

Dynamically constrained ensemble perturbations. Application to tides on the West Florida Shelf

Alexander Barth^{1,2}, Aida Alvera-Azcárate^{1,2}, Jean-Marie Beckers^{1,2}, Robert H. Weisberg³, Luc Vandenbulcke¹, Fabian Lenartz¹, and Michel Rixen⁴

¹GeoHydrodynamics and Environment Research (GHER), MARE, Allée du 6-Août 17, University of Liège, Sart-Tilman B5, 4000 Liège, Belgium

²National Fund for Scientific Research, Belgium

³Ocean Circulation Group, University of South Florida, 140 7th Avenue South, St. Petersburg, Florida 33701, USA

⁴NATO Undersea Research Centre (NURC), Viale San Bartolomeo 400, 19126 La Spezia, Italy

Correspondence to: A. Barth
(a.barth@ulg.ac.be)

Abstract. A method is presented to create an ensemble of perturbations that satisfies linear dynamical constraints. A cost function is formulated defining the probability of each perturbation. It is shown that the perturbations created with this approach take the land-sea mask into account in a similar way as variational analysis techniques. The impact of the land-sea mask is illustrated with an idealized configuration of a barrier island. Perturbations with a spatially variable correlation length can be also created by this approach. The method is applied to a realistic configuration of the West Florida Shelf to create perturbations of the M2 tidal parameters for elevation and depth-averaged currents. The perturbations are weakly constrained to satisfy the linear shallow-water equations. Despite that the constraint is derived from an idealized assumption, it is shown that this approach is applicable to a non-linear and baroclinic model. The amplitude of spurious transient motions created by constrained perturbations of initial and boundary conditions is significantly lower compared to perturbing the variables independently or to using only the momentum equation to compute the velocity perturbations from the elevation.

1 Introduction

In numerous modelling applications, the uncertainty of the model results needs to be estimated. An estimation of this uncertainty is often obtained by realizing a stochastic ensemble forecast: inputs to the model are perturbed within the bounds of their respective uncertainty and for each of those perturbations the model is integrated forward. For ocean models, the model uncertainty stems, among others, from the forcing fields and initial conditions (*e.g.* Lermusiaux et al., 2006). The spread of an

20 ensemble forecast using perturbed forcings fields and initial conditions gives the estimation of the
model uncertainty.

This uncertainty (often formally expressed as model error covariance) is also required in vari-
ous data assimilation techniques such as the Ensemble Kalman Filter (Evensen, 2003), the Error
25 Subspace Statistical Estimation (Lermusiaux and Robinson, 1999) and the Singular Evolutive Inter-
polated Kalman Filter (Pham, 2001). The success of the data assimilation depends crucially on the
realism of the model error covariance which in turns depends on the realism of the applied perturba-
tions.

30 Ensemble simulations are also very resource-intensive. There is the need, thus, to restrict explo-
ration of the state space only to model states which are physically reasonable and which respect
certain dynamical equilibria. Otherwise resources might be wasted to compute ensemble members
using unlikely initial or boundary conditions.

35 Physically unbalanced initial and boundary conditions produce transient motions during the ini-
tialization, often in form of spurious barotropic waves. Several techniques have been proposed to
reduce those waves in the context of data assimilation (*e.g.* Vallis, 1992; Dobricic et al., 2007; Barth
et al., 2007b). These methods, which basically damp barotropic waves, are difficult to apply to mod-
els which include tides, since tidal waves are also barotropic waves.

40

The objective of this study is to present a method to create smooth, Gaussian-distributed, mono-
variate or multivariate perturbations that have to satisfy a given linear balance. In particular, a
method to produce balanced perturbations for tidal models is presented here. As the method relies
on the definition of a cost function, the land-sea mask and a possibly variable correlation-length can
45 be taken into account.

The motivation for applying this technique to tidal boundary conditions is to assimilate a data set
resolving tides. Since part of the model error is associated with erroneous tidal boundary conditions
and initial conditions, those errors have to be taken into account. In an ensemble approach, the un-
50 certainty of those fields is represented by an ensemble of likely and physically balanced initial and
boundary conditions.

In section 2, the general method is introduced to generate smooth perturbations satisfying certain
constraints. Section 3 shows how the structure of the perturbations depends on the land-sea mask in
55 an idealized configuration. A case with a spatially variable correlation length is also presented. As
an example, the constraint from a shallow water model is introduced and implemented in a realistic

configuration of the West Florida Shelf (WFS) in section 4. This method is compared to other simpler techniques (section 5) and the results are discussed in section 6. In section 7, the numerical cost of this method is studied for different domain sizes. Section 8 summarizes our findings and provides the conclusions.

2 Methods

All model variables at all grid points that need to be perturbed are grouped into the perturbation vector \mathbf{x} which is composed by n elements. We will seek a method that provides smooth stochastic perturbations that satisfy approximately m -constraints given by the $m \times n$ matrix \mathbf{M} :

$$\mathbf{M}\mathbf{x} = 0 \tag{1}$$

We consider only linear constraints here. Non-linear constraints have to be linearized around the unperturbed state. This constraint can be any relationship between different elements of the vector \mathbf{x} known *a priori*, such as the geostrophic equilibrium, zero horizontal divergence of surface winds, a climatological TS diagram, stationary solution to the advection diffusion equation or the linear shallow water equations. The constraint can also be the requirement that the perturbations have to belong to a subspace defined by *e.g.* empirical orthogonal functions (EOFs). In this case the columns of \mathbf{M} would be all vectors orthogonal to the set of EOFs.

Only homogeneous constraints (*i.e.* the right-hand side of equation (1) is zero) are considered here because otherwise the perturbations would have a non-zero mean. The forcing fields to be perturbed are generally assumed to be unbiased, as it is also a requirement for most assimilation schemes.

To describe our *a priori* knowledge of what a realistic perturbation is, we introduce a cost function J , similar to the cost function in variational analysis techniques:

$$2J(\mathbf{x}) = (\mathbf{M}\mathbf{x})^T \mathbf{W}_M (\mathbf{M}\mathbf{x}) + (\mathbf{D}\mathbf{x})^T \mathbf{W}_D (\mathbf{D}\mathbf{x}) + \mathbf{x}^T \mathbf{W}_E \mathbf{x} \tag{2}$$

where \mathbf{W}_M , \mathbf{W}_D and \mathbf{W}_E are, for simplicity, diagonal weighting matrices. The first term penalizes the deviations from the linear constraint (1). It is not enforced strictly (strong constraint), but it has to be satisfied approximately (weak constraint). The diagonal elements of $\mathbf{W}_M^{-1/2}$ define the magnitude of an acceptable deviation from the linear constraint. The matrix \mathbf{D} is a diffusion operator that ensures the smoothness of the perturbation (Brasseur, 1991; Weaver and Courtier, 2001). The last term controls the amplitude (or total energy) of the perturbations. Later, this cost function will be related to the likelihood of a given perturbation. It is necessary that all constraints added in the cost function are compatible to obtain useful perturbations.

