Data Assimilation in Regional and Shelf Seas (DARSS)

Pierre De Mey, LEGOS/POC, Toulouse, France, PI, demey@neyak.org
Nadia Pinardi, INGV, Bologna, Italy, Co-PI
Francis Auclair, LA/POC, Toulouse, France, Co-I
Rémy Baraille, SHOM, Toulouse, France, Co-I
Claire Dufau, Baptiste Mourre, LEGOS/POC, Toulouse, France, Co-I's

Abstract

This OSTST proposal is about setting up specific advanced data assimilation tools to make use of both globally and locally available data, including present and future satellite altimeter measurements, in regional, open-boundary, free-surface numerical models at space scales intermediate between the global and coastal scales, and at time scales from hours to months. It is also about testing those ideas in several different regions of the Mediterranean and Northeast Atlantic and with several different types of models, all of them assimilating nadir altimetry in addition to other specific data (tide gauges, ADCP, coastal radars, AXBT, as well as simulated OSTM data and the "open ocean" data types).

Several categories of advanced multivariate estimation techniques are being developed and tested: (1) Ensemble methods play a central role in the preliminary exploration of the model error subspace and in the specification of assimilation statistics; (2) variational balanced analysis is used in the downscaling of the larger-scale solution in the regional free-surface model and in the projection onto a suitable dynamical attractor akin to the well known, "slow manifold"; (3) reduced-order optimal interpolation using Ensemble statistics drives the costlier models (e.g. high-resolution baroclinic); (4) Ensemble Kalman filter and similar schemes are used in studies involving highly non-stationary statistics. The tools developed in this proposal are being made available to colleagues in specific OSTST proposals and to the general community.

In addition to their scientific value, these developments are expected to be useful for the design of future satellite missions such as OSTM, and for the design of future Regional/Coastal Ocean Forecasting Systems (R/COFS) which are needed to address major environmental and societal issues in regional and coastal seas and which will make use of OSTM data.

More on us:
http://poc.omp.obs-mip.fr

VIFOP principle

A Variational Balanced Initialization technique has been implemented *(Auclair et al., 2000a,b)* to analyze the observations and/or the outputs of regional scale or basin scale circulation models used as initial field or open boundary forcing of high resolution ocean models. It was shown that such a method reduces drastically the amplitude of the numerical transient processes following the initialization by reducing the misfit of the initial field. As importantly, the adjustment of the initial geostrophic component of the current reduces the amplitude of the horizontal pressure gradient truncation errors due to the use of sigma coordinate in steep topography gradient areas. A short description of the method is now given.

The coastal high resolution model is written in a compact form:

$$\mathbf{x}^f(t_{i+1}) = M_i\left[\mathbf{x}^f(t_i)\right] \qquad (1)$$

where $\mathbf{x}$ is the model state vector, M the dynamics operator and t is time. The notations are the classic notations in which the superscripts f, a and o stand respectively for forecast, analysis and observation. A first "crude" interpolated "background or first guess state vector" $\mathbf{x}_b(t_0)$ is obtained using a smooth Gaussian interpolation on the high resolution grid:

$$x^b_{i,j} = \frac{\sum_{(\alpha,\beta)} \exp(-\frac{r^2}{R^2})\, y^d_{\alpha,\beta}}{\sum_{(\alpha,\beta)} \exp(-\frac{r^2}{R^2})} \qquad (2)$$

$$x^b_{i,j,k} = \frac{\sum_{(\alpha,\beta)} \exp(-\frac{r^2}{R^2})\,[y^d_{\alpha,\beta,\gamma_1} * d + y^d_{\alpha,\beta,\gamma_2} * (1-d)]}{\sum_{(\alpha,\beta)} \exp(-\frac{r^2}{R^2})} \qquad (3)$$

where *(i,j)* is a point of the high-resolution grid and *(a,b)* a point of the regional scale grid. The distance between the points *(i,j)* and *(a,b)* is given by r and R is typically equal to 5 km. The vertical distance between points *(i,j,k)* and *(a,b,g)* is given by *d* and the vertical interpolation is linear, g_1 and g_2 satisfying $z(\gamma_2) \le z(k) \le z(\gamma_1)$. The vector $\mathbf{y}^d$ is the observation vector including in-situ observations of model variables and/or output from regional or large scale models. Among the advantages of such a filter are its simplicity, its easy and well-understood tuning. However, it must be carefully tuned in so far as it can over-smoothed the regional scale fields.

