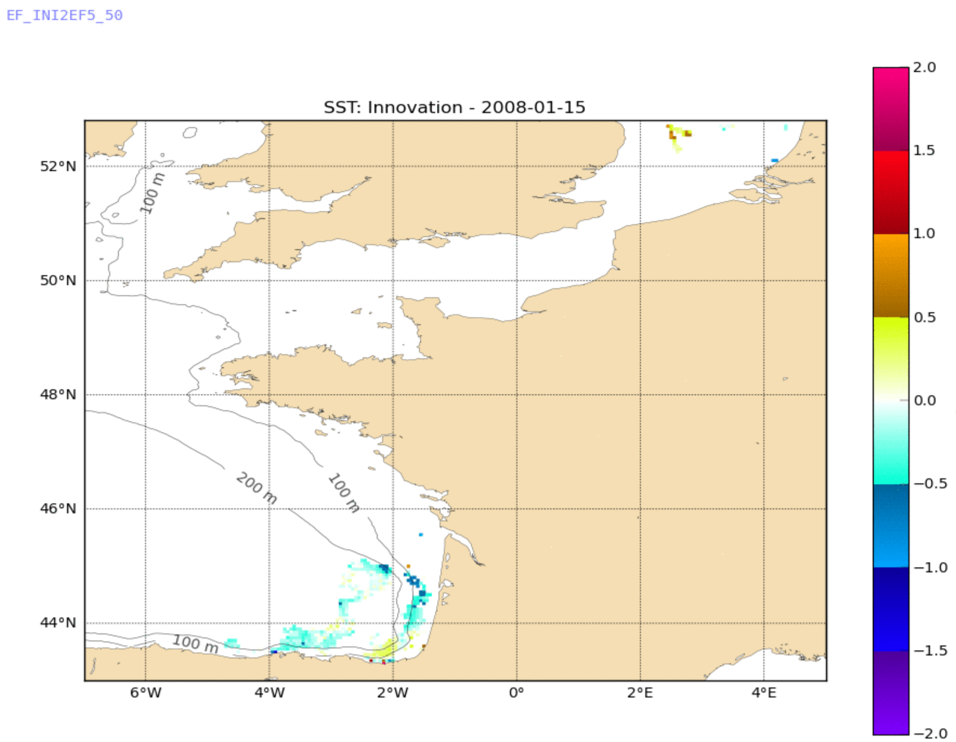
corr. scale = 3 units

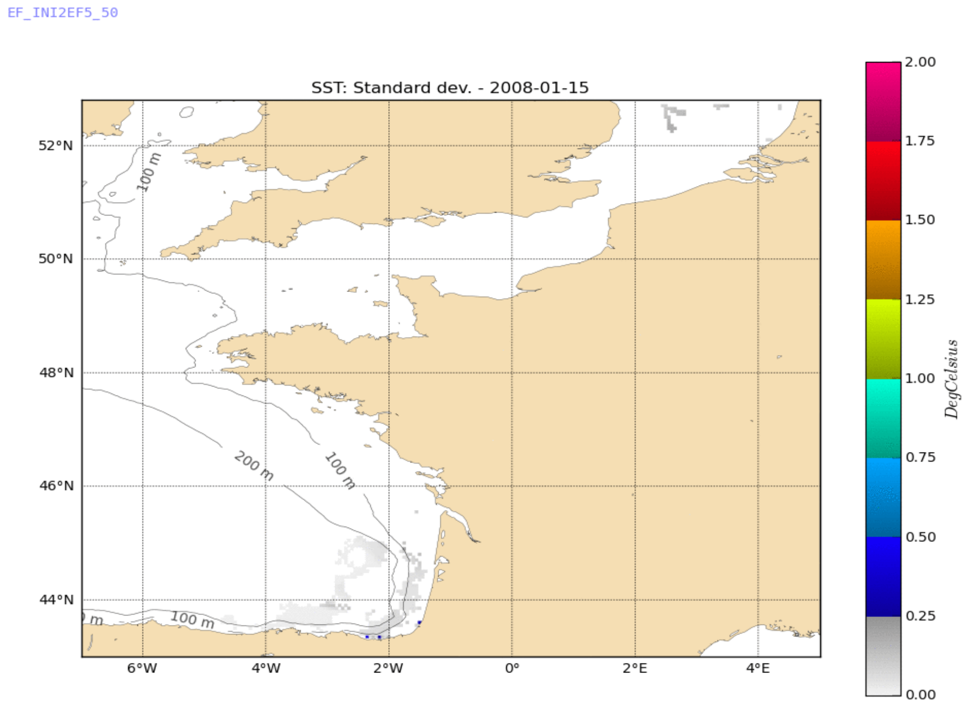
SWOT, Bay of Biscay



Impact of roll errors
(Le Hénaff, 2008)

Ensemble consistency analysis wrt. innovation statistics (Ifremer/Actimar)