90 The cost function is a quadratic function in \mathbf{x} and can thus be written as:

$$2J(\mathbf{x}) = \mathbf{x}^T (\mathbf{M}^T \mathbf{W}_M \mathbf{M} + \mathbf{D}^T \mathbf{W}_D \mathbf{D} + \mathbf{W}_E) \mathbf{x} \quad (3)$$

$$= \mathbf{x}^T \mathbf{B}^{-1} \mathbf{x} \quad (4)$$

where the matrix \mathbf{B}^{-1} (the Hessian matrix of $J(\mathbf{x})$) is defined as:

$$\mathbf{B}^{-1} = \mathbf{M}^T \mathbf{W}_M \mathbf{M} + \mathbf{D}^T \mathbf{W}_D \mathbf{D} + \mathbf{W}_E \quad (5)$$

95 The dynamically constrained covariance matrix \mathbf{B} could be used directly in assimilation schemes that are based on the inverse of the background covariance matrix (such as variational assimilation) without the need to create an ensemble. For ensemble-based assimilation schemes, the cost function can be used to define the probability of a perturbation \mathbf{x} (*e.g.* Kalnay, 2002):

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \|\mathbf{B}\|^{1/2}} \exp(-J(\mathbf{x})) \quad (6)$$

100 Perturbations resulting in a large value of the cost function J , meaning that the constraints are violated, have thus a low probability. The perturbations satisfying the weak constraint (1) are drawn from this pdf.

To generate an ensemble of perturbations that follows the previous pdf, the matrix \mathbf{B}^{-1} is decomposed in eigenvectors (rows of \mathbf{U}) and eigenvalues (diagonal elements of $\mathbf{\Lambda}$):

$$105 \quad \mathbf{B}^{-1} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \quad (7)$$

The larger an eigenvalue is, the stronger the corresponding eigenvector violates the dynamical and smoothness constraint. Indeed, if the perturbation is the eigenvector \mathbf{u}_i associated with eigenvalue λ_i , the cost function takes the value $2J = \mathbf{u}_i^T (\mathbf{U} \mathbf{\Lambda} \mathbf{U}^T) \mathbf{u}_i = \mathbf{u}_i^T \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \mathbf{u}_i = \lambda_i$

110 An ensemble of vectors $\mathbf{z}^{(k)}$ where the subscript k is the ensemble member, is created following a normal distribution.

$$\mathbf{z} \sim N(0, \mathbf{I}_n) \quad (8)$$

An ensemble of perturbations $\mathbf{x}^{(k)}$ following (6) can be obtained by:

$$\mathbf{x}^{(k)} = \mathbf{U} \mathbf{\Lambda}^{-1/2} \mathbf{z}^{(k)} \quad (9)$$

115 In practice, only the smallest eigenvalues (and their corresponding eigenvectors) need to be calculated. This enables the use of efficient software packages to compute selected eigenvalues and eigenvectors. In this work, we use the GNU Octave interface to the ARPACK package (Lehoucq et al., 1997). This approach for generating perturbations will be used in the subsequent experiments.

As a variant of the previous equation, perturbations can also be obtained by 2nd order re-sampling method used in the SEIK (Singular Evolutive Interpolated Kalman) Filter (Pham, 2001):

$$\mathbf{x}^{(k)} = \mathbf{U}\mathbf{\Lambda}^{-1/2}\mathbf{H}(\Omega)_k \quad \text{for } k = 1, \dots, N + 1 \quad (10)$$

where N is the number of eigenvectors retained, \mathbf{H} is a $N + 1 \times N$ matrix whose column vectors form an orthonormal basis perpendicular to the vector $\mathbf{1}_{N+1 \times 1}$ and $(\Omega)_k$ is the k th column of a $N \times N$ random orthogonal matrix Ω . This ensemble will have exactly a zero mean and a covariance equal to \mathbf{B} reduced to its N largest eigenvalues.

Instead of using an eigenvector decomposition, perturbations can also be created by a Cholesky decomposition of \mathbf{B}^{-1} into square root matrices (Fukumori, 2002):

$$\mathbf{B}^{-1} = \mathbf{R}\mathbf{R}^T \quad (11)$$

The perturbations can then be obtained by:

$$\mathbf{x}^{(k)} = \mathbf{R}^{-1}\mathbf{z}^{(k)} \quad (12)$$

Since \mathbf{R} is a triangular matrix, the product of its inverse and a vector can be efficiently calculated by back substitution.

3 Impact of the land-sea mask and correlation length

First we will show how the effect of land boundaries such as islands and peninsulas are taken into account to compute the perturbations. The effect induced by those barriers is difficult to include in methods which derive perturbations based on a given covariance matrix. For example, the method described in Evensen (2003) is widely used to generate ensemble perturbations. The perturbations are generated in Fourier space. Each Fourier mode is perturbed independently with an expected amplitude proportional to a Gaussian function of the wave number, producing a smooth field in physical space. This method is computationally very efficient since it can be implemented using the Fast Fourier Transform. Indeed, it can be shown that for any translation-invariant covariance matrix, its eigenfunctions are the Fourier modes (Barth et al., 2007a). In some circumstances, the approach might be appropriate for atmospheric fields and also for oceanographic fields far from the coastline. But in oceanographic applications, problems can arise near the coast. This problem will be illustrated with a narrow elongated island.

The effect of land boundaries is included in the smoothness constraint and is independent of any dynamic constraint. To show this we can think of the diffusion operator (on which the smoothness constraint is based) as a way to transfer information, such as tracer concentration, from one place to another. The method using constraints to generate perturbations can then easily take into account

a land-sea mask in the same way as diffusive fluxes of tracers. In continuous form, the operator D applied to a perturbation field ϕ can be written as:

$$D\phi = \nabla \cdot F \tag{13}$$

where the vector field F is given by:

$$\begin{aligned} F &= \nabla\phi \quad \text{for interior points} \\ F &= 0 \quad \text{at land-sea boundaries} \end{aligned} \tag{14}$$

The derivatives in (13) and (14) are approximated by centered finite differences, yielding a sparse matrix representation of \mathbf{D} . To facilitate the formulation of this sparse matrix, the operator is separated into sub-steps (calculating the vector field F , application of boundary conditions, calculating the divergence of the vector field). Each of the individual steps can be easily expressed as a sparse matrix. The matrix \mathbf{D} is then simply the product of all matrices corresponding to the sub-steps in the appropriate order. The same approach will also be used to create a sparse matrix representing the dynamical constraint.

To highlight the impact of this smoothness operator, we will not take into account a linear constraint in equation (2). Also, we chose the weighting matrices proportional to the identity matrix.

$$2J(\mathbf{x}) = L^4(\mathbf{D}\mathbf{x})^T(\mathbf{D}\mathbf{x}) + \mathbf{x}^T\mathbf{x} \tag{15}$$

The vector \mathbf{x} represents the values of the field ϕ on a set of grid points and the matrix \mathbf{D} is the discretization of the diffusion operator D . The exponent of L has been chosen such that L represents a length-scale. Indeed, this parameter is the horizontal correlation length of the perturbations. At this length scale both terms of the cost function have a similar magnitude (Brasseur et al., 1996). In the examples that will follow, the parameter has been adjusted such that the perturbations based on equation (15) have the same correlation length than the perturbations created using the Fourier transform without any land points. It may be worthwhile to note here that in this case the horizontal structure of the covariance is a Gaussian for the method using the Fourier transform; whereas the covariance of the perturbations based on the cost function is related to a modified Bessel function of order 1 in the continuous case, here discretized on a regular grid (Brasseur et al., 1996).