In the next step, an optimal perturbation of this first guess is computed by minimizing the cost function:

$$J = (\mathbf{x}^a - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x}^a - \mathbf{x}^b) + \left[\begin{pmatrix}\mathbf{y}^o \\ \delta_t \mathbf{y}^o\end{pmatrix} - \begin{pmatrix}\mathbf{H} \\ \mathbf{H}^\delta\end{pmatrix}\mathbf{x}^a\right]^T \mathbf{R}^{-1} \left[\begin{pmatrix}\mathbf{y}^o \\ \delta_t \mathbf{y}^o\end{pmatrix} - \begin{pmatrix}\mathbf{H} \\ \mathbf{H}^\delta\end{pmatrix}\mathbf{x}^a\right] \qquad (4)$$

$\mathbf{R}$ is the "observation" error covariance matrix (dynamical constraint + data), $\mathbf{B}$ is the error covariance matrix for the interpolated field $\mathbf{x}^b$, and $\mathbf{H}^\delta$ is the "observation" operator based on the tangent high resolution coastal model M in the vicinity of $\mathbf{x}^b$ and $\mathbf{F}_0$ is the forcing vector in (1). $\delta_t \mathbf{y}^0$ is the "tendency" vector: it indicates the true values of the constrains like the tendencies of the tangent linear equations, or some particular local or global observations. These tendencies can be specified from the regional scale models or from observations. In such a case, the analysis field is forced to have first order dynamics in agreement with these output data or observations. In the standard implementation of the method, the initial tendencies of sea level are forced to vanish ($\delta_t \mathbf{y}^0 = 0$). The analysis field is thus made close to an equilibrium field by filtering out initial high frequency processes. In the extended formulation, the tendencies are made to match an external solution, which comes back to specifying that the *residual* tendencies are forced to vanish.

$\mathbf{B}$ is in practice decomposed in two parts: one computed with homogeneity assumptions, the second being a parameterization of non-homogeneous and non-isotropic error structures due to the bathymetric constraints. This latest parameterization is composed of a component based on the difference of bathymetry between the regional and coastal models and an extra-diagonal component standing for the stronger correlation of the geostrophic current along the bathymetry. This non-homogeneous and non-isotropic error covariance matrix $\mathbf{B}$ leads to an analysis field that is very similar to the background interpolated solution $\mathbf{x}^b$ far from the shelf break and the regions with large bathymetry gradients while larger adjustments are computed anywhere else.

The initial field is successively optimized based on geostrophic constraints and then constraints related to the external and internal modes. The conservation of mass is imposed as a strong constraint in a "Sasaki" sense, i.e. the local tendency of the mass balance (the *residual* tendency in the extended formulation) is forced to be exactly zero. The complete computation of the analysis field requires approximately the equivalent of three to five days of computation with the high-resolution model over the same domain. The most important part of the computation is dedicated to the inversion of the optimal systems.

Auclair, F., P. Marsaleix, and C. Estournel, 2000a: Truncation errors in coastal modelling: evaluation and reduction by an inverse method. *J. Atmos. Oceanic Tech.*, **17,** 1348-1367.

Auclair, F., S. Casitas and P. Marsaleix, 2000b: Application of an inverse method to coastal modelling. *J. Atmos. Oceanic Tech.*, **17,** 1368-1391.

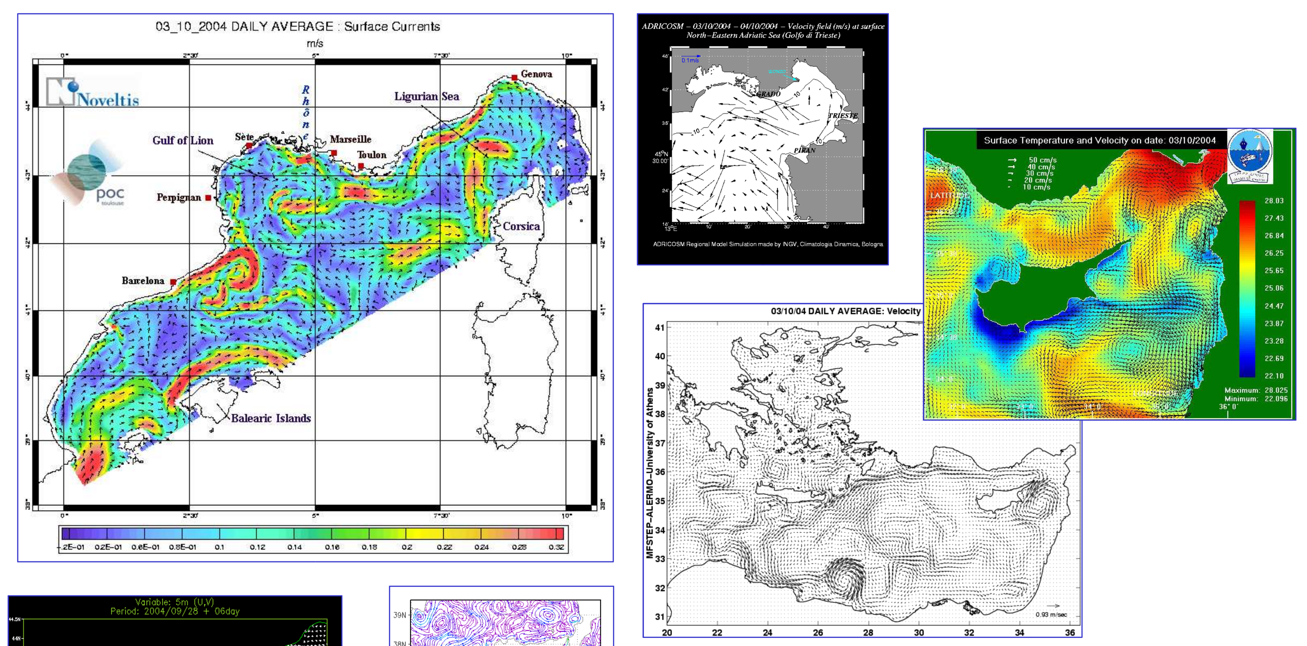
Downscaling from a larger-scale solution

