Boundary conditions control for a Shallow-Water model

Eugene Kazantsev

▶ To cite this version:

HAL Id: hal-00764540
https://hal.inria.fr/hal-00764540

Submitted on 13 Dec 2012

Boundary conditions control for a Shallow-Water model.

Eugene Kazantsev [*]

December 13, 2012

Abstract

A variational data assimilation technique was used to estimate optimal discretization of interpolation operators and derivatives in the nodes adjacent to the rigid boundary. Assimilation of artificially generated observational data in the shallow-water model in a square box and assimilation of real observations in the model of the Black sea are discussed. It is shown in both experiments that controlling the discretization of operators near a rigid boundary can bring the model solution closer to observations as in the assimilation window and beyond the window. This type of control allows also to improve climatic variability of the model.

Keywords: **Variational Data Assimilation; Boundary conditions; Shallow water model; Black sea model.**

1   Introduction

Variational data assimilation technique, first proposed in [1], [2], is based on the optimal control methods [3] and perturbations theory [4]. This technique allows us to retrieve an optimal data for a given model from heterogeneous observation fields. Since the early 1990's, many mathematical and geophysical teams are involved in the development of the data assimilation strategy.

---

[*]INRIA, projet MOISE, Laboratoire Jean Kuntzmann, BP 53, 38041 Grenoble Cedex 9, France

One can cite many papers devoted to this problem, as in the domain of development of different methods for the data assimilation and also in the domain of its applications to the atmosphere and oceans.

In the beginning, data assimilation methods were intended to identify and reconstruct an optimal initial state for the model. However, the idea that other model's parameters should also be identified by data assimilation has also been studied and discussed in numerous papers. One can cite several examples of using data assimilation to identify the bottom topography of simple models ([5], [6], [7]), to control open boundary conditions in coastal and regional models ([8], [9], [10], [11]) and to determine other parameters of a model ([12], [13], [14]) .

Together with these parameters, it seems to be interesting also to control boundary conditions on rigid boundaries. Although the boundary configuration of the ocean is steady and can be measured with much better accuracy than the model's initial state, it is not obvious how to represent it on the model's grid because of limited resolution. The coastal line of continents possesses a very fine structure and can only be roughly approximated by the model's grid. Consequently, boundary conditions are defined at the model grid's points which are different from the coast. Even the most evident impermeability condition being placed at a wrong point may lead to some error in the model's solution. If it can improve the model solution, we may accept the flux can cross the boundary in places where the boundary is in water, prescribing some integral properties on the flux.

Ocean models frequently include strong and thin boundary currents with intense velocity gradients. In this case, particular attention must be paid to the approximation of the boundary layer because a wrong representation of these currents may be responsible for drastic deformations of the global solution (see, for example [15]). This may lead us to control the discretization of the model's operators in the whole boundary layer rather than only at nodes directly adjacent to the boundary.

Alternative method that is frequently used in geophysical models consists in the grid refinement in zones where boundary layers might occur. However, this implies additional computational cost on each model run. Variational data assimilation may help us to determine the parametrization of the boundary layer once for all model runs and to save the computer time.

Boundary conditions on the rigid boundaries have been controlled by data assimilation for heat equation (see, for example, [16], [17]), but this control concerns the linear parabolic diffusion operator rather than hyperbolic trans-

port and advection operators that are more important in geophysical models.

Studies on the possibility to control boundary conditions on rigid boundaries for equations containing hyperbolic operators can be found in [18] on the example of non-linear balance equation, in [19] on the example of the wave equation, in [20] on the example of a linear shallow water model in a square box and in [21] and [22] on the example of the Burgers equation. The principal possibility to improve the model's solution controlling its boundary values are shown in all these papers. However, as it has been noted in [21], particular attention must be paid to the discretization process which must respect several rules. It is the discretization of the model's operators that takes into account the set of boundary conditions and introduces them into the model. Consequently, instead of controlling boundary conditions themselves, there has been proposed in [20] to identify optimal discretization of differential operators at points adjacent to boundaries. This allows us to control directly the way the boundary conditions influence the model and to control boundary parameters in a more general way. Boundary conditions participate in discretized operators, but considering the discretization itself, we take into account additional parameters like the position of the boundary, lack of resolution of the grid, etc. In [19], for example, it was shown that deplacing the boundary helps to correct numerical error resulting in a wrong wave velocity and this displacement has been directly introduced in the discretization of derivatives.