An ensemble of 10000 perturbations has been created with a horizontal correlation length of 20 km for a square domain with a narrow island. The ensemble covariance has been computed for a grid point near the end of the barrier (the black dot in figure 1). As expected, the structure of the ensemble covariance using the Fourier transform is isotropic. On the contrary, the covariance using the cost function is deformed by the presence of land points. Only grid points which can be connected

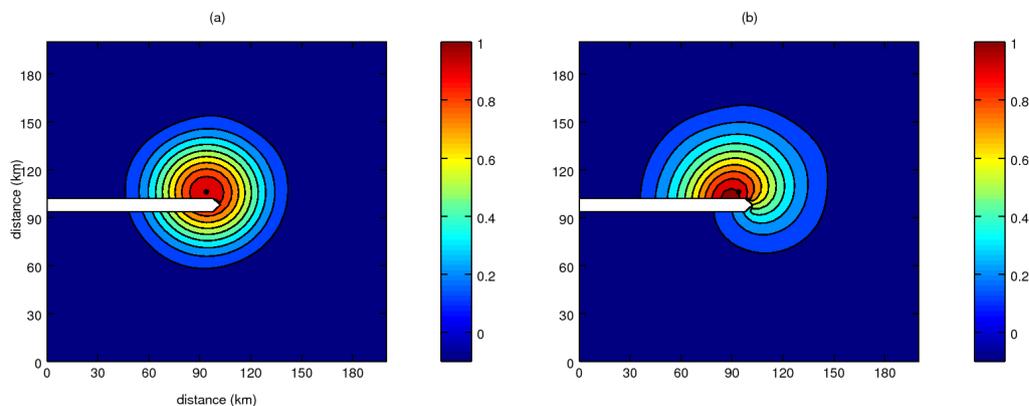


Fig. 1. Ensemble covariance using Fourier modes (a) and constrained perturbations based on the land-sea mask (b).

with a short path (relative to the correlation length-scale) not crossing land points have a significant correlation. In general, this is the expected physical behaviour.

185

This effect of land points is well known in the context of variational analysis where the covariance functions are constructed with differential operators (*e.g.* Brasseur, 1994). Here we use this desirable property for the generation of perturbations which can be used in sequential Kalman filters using ensembles.

190

It should be noted here that in the context of Ensemble Kalman filtering, covariance functions similar to panel (a) of Figure 1 are not directly used for assimilating observations. Perturbations with these covariance functions would be applied to the initial condition for example, and the covariance function will be inevitably transformed by the non-linear model (*e.g.* advected by currents).

195 However, a correlation across the barrier will still remain; only the shape of the correlation function is modified. Thus an observation made on one side of the barrier will also impact the other side.

The method based on the cost function can be applied to any model grid where the discrete diffusion operator can be formulated. This is an advantage compared to the Fourier method which can be
 200 only applied to rectangular grids and spherical grids (using spherical harmonics). For more general cases, the latter method requires an intermediate grid since the symmetry requirement (translation invariance) does not only apply to the domain but also to the model grid itself.

Another advantage of the presented method is that perturbations with a spatially varying correlation length can be created. Spatial structures of baroclinic flows in the ocean are related to the
 205 internal radius of deformation. This length-scale is often much smaller on the weakly stratified shelf

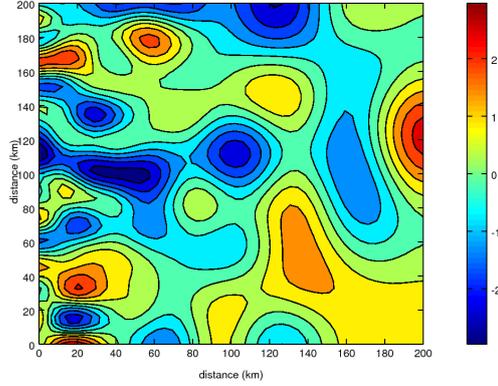


Fig. 2. Illustration of random field with a variable correlation length for a square domain without land points.

than in the deep ocean. A correlation length proportional to the internal radius of deformation is thus a reasonable choice. The effect of a spatially varying correlation length can be easily taken into account in the variational formulation (15). This is not possible with the Fourier method since
 210 a spatially varying correlation length is not translation-invariant. Figure 2 shows a random perturbation obtained for a correlation length varying linearly from 3 to 10 km in the zonal direction. As expected, the size of the structure gradually increases as the correlation length increases.

This aspect is also useful for a system of nested grid models and for models with unstructured
 215 grids. The length scale of the perturbations has to be properly resolved on the model grid. If only a unique length scale is used then it must be several times larger than the coarsest resolution. With a spatially varying correlation scale this limitation is lifted and areas with locally refined resolution can have smaller-scale perturbations. In a two-way nesting system, the model fields on different grids can also be regrouped in a single perturbation vector and the nesting feedback can be introduced as
 220 an additional constraint to ensure a coherent transition between model grids.

4 Application to tidal boundary conditions

This method is now applied to create multivariate perturbations of sea surface height and depth-averaged currents which have to be a harmonic solution of the shallow water equations. This example has a practical relevance since tides are governed by those equations. For simplicity, the shallow
 225 water equations are expressed here on a Cartesian grid, but they are implemented on a curvilinear grid as it will be applied to a curvilinear model grid.

$$\frac{\partial \zeta}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} = 0 \quad (16)$$

$$\frac{\partial u}{\partial t} - fv = -g \frac{\partial \zeta}{\partial x} \quad (17)$$

$$\frac{\partial v}{\partial t} + fu = -g \frac{\partial \zeta}{\partial y} \quad (18)$$

230 where h is the depth, f is the Coriolis parameter and g is the acceleration due to gravity. For simplicity we consider only one tidal constituent. The tidal component M2 is chosen because it is generally among the largest constituents. We require that the time dependency of the perturbations has the following form:

$$\zeta(x, y, t) = e^{i\omega t} \zeta'(x, y) \quad (19)$$

235 $u(x, y, t) = e^{i\omega t} u'(x, y) \quad (20)$

$$v(x, y, t) = e^{i\omega t} v'(x, y) \quad (21)$$

where ω is the angular frequency of the M2 tides. The perturbation vector \mathbf{x} is composed by the elevation ζ' and the depth-averaged currents u' and v' . It follows that,

$$i\omega \zeta' + \frac{\partial(hu')}{\partial x} + \frac{\partial(hv')}{\partial y} = 0 \quad (22)$$

240 $i\omega u' - fv' + g \frac{\partial \zeta'}{\partial x} = 0 \quad (23)$

$$i\omega v' + fu' + g \frac{\partial \zeta'}{\partial y} = 0 \quad (24)$$

The discrete operator M is obtained by discretizing the spatial derivatives (22)-(24) on an Arakawa C grid (Arakawa and Lamb, 1981) using finite volumes. Only wet points are included in the state vector \mathbf{x} . In this sense, boundary effects are already included in M . The normal velocity at the
245 land-sea boundary is prescribed to be zero and is not part of the state vector \mathbf{x} .

Open ocean boundary values are not constrained by the shallow water equations. If no smoothness constraint would be present, the resulting ensemble members would be discontinuous at the boundary in the direction parallel to the boundary. The explicit smoothness constraint (Laplacian
250 with a zero gradient at the open boundary in a direction perpendicular to the boundary) is added to the cost-function to avoid those discontinuities at the boundary.