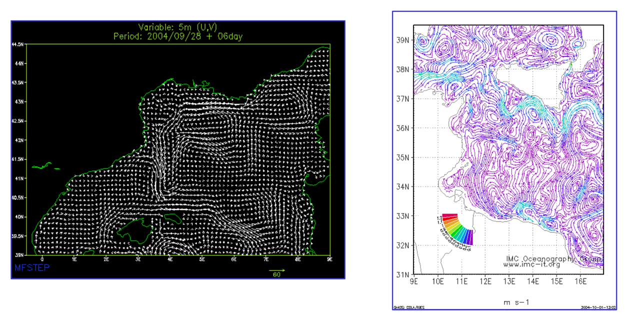
We obtain regional and coastal ocean estimates as part of "downscaling" from a larger-scale solution. This requires developing and using specific tools:

- One- or two-way nesting techniques with support for a variety of models
- Techniques to initialize and project the large-scale solution on the regional/coastal dynamical attractor: this is done together with nesting with the VIFOP tool (see below)
- Techniques to blend the results of different models (e.g., 2D-FE and rigid-lid GCM): this will be done with VIFOP in the cours of the DARSS project
- Techniques to make use of local data not used (or incompletely used) in the GCM: this is in the works at POC -- results at the next OSTST meeting!

These tools are used in operational oceanography projects such as MERCATOR and MFS (Mediterranean Forecasting System).

The MFSTEP Target Operational Period (TOP) in the Mediterranean: Regional and shelf models are weekly downscaled from GCM with VIFOP