In this paper, we apply 4D-Var data assimilation to control the discretization of derivatives and interpolation operators in the boundary regions of a non-linear shallow-water model. We use methods proposed in [20] and study the assimilation results obtained with a model that exhibits a chaotic behavior. Two examples are considered in this paper: an academic case of the model in a square box with artificially generated observations and more realistic case of assimilation of real observational data in the Black sea model.

2 Shallow Water Model

2.1 Model's equations and discretization

The data are assimilated into the shallow-water model on the β-plane [23], [24]:

$$
\begin{aligned}
\frac{\partial u}{\partial t} - (f_0 + \beta y - \frac{\partial u}{\partial y})v + \frac{\partial(u^2/2 + g\eta)}{\partial x} &= -\sigma u + \mu \Delta u + \frac{\tau_x}{\rho_0 H_0}, \\
\frac{\partial v}{\partial t} + (f_0 + \beta y + \frac{\partial v}{\partial x})u + \frac{\partial(v^2/2 + g\eta)}{\partial y} &= -\sigma v + \mu \Delta v + \frac{\tau_y}{\rho_0 H_0}, \quad (1) \\
\frac{\partial \eta}{\partial t} + \frac{\partial \eta u}{\partial x} + \frac{\partial \eta v}{\partial y} &= 0.
\end{aligned}
$$

where $u(x, y, t)$ and $v(x, y, t)$ are two velocity components, $\eta(x, y, t)$ is the sea surface elevation, ρ_0 is the mean density of water, H_0 is the characteristic depth of the basin and g is the reduced gravity. The model is driven by the surface wind stress with components $\tau_x(x, y)$ and $\tau_y(x, y)$ and subjected to the bottom drag that is parametrized by linear terms σu, σv. Horizontal eddy diffusion is represented by harmonic operators $\mu \Delta u$ and $\mu \Delta v$. Coriolis parameter is supposed to be linear in y coordinate and is presented as $(f_0 + \beta y)$. The system is defined in some domain Ω with characteristic size L requiring that both u and v vanish on the whole boundary of Ω. No boundary conditions is prescribed for η. Initial conditions are defined for all variables u, v and η.

We discretize all variables of this equation on the regular Arakawa's C-grid [25] with constant grid step $h = \frac{L}{N}$ in both x and y directions (see fig.1)

$$
\begin{aligned}
u_{i,j-1/2}(t) &= u(ih, jh - h/2, t) \text{ for } i = 0, \ldots N, j = 0, \ldots, N+1 \\
v_{i-1/2,j}(t) &= v(ih - h/2, jh, t) \text{ for } i = 0, \ldots N+1, j = 0, \ldots N \\
\eta_{i-1/2,j-1/2}(t) &= \eta(ih - h/2, jh - h/2, t) \text{ for } i = 0, \ldots N+1, j = 0, \ldots N+1
\end{aligned}
$$