The method is tested using the WFS ROMS model (Barth et al., 2008c) which is nested in the Atlantic HYCOM model (Chassignet et al., 2007). The nesting procedure is explained in Barth et al.
255 (2008a). The model uses a curvilinear grid with a resolution of about 3.5 km near the coast and 10 km near the open boundary. The model is initialized on the 1st January 2005 using the elevation, velocity, temperature and salinity from HYCOM. The boundary conditions from HYCOM mainly impose the path of the Gulf of Mexico Loop Current generating mesoscale eddies and filaments

inside the model domain (Barth et al., 2008b). The tidal boundary conditions produce tidal waves
 260 propagating inside the model domain and increase their amplitude near the coast as predicted by
 linear tidal wave theory for wide continental shelves (Clarke, 1991) and by a numerical ocean model
 of the WFS (He and Weisberg, 2002).

5 Experiments

265 Three methods, of increasing complexity, for the generation of perturbations of the tidal motions are
 examined:

- The elevation and the two horizontal velocities are perturbed independently. Instead of per-
 turbing the amplitude and phase directly, it is preferable to work here with the complex repre-
 sentation of these tidal parameters. The real and imaginary parts are perturbed independently.

270 The horizontal correlation length of the perturbations is 300 km which is the typical length-
 scale of the tidal maps. The complex perturbations are added to the TPXO6.2 (Egbert et al.,
 1994; Egbert and Erofeeva, 2002) complex tidal parameters and converted into amplitude and
 phase, as tides are usually characterized. If the amplitudes would be directly perturbed then
 one could not ensure that they remain positive.

- 275 – The elevation is perturbed as described in the previous experiment, but now the corresponding
 velocity perturbations are diagnosed using the momentum balance:

$$u = -\frac{g}{f^2 - \omega^2} \left(f \frac{\partial \zeta}{\partial y} + i\omega \frac{\partial \zeta}{\partial x} \right) \quad (25)$$

$$v = \frac{g}{f^2 - \omega^2} \left(f \frac{\partial \zeta}{\partial x} - i\omega \frac{\partial \zeta}{\partial y} \right) \quad (26)$$

280 Those equations become the geostrophic equilibrium for $\omega = 0$. The geostrophic equilibrium
 is often used to model error covariances in data assimilation (*e.g.* Dee, 1991; Brankart et al.,
 2003) and also to compute perturbations of the velocity based on the perturbation of the den-
 sity field (Barth et al., 2007a). Equations (25)-(26) are more complete than the geostrophic
 equilibrium since the former do not assume small accelerations.

- 285 – In the third method, perturbations are weakly constrained by the equations (22)-(24) following
 the approach described in section 2. The weighting matrix \mathbf{W} is defined by the energy E of
 shallow water waves:

$$\mathbf{x}^T \mathbf{W} \mathbf{x} = E = \frac{1}{2} \int_S g \zeta^2 + hu^2 + hv^2 dS \quad (27)$$

290 where on the right-hand side, the discretized version defines the weighting matrix (which ensures also proper dimensional relationships). Up to a constant factor ρ_0 , E corresponds to the total energy of a system governed by the shallow water equations.

The weighting matrices for the dynamical constraint \mathbf{W}_M , the smoothness constraint \mathbf{W}_D and the total norm \mathbf{W}_E are proportional to \mathbf{W} . Since the perturbations are scaled afterwards, one proportionality coefficient can be arbitrarily fixed without loss of generality. Here the matrices are defined by:

$$295 \quad \mathbf{W}_M = \omega^{-2} \mathbf{W} \quad (28)$$

$$\mathbf{W}_D = L^4 \mathbf{W} \quad (29)$$

$$\mathbf{W}_E = \alpha \mathbf{W} \quad (30)$$

300 where L (dimension of a length scale) and α (adimensional) are parameters to be chosen. The smaller these values are, the stronger the dynamical constraint will be enforced. However, we do not require that the constraint will be exactly satisfied since it is derived using a series of assumptions.

After several values of those parameters were tested, α was set to 0.001 and L to 10 km. The preferential length-scale taking only the smoothness and total energy constraint into account is thus $\alpha^{-1/4} L = 56$ km which is small compared to the length-scale of the tidal structures.
 305 This indicates that the structure of the perturbations is mainly determined by the dynamical constraint and not by the two other constraints.

When this method for generating ensemble perturbations is applied to data assimilation, one can use cross-validation (Wahba and Wendelberger, 1980) to estimate the parameters L and α more objectively in a similar way than parameters are optimized in *e.g.* DIVA (Data-Interpolating Variational Analysis) Brankart and Brasseur, 1996) and DINEOF (Data Interpolating Empirical Orthogonal Functions; Beckers and Rixen, 2003; Alvera-Azcárate et al., 2007): a small subset of observations are not used during the data assimilation and reserved for the validation of the results. The ensemble run and assimilation steps are repeated for different values of the parameters L and α . The best set of parameters is the one that minimizes the root mean squared error compared to the validation data set. This approach is not used in
 315 this present work but might be explored in further data assimilation studies.

The amplitude of the perturbations is chosen such that, in all the experiments, the expected value of the energy norm $\mathbf{x}^T \mathbf{W} \mathbf{x}$ is a given constant. Here we use $0.01 \text{ m}^3 \text{s}^{-1}$ times the total surface of

320 the model domain. This produces perturbations of the elevation of the order of cm which is comparable to the error of tidal models of this region (He and Weisberg, 2002). All methods produce thus perturbations which are equally energetic.

The M2 tidal parameters for the elevation and the depth-averaged velocity are obtained from the
325 TPXO6.2 global inverse tide model. Those tidal parameters are perturbed using the three different methods. For each method, an ensemble of 32 members has been created. The tidal surface elevation and depth-averaged velocity are added to the initial conditions and to the HYCOM boundary conditions. All ensemble members are integrated for 10 days with the full, non-linear ROMS model.

330 The first method produces isotropic covariances, but this is not the case for the constrained perturbations since they have to satisfy (approximately) the shallow water equations. The horizontal covariance of the constrained perturbations for a point near the open boundary is shown in Figure 3. This covariance represents how a hypothetical error (or an observation if this covariance is used for data assimilation) near the boundary would affect the solution within the model domain. In the
335 vicinity of this point the covariance decreases monotonically with a preferred direction to the North-East inward to the model domain. However the covariance increases in Florida's Big Bend and near the Mississippi Delta. Thus the impact of an error in the boundary is not only local but it can be seen remotely due to the propagation of tidal waves. To include such aspects in error covariance models, dynamical constraints have to be taken into account.

340 **6 Results**

To compare the three different examined methods, it is necessary to quantify the amount of transient motions created due to the perturbations. The surface elevation of an unperturbed (central) run was subtracted from each ensemble member to isolate the impact of the initial and boundary perturbations. A tidal analysis was performed to subtract the tidal variations. This residual is then averaged
345 over the 10-day time period which is shown in figure 4. To compute the total amount of transient motions those maps are averaged in space and the results are given in table 1. If this residual would be zero, then the perturbations would have only created tidal waves at the M2 frequency (unless the amplitude of the perturbations is so large that it creates harmonics).