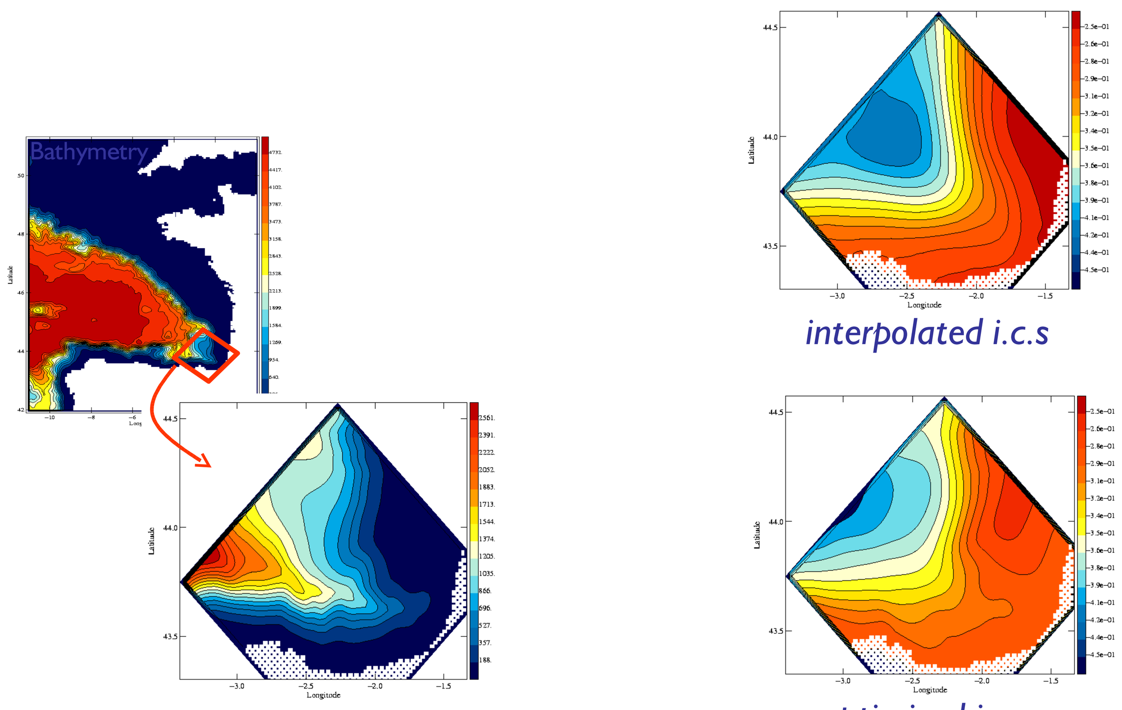


More info: http://www.bo.ingv.it/mfstep

VIFOP: a Variational Balanced Analysis tool for downscaling

Care must be exercised when controlling the initial and boundary values of an embedded free-surface 3D model from a large-scale solution or an assimilated solution since unphysical gravity transients may be triggered. This is due to the fact that the large-scale solution is unbalanced with respect to the local physics, to the different resolution, to the different bathymetry, to the numerical boundary conditions, etc. Investigators in this project have developed a Variational Balanced Analysis method based on the minimization of a cost function involving data constraints as well as a dynamical penalty involving the tangent linear model. This approach leads to several improvements of the free-surface coastal model solution such as a drastic decrease of the spurious numerically generated external gravity waves and a decrease of the amplitude some of the model biases such as the horizontal pressure gradient truncation errors. VIFOP has been set up for several free surface models including SYMPHONIE and POM in the Gulf of Lions, in the Northern Adriatic, in the Golfe de Gascogne and in other areas.

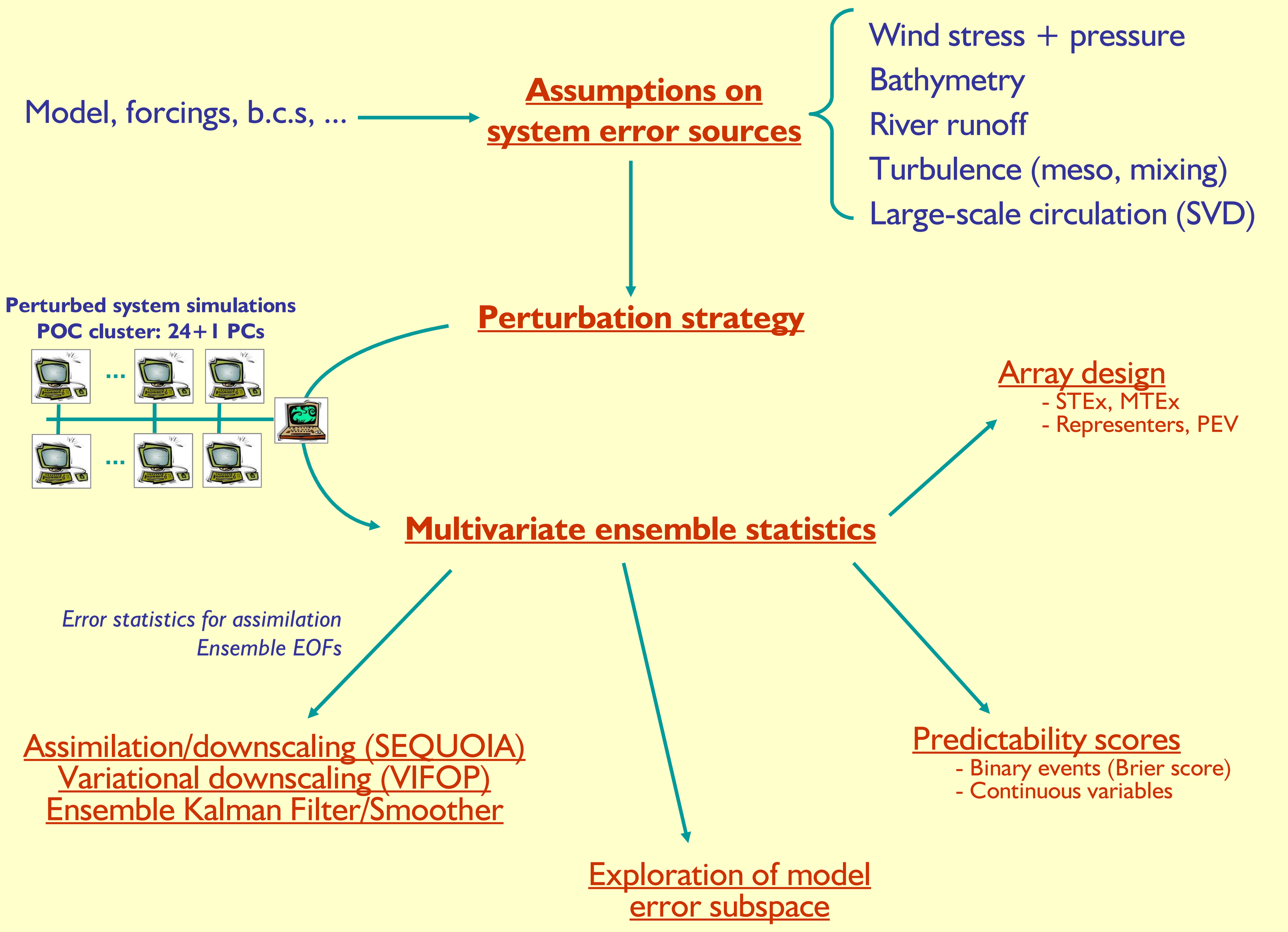
3DFD free-surface model in the Bay of Biscay with VIFOP initialization from MERCATOR GCM solution