In order to discretize the system (1), we replace the derivatives by their finite difference representations D_x and D_y and introduce two interpolations in x and y coordinates S_x and S_y. Interpolations are necessary on the staggered grid to calculate the variable's values in nodes where other variables
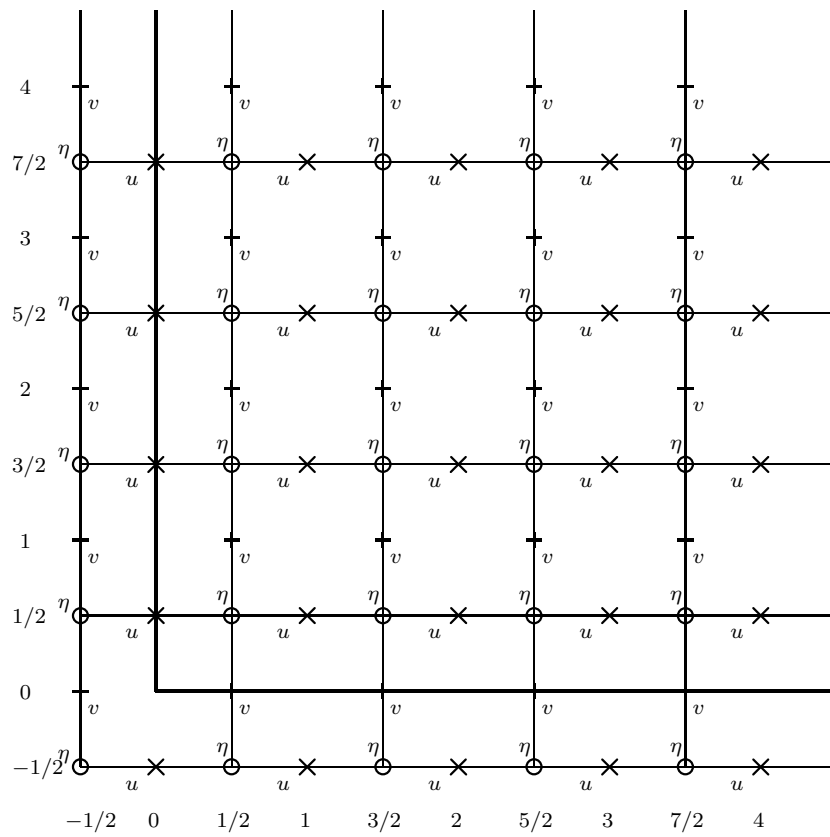
4

Figure 1: Arakawa C-grid

are defined. Following [26], we write the discretized system (1) as

$$\frac{\partial u}{\partial t} - (f_0 + \beta y - S_y D_y u) S_x S_y v + D_x((S_x u)^2/2 + g\eta) = -\sigma u + \mu \Delta^h u + \frac{\tau_x}{\rho_0 H_0},$$

$$\frac{\partial v}{\partial t} + (f_0 + \beta y + S_x D_x v) S_y S_x u + D_y((S_y v)^2/2 + g\eta) = -\sigma v + \mu \Delta^h v + \frac{\tau_y}{\rho_0 H_0} \quad (2)$$

$$\frac{\partial \eta}{\partial t} + D_x(u S_x \eta) + D_y(v S_y \eta) = 0.$$

Discretized operators D_x, D_y and S_x, S_y are defined in a classical way at all internal points of the domain. For example, the derivative and the interpolation operator of the variable u defined at corresponding points write

$$(D_x u)_{i-1/2,j-1/2} = \frac{u_{i,j-1/2} - u_{i-1,j-1/2}}{h} \text{ for } i = 2, \ldots, N-1,$$

$$(S_x u)_{i-1/2,j-1/2} = \frac{u_{i,j-1/2} + u_{i-1,j-1/2}}{2} \text{ for } i = 2, \ldots, N-1. \quad (3)$$

These expressions represent a well known second order approximation. Laplace operator is discretized in a common way $\Delta^h v = \frac{v_{i+1,j} + v_{i-1,j} - 2v_{i,j}}{h^2} + \frac{v_{i,j+1} + v_{i,j-1} - 2v_{i,j}}{h^2}$.

The boundary region is considered as a band of adjacent to boundary nodes. Discretizations of operators in this band are different from (3) and represent the control variables in this study. In order to obtain their optimal values assimilating external data, we suppose nothing about derivatives near the boundary and write them in a general form

$$(D_x u)_{1/2,j-1/2} = \frac{\alpha_0^{D_x u} + \alpha_1^{D_x u} u_{0,j-1/2} + \alpha_2^{D_x u} u_{1,j-1/2}}{h} \quad (4)$$

This formula represents a linear combination of values of u at two points adjacent to the boundary with coefficients α. These coefficients are considered as particular for each operator and for each variable. The constant α_0, which is also particular for each operator, may be added in some cases to simulate non uniform boundary conditions like $u(0, y) = \alpha_0^{D_x u} \neq 0$.

We distinguish α for different variables and different operators allowing different controls of derivatives because of the different nature of these variables and different boundary conditions prescribed for them. It is obvious,

for example, that the approximation of the derivative of η in the gradient may differ from the approximation of the derivative of u in the divergence. Although both operators represent a derivative, boundary conditions for u and η are different, these derivatives are defined at different points, at different distance from the boundary. Consequently, it is reasonable to let them be controlled separately and to assume that their optimal approximation may be different with distinct coefficients $\alpha^{D_x u}$ and $\alpha^{D_x \eta}$.