350 The largest transient motions are created by perturbing elevation and velocity independently (panel (a) on Figure 4). The structure of the standard deviation of the elevation residual is quite similar to the amplitude of the M2 tide in the WFS. In cases 1 and 2, the highest values are observed in Florida's Big Bend and in the southern part of the domain. This suggest that the amplitude of the transient motions increases near the coast in the same way as the barotropic tidal waves increase its

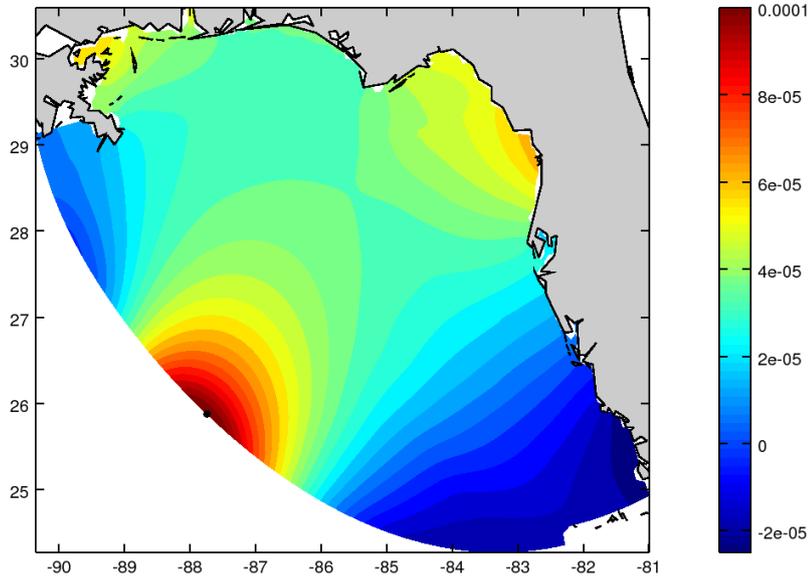


Fig. 3. Horizontal covariance of the constrained perturbations between the point near the open boundary marked by a black dot and all other grid points.

355 amplitude when the depth decreases and when the wave is reflected from the coastline. The fact that the transient motions are amplified near the coast is particularly problematic since this is often the region of interest for regional models.

To quantify the overall magnitude of the transient motions, the variance of the residual is averaged
 360 over the model domain. The square root of the averaged variance (or average standard deviation) of all three experiments is shown in table 1. The experiment with independent perturbations has indeed the largest average variance of transient motions.

Using the momentum equation to compute the velocity perturbation based on the elevation pertur-
 365 bation does not significantly improve the dynamical balance of the perturbed model simulation since the average standard deviation is only slightly reduced. This result is surprising since the geostrophic equilibrium, which is related to equations (25) - (26), is used successfully in data assimilation. The difference here is that the (pure) geostrophic equilibrium does satisfy the horizontal continuity equation (on an f-plane) and does not require vertical motions. But this is not the case for equations (25)
 370 - (26) which can thus entrain a movement of the free surface generating spurious gravity waves.

The smallest transient motions are obtained by using the constrained perturbations. The variance

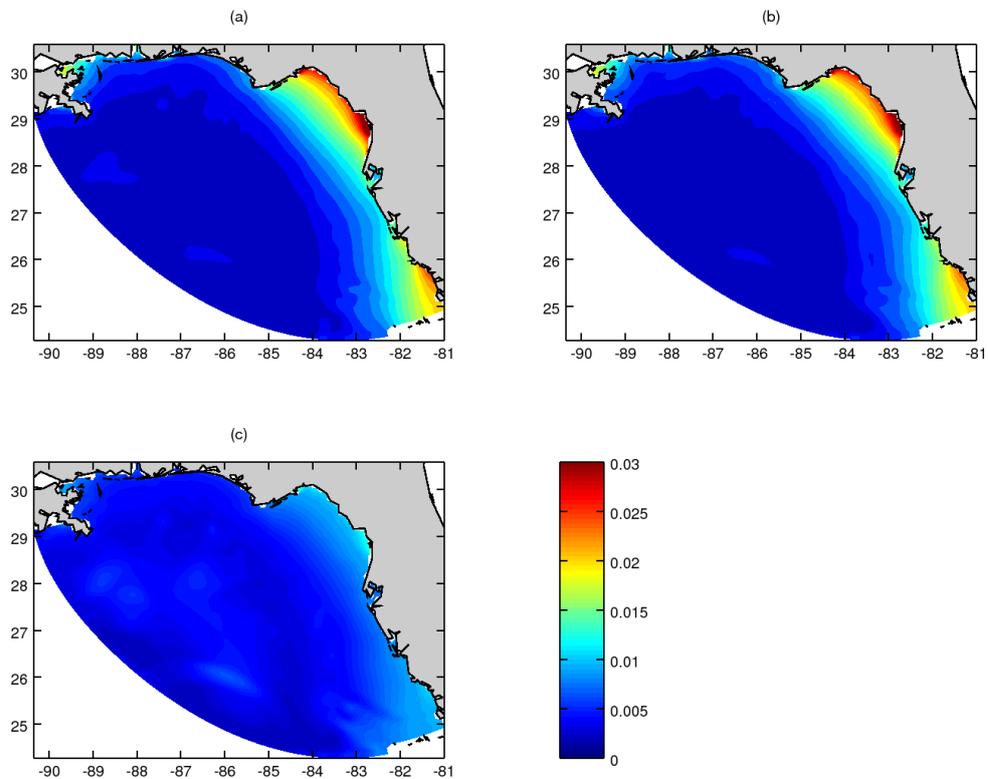


Fig. 4. Standard deviation of the elevation residual. (a) using independent perturbations, (b) using the momentum equation and (c) using constrained perturbations.

Experiment	description	average standard deviation
1	independent perturbations	0.00763
2	perturbations using the momentum balance	0.00753
3	dynamically constrained perturbations	0.00480
2bis	same as 2 but with bottom drag	0.00707
3bis	same as 3 but with bottom drag	0.00430

Table 1. The average standard deviation (in m) of the elevation residual. This quantity indicates the magnitude of the transient motions created by different perturbations.

of the residual is reduced by 60% compared to the previous two approaches. While the standard deviation of the residual is still the largest near the coast compared to the open ocean, it is now significantly reduced. This shows that although the perturbations are derived using several assumptions (in particular, no bottom drag, purely barotropic and linear dynamics), they are applicable to an implementation of a baroclinic and non-linear model with realistic friction.