interpolated i.c.s

optimized i.c.s: unphysical transients eliminated, better adjustment to bathymetry

Ensemble statistics and their use in coastal estimation problems



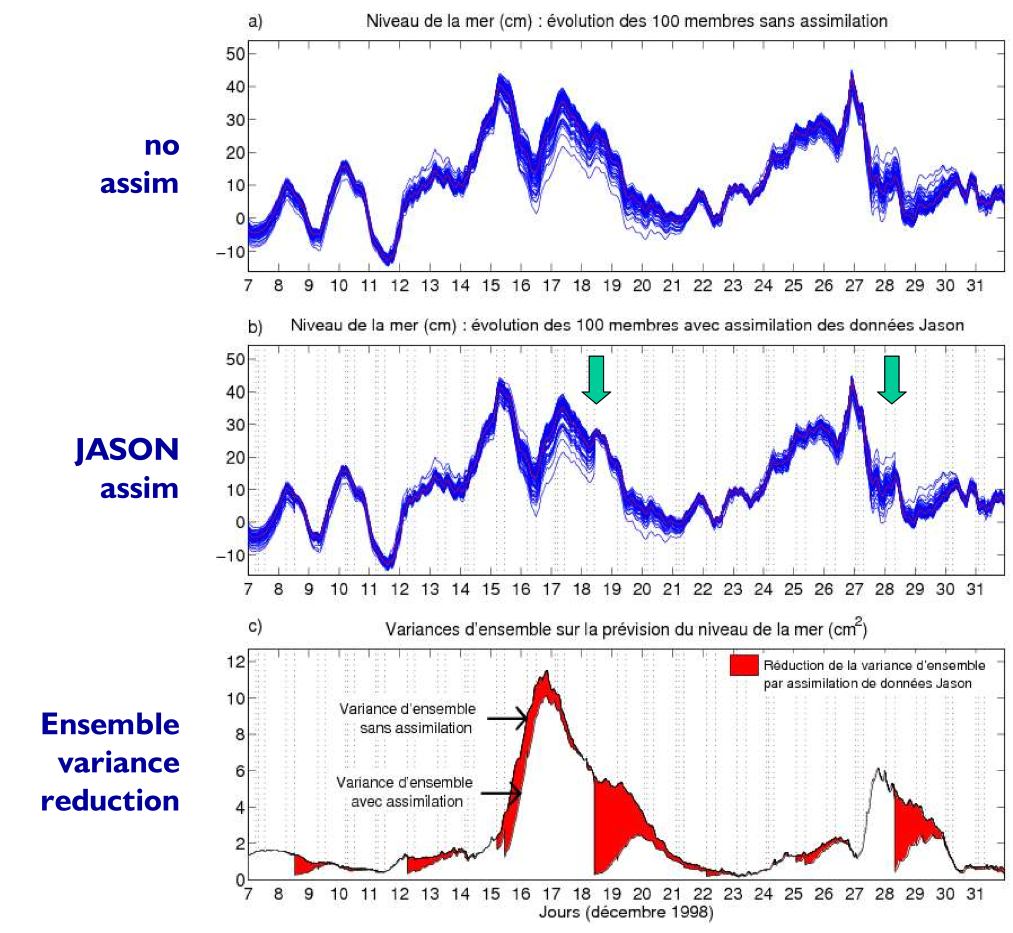
Assessment of observing networks: comparison of forthcoming altimeter system configurations

Because the error statistics in shelf and coastal seas are non-stationary (see bottom left), non-homogeneous, non-isotropic and non-separable, any detailed assessment of an observing network through assimilation warrants the use of a so-called "advanced" method, such as variational or ensemble-based filters.