Time stepping of this model is performed by the leap-frog scheme for all hyperbolic terms and Euler scheme for the dissipative terms. The first time step is splitted into two Runge-Kutta stages in order to ensure the second order approximation.

2.2 Tangent and adjoint models

The approximation of the derivative introduced by (3) and (4) depends on control variables α. Operators are allowed to change their properties near boundaries in order to find the best fit with requirements of the model and data. To assign variables α we shall perform data assimilation procedure and find their optimal values. Variational data assimilation is usually performed by minimization of the specially introduced cost function. The minimization is achieved using the gradient of the cost function that is usually determined by the run of the adjoint to the tangent linear model.

Developing the tangent and adjoint models in this case, we follow the procedure presented in [20]. The tangent model is equal to the Gateaux derivative of the original model (2) with respect to the control parameters. To calculate this derivative we suppose that the control variables α can have small variations $\delta\alpha$ and we determine the linear model that governs the evolution of the perturbations δu, δv, $\delta\eta$ of the solution u, v, η of the model (2). Skipping the detailed development of the tangent model (see [20]), we write the model as

$$\frac{\partial}{\partial t}\begin{pmatrix}\delta u\\\delta v\\\delta \eta\\\delta \alpha\end{pmatrix}=\begin{pmatrix}\begin{smallmatrix}S_xS_yv\cdot S_yD_y\\+D_x(S_xu\cdot S_x)-\sigma+\mu\Delta^h\end{smallmatrix} & -(f_0+\beta y-S_yD_yu)\cdot S_xS_y & -gD_x & R_u\\ (f_0+\beta y+S_xD_xv)\cdot S_yS_x & \begin{smallmatrix}S_yS_xu\cdot S_xD_x\\+D_y(S_yv\cdot S_y)-\sigma+\mu\Delta^h\end{smallmatrix} & -gD_y & R_v\\ D_x(S_x\eta\cdot) & D_y(S_y\eta\cdot) & 0 & R_\eta\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\delta u\\\delta v\\\delta \eta\\\delta \alpha\end{pmatrix}$$

(5)

where

$$\begin{aligned}R_u\delta\alpha &= -(f_0+\beta y-S_yD_yu)\cdot(\delta S_x(S_yv)+S_x(\delta S_yv))+S_xS_yv\cdot(\delta S_y(D_yu)+S_y(\delta D_yu))+\\ &+ \delta D_x((S_xu)^2/2+g\eta)+D_x(S_xu\cdot\delta S_xu)\end{aligned}$$

7

$$\begin{aligned}
R_v\delta\alpha &= (f_0 + \beta y + S_x D_x v) \cdot (\delta S_y(S_x u) + S_y(\delta S_x u)) + S_y S_x u \cdot (\delta S_x(D_x v) + S_x(\delta D_x v)) + \\
&\quad + \delta D_y((S_y v)^2/2 + g\eta) + D_y(S_y v \cdot \delta S_y v)) \\
R_\eta\delta\alpha &= \delta D_x(u \cdot S_x\eta) + D_x(u \cdot \delta S_x\eta) + \delta D_y(v \cdot S_y\eta) + D_y(v \cdot \delta S_y\eta)
\end{aligned} \tag{6}$$

Operators δS and δD denote $S(\alpha + \delta\alpha) - S(\alpha)$ and $D(\alpha + \delta\alpha) - D(\alpha)$ respectively. These operators are of implicit structure. Their argument (that is, in fact, $\delta\alpha$) is contained in the matrix itself. This representation is not convenient in the development of the adjoint model, that's why we would better rewrite them in a more classical form: a constant operator (which does not depend on $\delta\alpha$) multiplied by a variable vector (which is $\delta\alpha$). So, each product like $\delta D_x u$, $\delta S_x\eta$ etc. is replaced by another product: $\delta D_x u = \widehat{[u]}_{D_x}\vec{\delta\alpha}$, $\delta S_x\eta = \widehat{[\eta]}_{S_x}\vec{\delta\alpha}$ etc. where operators $\widehat{[u]}$, $\widehat{[\eta]}$, ... are constructed from the solution u, η of the original equation and the vector $\vec{\delta\alpha}$ is extracted from matrices δD or δS as it is described in [19]. We shall further use hats to denote matrices that have been constructed from vectors. These matrices are also block-matrices. All elements of their blocks are equal to zero except one line in the beginning and one line at the end of each block. These lines are composed of two values of corresponding vector, and, namely, values of approximated function in two nodes near the boundary.