6.1 Impact of bottom drag

In the previous section, only linear terms were examined for methods 2 and 3. The bottom drag was thus ignored. However, it is well known that it plays an important role for modelling tides in shelf seas. Including the bottom friction, the governing equations are:

$$\frac{\partial \zeta}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} = 0 \quad (31)$$

$$\frac{\partial}{\partial t} u - fv = -g \frac{\partial \zeta}{\partial x} - \frac{\tau_x}{h} \quad (32)$$

$$\frac{\partial}{\partial t} v + fu = -g \frac{\partial \zeta}{\partial y} - \frac{\tau_y}{h} \quad (33)$$

The WFS ROMS uses a quadratic parametrization for the bottom drag τ :

$$\tau = r|\mathbf{u}|\mathbf{u} \quad (34)$$

where the drag coefficient r is 10^{-3} . This bottom drag is linearized around the velocity vector (\bar{u}, \bar{v}) :

$$\tau_x = \left(r \frac{2\bar{u}^2 + \bar{v}^2}{\sqrt{\bar{u}^2 + \bar{v}^2}} \right) u \quad (35)$$

$$\tau_y = \left(r \frac{\bar{u}^2 + 2\bar{v}^2}{\sqrt{\bar{u}^2 + \bar{v}^2}} \right) v \quad (36)$$

The velocities \bar{u} and \bar{v} are set to the half of the maximum tidal velocity in the x and y directions respectively as this corresponds to the average of the tidal velocity squared. The dynamical constraint with bottom drag becomes:

$$i\omega \zeta' + \frac{\partial(hu')}{\partial x} + \frac{\partial(hv')}{\partial y} = 0 \quad (37)$$

$$i\omega u' - fv' + g \frac{\partial \zeta'}{\partial x} + c_u u' = 0 \quad (38)$$

$$i\omega v' + fu' + g \frac{\partial \zeta'}{\partial y} + c_v v' = 0 \quad (39)$$

where c_u and c_v are given by:

$$c_u = \left(r \frac{2\bar{u}^2 + \bar{v}^2}{\sqrt{\bar{u}^2 + \bar{v}^2}} \right) / h \quad (40)$$

$$c_v = \left(r \frac{\bar{u}^2 + 2\bar{v}^2}{\sqrt{\bar{u}^2 + \bar{v}^2}} \right) / h \quad (41)$$

Methods 2 and 3 are repeated by including a linearized drag term. With the drag term, the equations (25) and (26) become:

$$u = -A \left(f \frac{\partial \zeta}{\partial y} + (i\omega + c_v) \frac{\partial \zeta}{\partial x} \right) \quad (42)$$

$$v = A \left(f \frac{\partial \zeta}{\partial x} - (i\omega + c_u) \frac{\partial \zeta}{\partial y} \right) \quad (43)$$

where A is given by,

$$405 \quad A = \frac{g}{f^2 - \omega^2 + c_u c_v + i\omega(c_u + c_v)}. \quad (44)$$

From this denominator, it is evident that the drag reduces in average the size of the velocity perturbations (for a given perturbation of the elevation) and alters also the phase.

Figure 5 shows the standard deviation of the residual using perturbations taking the bottom drag
 410 into account. In panel (a), the perturbations we created similar to experiment 2 but using the equations (42) and (43). The results in panel (b) are obtained by perturbations constrained by the spatial discretization equations (37)-(39) similar to experiment 3.

The spatial structure of the residual is comparable to the residual in the previous experiments with
 415 an increase of the amplitude of transient motions near the coast. However, the amplitude is indeed reduced in both cases, which confirms our expectation that the more complete our dynamical constraint is, the more balanced our perturbations are. The amplitude is actually reduced by approximately the amount relative to the corresponding experiment without bottom drag. The smallest transient motions are generated by using the perturbations constrained by the shallow water equations including
 420 the bottom drag.

6.2 Motion induced by the constrained perturbations

The main purpose of the proposed procedure is the generation of balanced perturbations. In the
 previous section the magnitude of the transition motions were used to quantify how close the model
 state is to an equilibrium state. To further study the tidal perturbations we can compare the amplitude
 425 and phase of the perturbations with the corresponding parameters diagnosed from a perturbed run. Since the model is derived from the non-linear baroclinic equations, we will examine how relevant the simplified shallow water equations are compared to a realistic model. Also, the constraints still use continuous time while the model is discretized in time.

430 Panel (a) of figure 6 shows the elevation of a realization of the constrained perturbations. The colors represent the amplitude and the isolines are the phase. This perturbation shares indeed common characteristics known from tidal propagation. The tidal amplitude is amplified in some regions on the shelf. It also contains an amphidromic point in the south eastern part of the model domain.

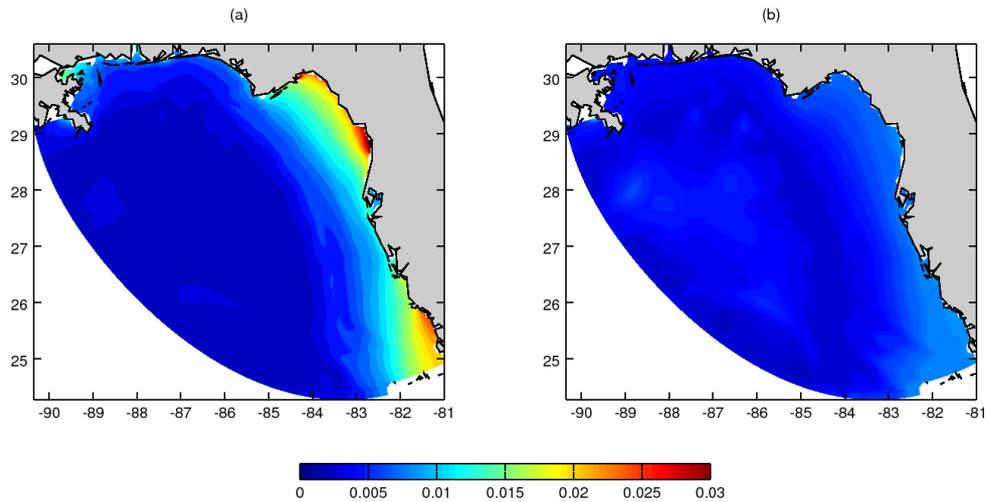


Fig. 5. Standard deviation of the elevation residual (a) using the momentum equation and (b) using constrained perturbations. Bottom drag is taken into account.

435 This perturbation (and the corresponding velocity perturbation) is applied to initial conditions and boundary conditions of the WFS ROMS model. The model is then integrated for 10 days using realistic forcings. The amplitude and phase of the difference between the elevation of the perturbed run and the unperturbed run is computed and shown in panel (b) of figure 6.

440 The location and value of the amplitude maxima in the perturbed model run are comparable to the applied perturbation. The model results contain also the amphidromic point as it can be seen from the phase. However, some differences in phase can be seen especially in Florida's Big Bend, where the phase difference is about 40° . The agreement between both fields is reasonable given the number of assumptions that were necessary for the shallow water equations that were not made in the WFS
 445 ROMS.

In the context of an ensemble forecast, if one ensemble member is particularly close to the observations (closer for example than the central, unperturbed simulation), we can therefore assume that the perturbations produce indeed initial and boundary conditions which are more realistic than the
 450 unperturbed initial and boundary conditions. Formally, this can be done by including the perturbation in the data assimilation state vector. The analysis will then provide an improved estimation of those boundary conditions.