Here we use an Ensemble Kalman Filter (Evensen, 2003) to perform Observing System Simulation Experiments. Our approach is based on a generalization on the twin experiments approach in which observations simulated from one member are assimilated in all the other members (Mourre et al., 2004). The figure of merit associated with one given observational network under specific assumptions of model error is the average ensemble spread after assimilation. The most efficient observing schemes will constrain the model error subspace better and will reduce this spread.

We chose to perturb the bathymetry in the simulations (this is the only forcing difference between members). Typical bathymetry errors are generated by combining generic error patterns derived from the differences between 6 bathymetry solutions.

Principle: assimilation reduces ensemble spread



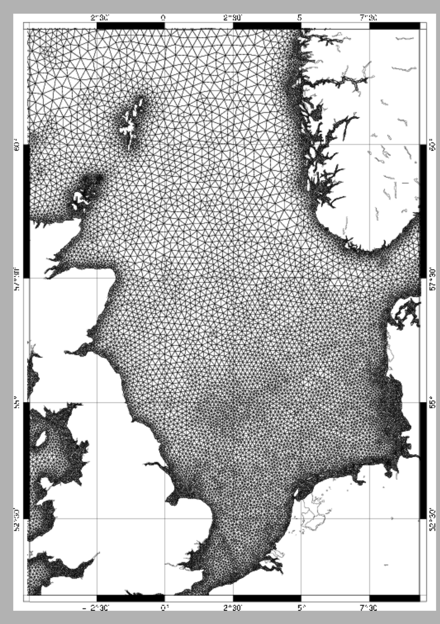
SLA, 7-31 Dec. 1998, point N

Exploration of the model error subspace and use of ensemble statistics in assimilation

Ensemble statistics under specific assumptions on system error sources permit the exploration of the model error subspace along its dynamical attractor, in the form of sensitivity to perturbations:

- Time series of ensemble spread as a function of time quantify predictability limits for selected variables at selected key locations. (The so-called Lorenz time scales are useful in defining a data assimilation cycle.)
- Multivariate representer functions tells us about how observations project on the model state variables, and hence if and how well we can hope to correct the model and with which data. (Representer functions and the eigenvalues of the representer matrix are useful in testing new observational networks.)
- A multivariate SVD or EOF analysis of those second-order moments gives access to coherent error structures, their signature in terms of the physical variables, dynamical contents, magnitude and location.
- The projection of the representer functions on a dominant set of modes (the "reduced state space") will tell us if the reduced state space chosen is compatible with our observational network.
- Time series of error variances as a function of time along each direction associated with a coherent error mode will quantify predictability limits in the reduced state space.

Ensemble modelling in the MOG2D 2D-FE model



Finite-element
Barotropic
Non-linear
Entire European shelf (only North Sea shown)
(τ,p) forcing
HF processes
Perturbed: bathymetry, (τ,p) atm. forcings
Runs on a Linux cluster

Assessment results

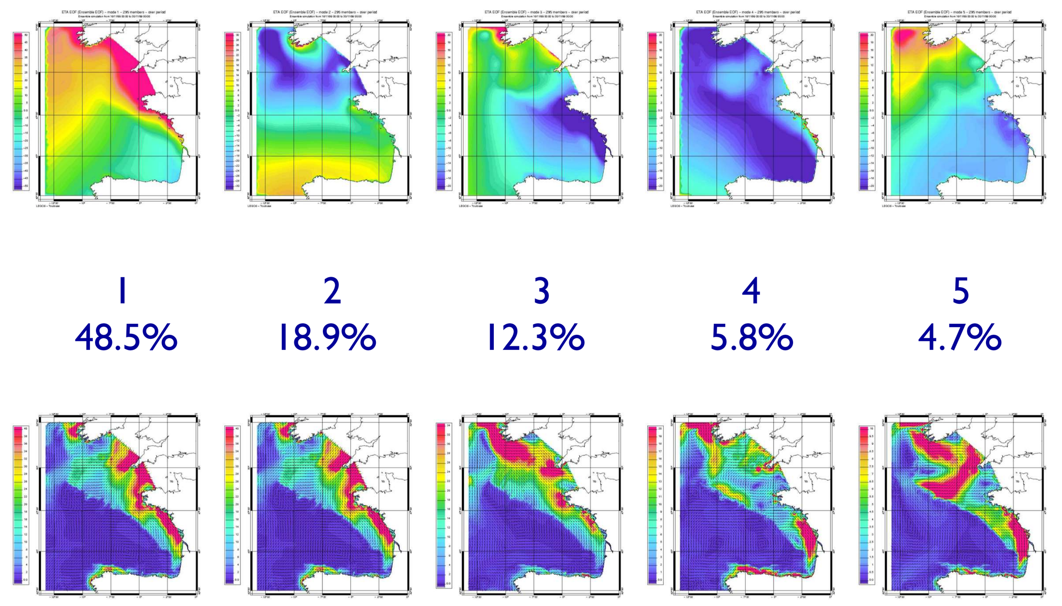
The performance of forthcoming altimeter systems (wide-swath, constellations) and of enhanced tide gauge networks is compared by means of ensemble spread (more rigorously put "posterior ensemble variance").