Using these notations, operators R (6) are rewritten as

$$\begin{aligned}
R_u &= -(f_0 + \beta y - S_y D_y u) \cdot ([\widehat{S_y v}]_{S_x} + S_x\widehat{[v]}_{S_y}) + S_x S_y v \cdot ([\widehat{D_y u}]_{S_y} + S_y\widehat{[u]}_{D_y}) + \\
&\quad + [\widehat{(S_x u)^2}/2 + g\eta]_{D_x} + D_x(S_x u \cdot \widehat{[u]}_{S_x}) \\
R_v &= (f_0 + \beta y + S_x D_x v) \cdot ([\widehat{S_x u}]_{S_y} + S_y\widehat{[u]}_{S_x}) + S_y S_x u \cdot ([\widehat{D_x v}]_{S_x} + S_x\widehat{[v]}_{D_x}) + \\
&\quad + [\widehat{(S_y v)^2}/2 + g\eta]_{D_y} + D_y(S_y v \cdot \widehat{[v]}_{S_y}) \\
R_\eta &= [u \cdot \widehat{S_x\eta}]_{D_x} + D_x(u \cdot \widehat{[\eta]}_{S_x}) + [v \cdot \widehat{S_y\eta}]_{D_y} + D_y(v \cdot \widehat{[\eta]}_{S_y})
\end{aligned} \tag{7}$$

The tangent model is subjected to the same zero boundary conditions for δu and δv as for u and v. No boundary conditions is prescribed for $\delta\eta$ as well as for η. At initial time $\delta u, \delta v$ and $\delta\eta$ are equal to $\delta u_0, \delta v_0$ and $\delta\eta_0$ respectively. These variables are now classical in controlling initial conditions of the model. They will be used for the same purpose and for the joint control of a boundary scheme and initial conditions of the model as well.

We should note, the matrix of the tangent linear model (5) is composed by two parts: the 3×3 block composed of operators acting in the space of the model's variables and the fourth column composed of operators R (7).

The 3×3 block is responsible for the evolution of a small perturbation by the model's dynamics and is similar for any data assimilation. The column determines the way how this perturbation is introduced into the model and is specific to the particular variable under control. This column is absent when the goal is to identify the initial conditions of the model because the uncertainty in initial conditions is introduced only once, at the beginning of the model integration. But, when the uncertainty is presented in an internal model parameter, like in this case, the perturbation is introduced at each time step of the model.

In order to develop the adjoint model, we need to introduce the scalar product in the space defined by tangent model. Each element in this space is composed of discretized functions u, v and η and also the whole set of the control coefficients α. A vector in this space has four components $\phi = (u, v, \eta, \alpha)$.

Following [20], we consider a weighted Euclidian scalar product in this space

$$
\ll \phi, \phi^* \gg \;=\; \ll \begin{pmatrix} u \\ v \\ \eta \\ \alpha \end{pmatrix}, \begin{pmatrix} u^* \\ v^* \\ \eta^* \\ \alpha^* \end{pmatrix} \gg =
$$

$$
= \; w_u^2 \sum_{i,j} u_{i,j} u_{i,j}^* + w_v^2 \sum_{i,j} v_{i,j} v_{i,j}^* + w_\eta^2 \sum_{i,j} \eta_{i,j} \eta_{i,j}^* + \sum_{k,operator} \alpha_k^{operator} (\alpha_k^{operator})^*
$$

(8)