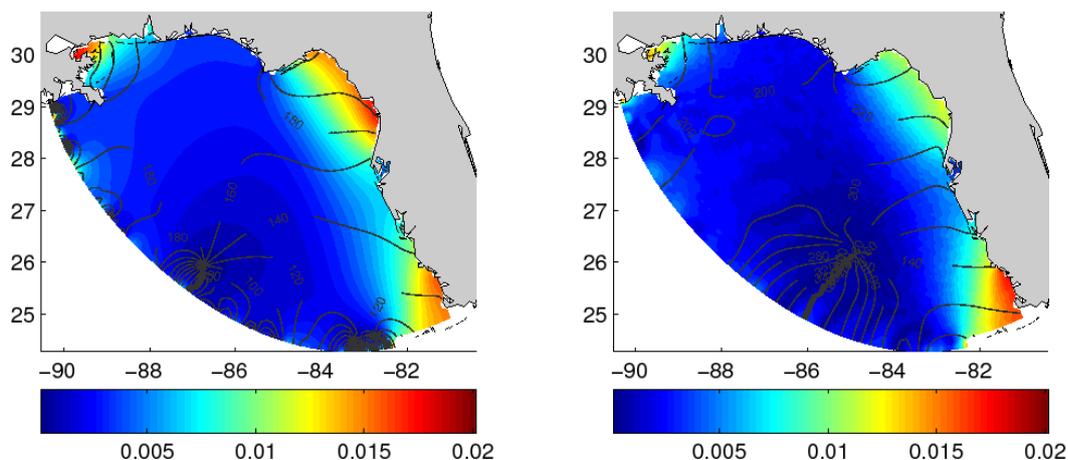


Fig. 6. Left panel: a realization of a constrained perturbation. Only the elevation amplitude and phase are shown. Right panel: tidal analysis of the difference between the elevation of the perturbed and the unperturbed simulation.

7 Numerical cost

To assess the numerical cost of this scheme and to test its feasibility with a high resolution ocean
 455 model, the method was applied to a square domain of 200 km length with different resolutions. In a
 first series of experiments only 50 eigenvectors are retained during the eigenvector decomposition.
 All other parameters were the same than in the perturbation for the WFS case.

The code is tested on a single core of an Intel Xeon E5420 CPU. The code is run in Octave 3.0.5
 460 compiled among others with SuiteSparse 3.4.0 (Davis, 2004a,b), GotoBLAS 1.26 and ARPACK 96
 (Lehoucq et al., 1997). Those libraries are used in the eigenvector decomposition. The time in sec-
 onds of different steps in the algorithm are shown in table 2 for different grid sizes. A domain size of
 512x512 was tested but the method required more than the available 16 GB of RAM. As expected,
 the creation of the matrix \mathbf{B}^{-1} , and the ensemble creation increases essentially linearly with the
 465 number of grid points while the eigenvector decomposition increases faster than linearly with the
 number of grid points. The slope of a linear regression in log-log space is 1.3. This progression is
 still quite similar to a linear increase because the matrix \mathbf{B}^{-1} is sparse and because a fixed number
 of eigenvectors are retained.

470 In a second series of tests, the domain size is fixed (128x128 grid points) and the number of eigen-
 vectors are increased from 50 to 200 (table 3). The numerical cost of the eigenvector decomposition
 shows a linear progression relative to the number of eigenvectors. The numerical cost of the ensem-
 ble creation increases only slowly and is marginal in the overall cost.

Size of domain	Creation of matrix \mathbf{B}^{-1}	Eigenvector decomposition	Ensemble
32x32	0.28	1.31	0.09
64x64	0.38	8.77	0.39
128x128	1.36	51.23	1.69
256x256	5.67	296.28	7.07
300x300	7.96	451.66	9.79
400x400	14.77	983.77	18.36

Table 2. Time in seconds as a function of the number of grid points of the different steps involved in the creation of the perturbations.

Number of eigenvectors	Eigenvector decomposition	Ensemble
50	51.23	1.69
100	112.87	2.47
150	151.90	3.42
200	227.19	4.60

Table 3. Time in seconds as a function of number of eigenvector retained for a domain of 128x128 grid points.

475 By increasing the resolution of a domain, one can argue that the number of eigenvectors and ensemble members should also increase proportional to the number of grid points. In this case, the progression rate of both series have to be combined and the numerical cost scales approximately by the number of grid points elevated to 2.3.

480 The program code has been written such that it can run unmodified also on MATLAB. The benchmark was repeated on the same machine with MATLAB R2008a (64-bit version). The numerical cost in function of the number of grid points and the number of eigenvectors retained varied in a similar way than with Octave. Overall, Octave was 13% faster than MATLAB in completing the two series of tests. In summary, the CPU time of this method is acceptable since it is very small
485 compared to the CPU time needed for the ensemble run. More limiting than the CPU time, can be the required amount of RAM memory for large model configurations.

8 Conclusions

A new method to produce ensemble perturbations is presented. It allows to take linear constraints into account. The formalism similar to weak constraints in variational analysis can be used in an ensemble scheme for generating ensemble perturbations. The constraints can be chosen such that
490 the perturbations are dynamically balanced. This is useful for ensemble forecasts and in particular for data assimilation where the ensemble spread is supposed to reflect the uncertainty in the model forecast and should not include the variability of transient motions generated by an unbalanced state.

495 The method has been tested to create elevation and velocity perturbations of tidal motions. The shallow water equations constitute a constraint that relate elevation and velocity. In a first test, those variables are perturbed independently. In a second simulation of intermediate complexity, only the momentum equation is enforced. This procedure allows to derive the velocity perturbation from an elevation perturbation. The third tests uses the linear shallow water equations (momentum equations
500 and continuity equations) to create the perturbations. For each of those simulations, the amount of transient motions created by the tidal analysis is computed by applying a tidal analysis. In all cases the amplitude of the transient motions were highest near the coast. The transient motions created using the linear shallow water equations were significantly smaller than those generated with the two other methods showing that the constrained ensemble was dynamically more balanced. A further
505 reduction of the transient motions could be obtained by including the effect of bottom drag into the dynamic constraint. The importance of drag on tidal dynamics is well known, but in this context it highlights the fact that important non-linear terms cannot be simply neglected but must be linearized around an appropriate mean-state.

510 The challenge of this approach will be to formalize an appropriate dynamical constraint to quantify if a perturbation is realistic or not. It will require a good knowledge of the dynamical behavior of the studied system and of its error sources. Some constraints (involving for example density) are non-linear. But for small errors, a useful local linearization can be generally obtained.

515 Even if no linear constraint or balance can be formulated, the presented procedure can still be useful to create perturbations that are aware of the land-sea boundary. Spurious correlation across land points are thus avoided. Also, a spatially varying correlation length can be used with this method. For example, one can specify the correlation length as a multiple of the radius of deformation. Those two aspects are not possible in the Fourier-based method to generate perturbations. Despite the numerical cost is higher than for the Fourier-based method, its cost is still acceptable for most model
520 setups.

Althought only two-dimensional perturbations have been shown, the ensemble generation code is written in a way that it can generate n-dimensional perturbations given a user-specified dynamical
525 constraint. The source code which runs on MATLAB and GNU Octave is freely available at <http://modb.oce.ulg.ac.be/mediawiki/index.php/WCE>.

In future works, the method will be used for assimilating surface current observations. By including the boundary perturbations into the model state vector, the assimilation can provide also an
530 improved estimation of the boundary values.

Acknowledgements. This work was supported by the “Global Ocean Prediction with the Hybrid Coordinate Ocean Model (HYCOM)” program and the “HYCOM Coastal Ocean Hindcasts and Predictions: Impact of Nesting in HYCOM GODAE Assimilative Hindcasts” program of the US Office of Naval Research (Grant numbers: N00014-04-1-0676 and N00014-99-1-1051 respectively) and the ECOOP project (European Coastal-shelf sea Operational observing and forecasting system) of the European Union. The National Fund for Scientific Research, Belgium is acknowledged for funding the post-doctoral positions of Alexander Barth and Aida Alvera Azcárate. The Center for Prediction of Red Tides is acknowledged for granting computer time used in this study. We thank also the US Naval Research Laboratory for providing HYCOM model results used as initial and boundary conditions. G.D Egbert and S.Y. Erofeeva are acknowledged for providing the TPXO6.2 global inverse tide model solution. This is MARE publication 174.