All observing systems were found to do better with SLA than with velocities. This is mostly due to the fact that the bathymetry itself is not corrected.

As a rule of thumb, and considering the particular perturbation strategy chosen, one wide-swath instrument was found to do as well as two Jasons for sea-level, and three Jasons for velocities.

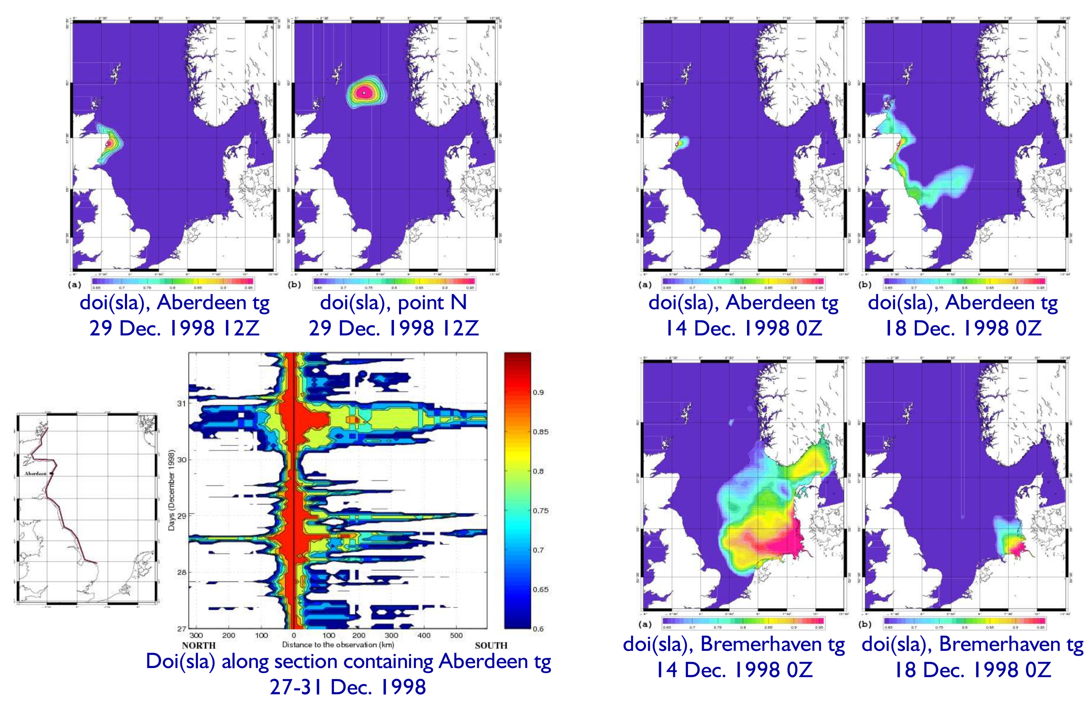
The best results were obtained by combining wide-swath altimetry with tide gauges.

Bay of Biscay multivariate ensemble EOFs in response to wind/pressure perturbations

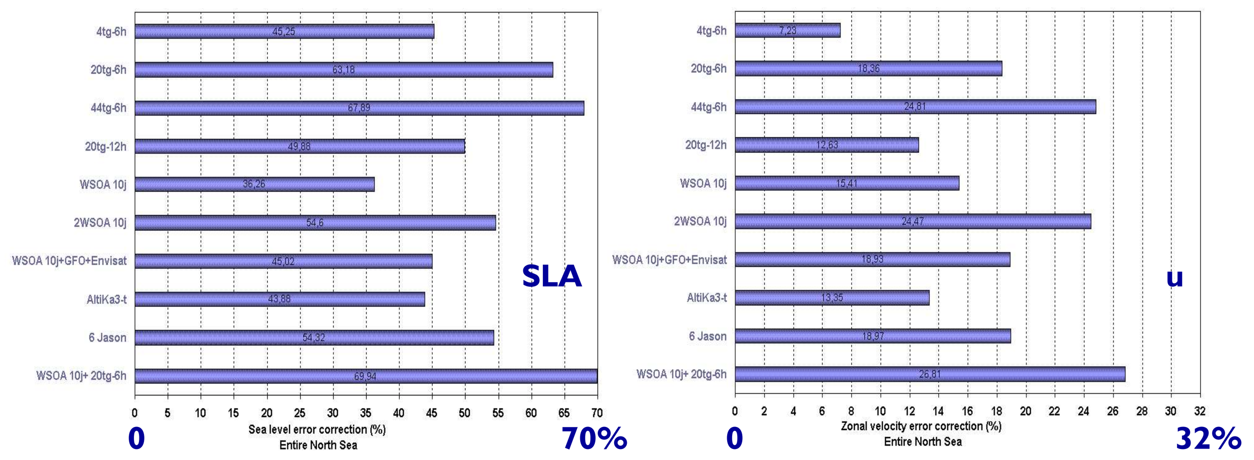


November 1999 (Lamouroux et al., 2004)