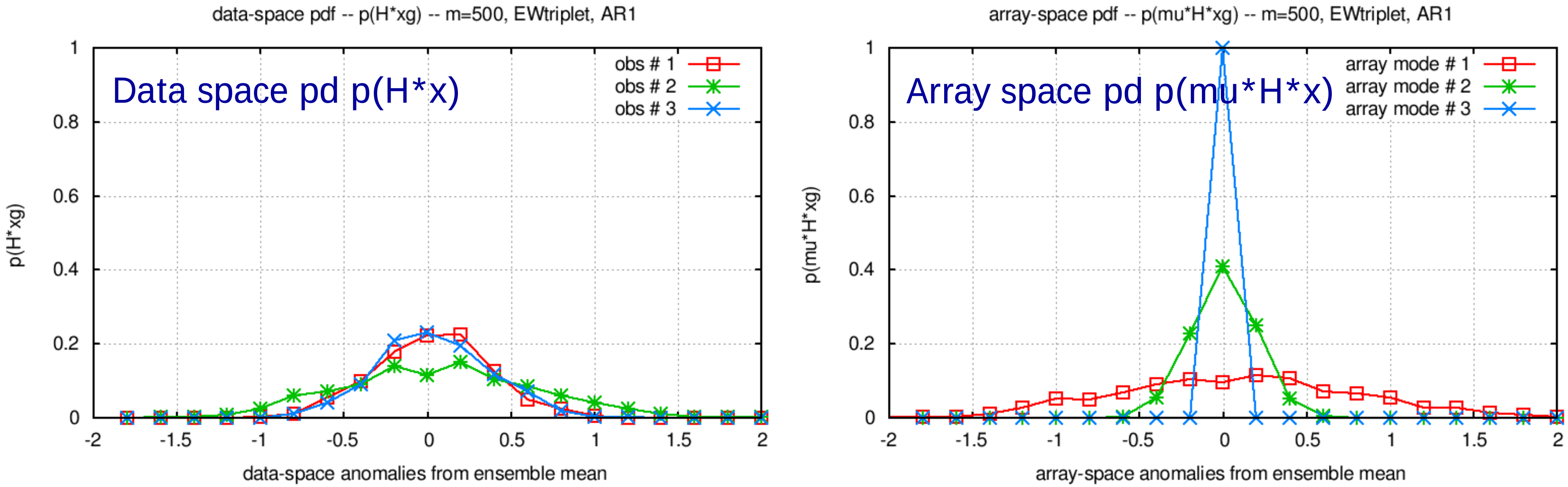


Jan 15 – Feb 15, 2008; SST data space

(Cranéguy, et al. 2011)

(Towards) array-space consistency analysis

- Goal: check whether pd's of model forecast and observations are consistent weo (visually, through reliability scores, etc. – not the topic here)
  - Data space vs. array space
- Low-order array-space forecast pd's have broadest base (by design)
  - Hierarchize consistency checks from easiest to hardest to pass



EW triplet, stationary/correlated AR process, 500 members

Ongoing work (non-SANGOMA)

Tasks	Status
SEQUOIA update • Unstable branch v1.3 (MyOcean/WP3,8) • Upgrade several SEQUOIA components from previous version (v1.2.1sirocco1) reusing 1.3 elements → v1.2.10sirocco2	sirocco2 80% through
Build RM analysis into BELUGA (SEQUOIA v1.3) (MyOcean/WP3,8) • Used to test capacity of obs. arrays to detect prior state error • On-line analysis • Use **S** matrix as in Sakov paper	70% through (prototype version presented in Portland)
Build array-space consistency analysis into BELUGA (SEQUOIA v1.3) • Used to formalize consistency checks btw. observational and Ensemble pdfs • Compare data-space to array-space checks	Started (outline presented in Portland)
Test BELUGA/sirocco2 in BoB region • Use existing interface with G. Herbert's PhD thesis prototype	30% through (Nadia's LEFE proposal)
NEMO interface (MyOcean/WP8) • Only: testing needed ensemble size, feedback onto SAM-2	Ongoing

What we plan to do in SANGOMA (tbd)

- WP1 (1mo in DoW)
  - participate in specs (easier in Dec)
  - possibly submit stochastic array performance assessment tool
- WP2 (2mo)
  - depends on WP1
- WP3 (7mo)
  - implement common new non-linear DA method in SEQUOIA framework
- WP4 (6mo)
  - benchmark common new DA method in SEQUOIA framework: Lorenz-40 and another configuration – coastal/Bay of Biscay or double-gyre (tbd)
- WP5 (6mo)
  - stochastic array performance assessment with RMSpectrum (implementation partly depending on our WP4 configuration – for the BoB: SLA, SST, radars, SWOT, other data sources)
  - array-space consistency analysis (implementation partly depending on our WP4 configuration)
  - D5.6 library + report

Toulouse personnel

- P De Mey, DR CNRS
- N Ayoub, CR CNRS
- Post-doc, €€ for O(28m), mismatch?
- Help from C Nguyen, SIROCCO engineer

To finish, a few photos I took last weekend of cranes from the 10000+ individuals on stopover at the Lac de Der on their way migrating South!

Crane
= Grue cendrée
= Kraanvogel

I have better shots but they are still in my camera!

Of sangoma and cranes

« L'étude approfondie de la danse rituelle *semah* des Alévis-Bektachis éclaire d'un jour nouveau la question de l'origine chamanique de leur système religieux. Chorégraphies et chants sont révélateurs d'un attachement, par-delà les siècles, à la symbolique générale de l'ancienne société de chasse altaïque. Les couples se réunissent l'hiver dans *l'ayn-i cem* où ils chantent et dansent pour appeler en migration les grues cendrées, dont le retour printanier annonce la régénération vitale de la nature nourricière. »

Françoise Arnaud-Demir, *Quand passent les grues cendrées*, Turcica, 2002