References

- Alvera-Azcárate, A., Barth, A., Beckers, J., and Weisberg, R. H.: Multivariate reconstruction of missing data in sea surface temperature, chlorophyll and wind satellite field, *Journal of Geophysical Research*, 112, C03 008, doi:10.1029/2006JC003660, 2007.
- 545 Arakawa, A. and Lamb, V.: A potential enstrophy and energy conserving scheme for the shallow water equations, *Monthly Weather Review*, 109, 18–36, 1981.
- Barth, A., Alvera-Azcárate, A., Beckers, J.-M., Rixen, M., and Vandenbulcke, L.: Multigrid state vector for data assimilation in a two-way nested model of the Ligurian Sea, *Journal of Marine Systems*, 65, 41–59, doi:10.1016/j.jmarsys.2005.07.006, 2007a.
- 550 Barth, A., Beckers, J.-M., Alvera-Azcárate, A., and Weisberg, R. H.: Filtering inertia-gravity waves from the initial conditions of the linear shallow water equations, *Ocean Modelling*, 19, 204–218, doi:10.1016/j.ocemod.2007.06.007, 2007b.
- Barth, A., Alvera-Azcárate, A., and Weisberg, R. H.: Benefit of nesting a regional model into a large-scale ocean model instead of climatology. Application to the West Florida Shelf, *Continental Shelf Research*, 28, 561–573, doi:10.1016/j.csr.2007.11.004, 2008a.
- 555 Barth, A., Alvera-Azcárate, A., and Weisberg, R. H.: A nested model study of the Loop Current generated variability and its impact on the West Florida Shelf, *Journal of Geophysical Research*, 113, C05 009, doi:10.1029/2007JC004492, 2008b.
- Barth, A., Alvera-Azcárate, A., and Weisberg, R. H.: Assimilation of high-frequency radar currents in a nested model of the West Florida Shelf, *Journal of Geophysical Research*, 113, C08 033, doi:10.1029/2007JC004585, 2008c.
- 560 Beckers, J.-M. and Rixen, M.: EOF calculation and data filling from incomplete oceanographic datasets, *Journal of Atmospheric and Oceanic Technology*, 20, 1839–1856, 2003.
- Brankart, J.-M. and Brasseur, P.: Optimal analysis of in situ data in the Western Mediterranean using statistics and cross-validation., *Journal of Atmospheric and Oceanic Technology*, 13, 477–491, 1996.
- 565 Brankart, J. M., Testut, C. E., Brasseur, P., and Verron, J.: Implementation of a multivariate data assimilation scheme for isopycnic coordinate ocean models: application to a 1993-96 hindcast of the North Atlantic Ocean circulation, *Journal of Geophysical Research*, 108, 3074, 2003.
- Brasseur, P.: A Variational Inverse Method for the Reconstruction of General Circulation Fields in the Northern Bering Sea., *Journal of Geophysical Research*, 96, 4891–4907, 1991.
- 570 Brasseur, P.: Reconstitution de champs d’observations océanographiques par le Modèle Variationnel Inverse: Méthodologie et Applications., Ph.D. thesis, Université de Liège, Collection des publications, Sciences appliquées, 1994.
- Brasseur, P., Beckers, J.-M., Brankart, J.-M., and Schoenauen, R.: Seasonal Temperature and Salinity Fields in the Mediterranean Sea: Climatological Analyses of an Historical Data Set., *Deep-Sea Research*, 43, 159–192, 1996.
- 575 Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., Baraille, R., and Bleck, R.: The HYCOM (HYbrid Coordinate Ocean Model) Data Assimilative System, *Journal of Marine Systems*, 65, 60–83, 2007.
- 580 Clarke, A. J.: Tidal Hydrodynamics, chap. The dynamics of barotropic tides over the continental shelf and slope

- (review), John Wiley and Sons, 1991.
- Davis, T. A.: A column pre-ordering strategy for the unsymmetric-pattern multifrontal method, *ACM Transactions on Mathematical Software*, 30, 165–195, 2004a.
- Davis, T. A.: Algorithm 832: UMFPACK, an unsymmetric-pattern multifrontal method, *ACM Transactions on*
585 *Mathematical Software*, 30, 196–199, 2004b.
- Dee, D.: Simplification of the Kalman filter for meteorological data assimilation, *Quarterly Journal of the Royal Meteorological Society*, 117, 365–384, 1991.
- Dobricic, S., Pinardi, N., Adani, M., Tonani, M., Fratianni, C., Bonazzi, A., and Fernandez, V.: Daily oceanographic analyses by Mediterranean Forecasting System at the basin scale, *Ocean Science*, 3, 149–157, 2007.
- 590 Egbert, G. and Erofeeva, S.: Efficient inverse modeling of barotropic ocean tides, *J. Atmos. Oceanic Technol.*, 19, 183–204, 2002.
- Egbert, G., Bennett, A., and Foreman, M.: TOPEX/POSEIDON tides estimated using a global inverse model, *Journal of Geophysical Research*, 99, 24 821–24 852, 1994.
- Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, *Ocean Dy-*
595 *namics*, 53, 343–367, 2003.
- Fukumori, I.: A partitioned Kalman filter and smoother, *Monthly Weather Review*, 130, 1370–1383, 2002.
- He, R. and Weisberg, R. H.: Tides on the West Florida Shelf, *Journal of Physical Oceanography*, 32, 3455–3473, 2002.
- Kalnay, E.: *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, 1 edn.,
600 2002.
- Lehoucq, R. B., Sorensen, D. C., and Yang, C.: *ARPACK user’s guide: solution of large scale eigenvalue problems with implicitly restarted Arnoldi methods*, Computational and Applied Mathematics, Rice University, <http://www.caam.rice.edu/software/ARPACK/>, 1997.
- Lermusiaux, P., C.-S., C., Gwarkiewicz, G., Abbot, P., Robinson, A., Miller, R., Haley, P., Leslie, W.,
605 Majumdar, S., Pang, A., and Lekien, F.: Quantifying uncertainties in ocean predictions, *Oceanography*, Special issue on “Advances in Computational Oceanography”, 19, 90–103, 2006.
- Lermusiaux, P. F. J. and Robinson, A. R.: Data assimilation via error subspace statistical estimation. Part II. Middle Atlantic Bight shelfbreak front simulations and ESSE validation., *Monthly Weather Review*, 127, 1408–1432, 1999.
- 610 Pham, D. T.: Stochastic methods for sequential data assimilation in strongly nonlinear systems, *Monthly Weather Review*, 129, 1194–1207, 2001.
- Vallis, G. K.: Mechanisms and Parameterizations of Geostrophic Adjustment and a Variational Approach to Balanced Flow, *Journal of Atmospheric Science*, 49, 1144–1160, 1992.
- Wahba, G. and Wendelberger, J.: Some new mathematical methods for variational objective analysis using
615 splines and cross validation., *Monthly Weather Review*, 108, 1122–1143, 1980.
- Weaver, A. T. and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, *Quarterly Journal of the Royal Meteorological Society*, 127, 1815–1842, 2001.