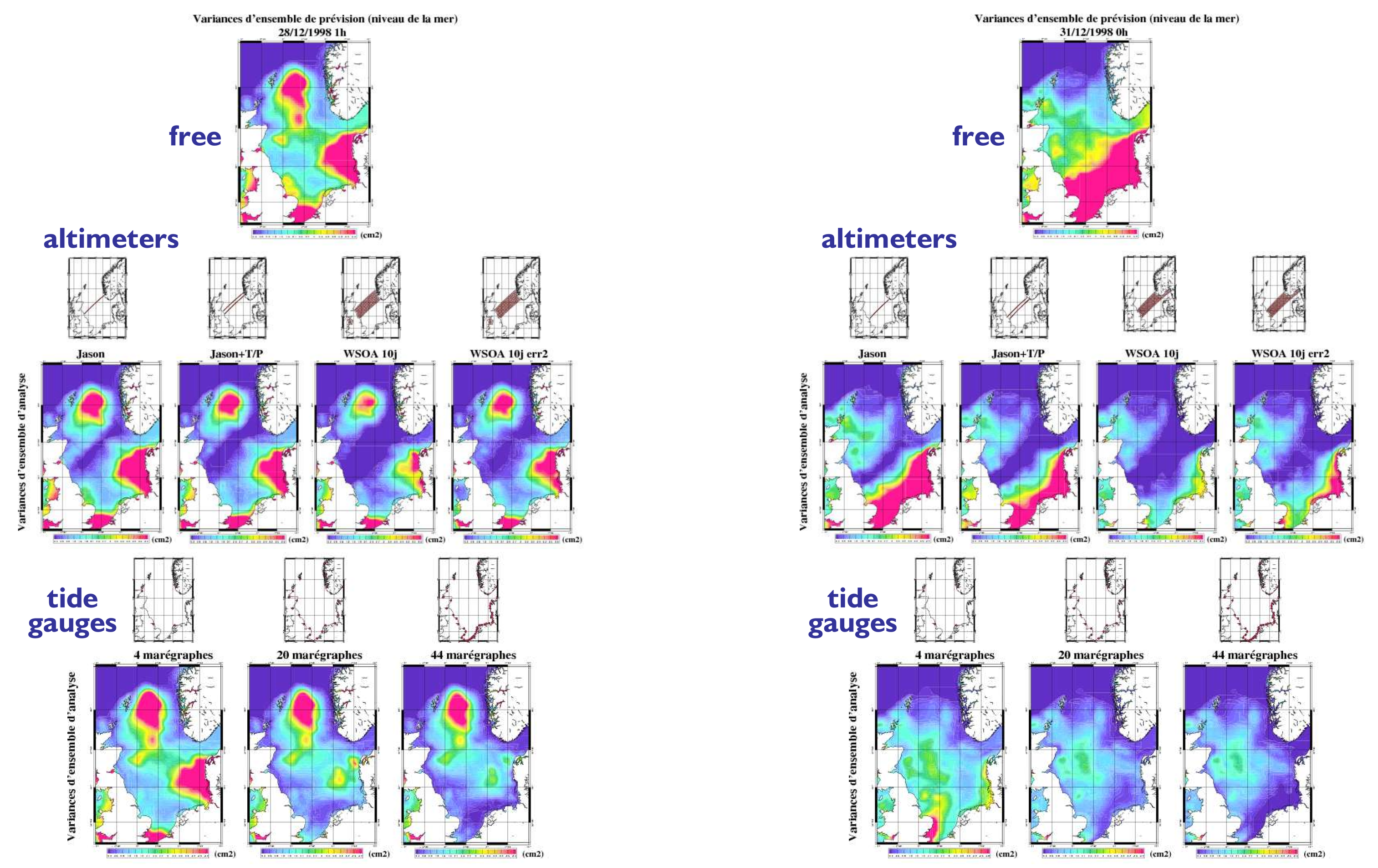
North Sea ensemble representer functions derived from bathymetry perturbations: error statistics in shelf seas are non-stationary!!



Average PEV reduction over ensemble + period + domain



Impact of altimeter and tide gauges scenarios in the North Sea: Posterior Ensemble Variance analysis



Dec 28, 1998 **Dec 31, 1998**