The sums in the scalar product is performed over all nodes i, j of the grid of all model's variables u, v and η. The sum of control coefficients α is performed over all $k : 0 \leq k \leq 2$, (4) and over all operators "*operator*" controlled by these coefficients. Weights $w_u = w_v = \dfrac{1}{\sqrt{gH_0}}$ and $w_\eta = \dfrac{1}{H_0}$ are introduced to bring all variables to dimensionless form. Applying the development described in [20] we get the formulation of the adjoint model:

$$
-\frac{\partial}{\partial t} \begin{pmatrix} u^* \\ v^* \\ \eta^* \\ \alpha^* \end{pmatrix} = \begin{pmatrix} \begin{array}{c} D_y^* S_y^* (S_x S_y v \cdot) + \\ + S_x^* (S_x u \cdot D_x^*) - \sigma + \mu \Delta^h \\ -S_y^* S_x^* (f_0 + \beta y - S_y D_y u) \cdot \\ g D_x^* \\ R_u^* \end{array} & \begin{array}{c} S_x^* S_y^* (f_0 + \beta y + S_x D_x v) \cdot \\ D_x^* S_x^* (S_y S_x u \cdot) + \\ + S_y^* (S_y v \cdot D_y^*) - \sigma + \mu \Delta^h \\ g D_y^* \\ R_v^* \end{array} & \begin{array}{c} S_x \eta \cdot D_x^* \\ S_y \eta \cdot D_y^* \\ 0 \\ R_\eta^* \end{array} & \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array} \end{pmatrix} \begin{pmatrix} u^* \\ v^* \\ \eta^* \\ \alpha^* \end{pmatrix}
$$

(9)

where

$$
R_u^* \;=\; -([\widehat{S_y v}]_{S_x}^* + [\widehat{v}]_{S_y}^* S_x^*)((f_0 + \beta y - S_y D_y u) \cdot) + ([\widehat{D_y u}]_{S_y}^* + [\widehat{u}]_{D_y}^* S_y^*)(S_x S_y v \cdot) +
$$

$$+ \quad [(S_x u)\widehat{{}^2/2} + g\eta]^*_{D_x} + \widehat{[u]}^*_{S_x}(S_x u \cdot D_x^*)$$

$$R_v^* = ([\widehat{S_x u}]^*_{S_y} + \widehat{[u]}^*_{S_x} S_y^*)((f_0 + \beta y - S_x D_x v)\cdot) + ([\widehat{D_x v}]^*_{S_x} + \widehat{[v]}^*_{D_x} S_x^*)(S_y S_x u\cdot) +$$

$$+ \quad [(S_y v)\widehat{{}^2/2} + g\eta]^*_{D_y} + \widehat{[v]}^*_{S_y}(S_y v \cdot D_y^*) \tag{10}$$

$$R_\eta^* = [u \cdot \widehat{S_x \eta}]^*_{D_x} + \widehat{[\eta]}^*_{S_x}(u \cdot D_x^*) + [v \cdot \widehat{S_y \eta}]^*_{D_y} + \widehat{[\eta]}^*_{S_y}(v \cdot D_y^*)$$

2.3   Cost function

One of the principal purposes of variational data assimilation consists in the variation of control parameters in order to bring the model's solution closer to the observational data. This implies the necessity to measure the distance between the trajectory of the model and data. Introducing the cost function, we define this measure. Generally speaking, the cost function is represented by some norm of the difference between model's solutions and observations, eventually accompanied by some regularization term.

To characterize the difference between the model's solution and the observational data we use the norm based on the scalar product (8) with the fourth component α equal to zero:

$$\xi^2 = w_u^2 \sum_{i,j}(u_{i,j} - u_{i,j}^{obs})^2 + w_v^2 \sum_{i,j}(v_{i,j} - v_{i,j}^{obs})^2 + w_\eta^2 \sum_{i,j}(\eta_{i,j} - \eta_{i,j}^{obs})^2. \tag{11}$$

Expressing ξ in terms of the scalar product, we emphasize its dependence on time and on control coefficients α:

$$\xi^2 = \xi^2(\alpha, t) = \ll \phi(\alpha, t) - \phi^{obs}(t), \phi(\alpha, t) - \phi^{obs}(t) \gg =$$

$$= \ll \begin{pmatrix} u(\alpha, t) - u^{obs}(t) \\ v(\alpha, t) - v^{obs}(t) \\ \eta(\alpha, t) - \eta^{obs}(t) \\ 0 \end{pmatrix}, \begin{pmatrix} u(\alpha, t) - u^{obs}(t) \\ v(\alpha, t) - v^{obs}(t) \\ \eta(\alpha, t) - \eta^{obs}(t) \\ 0 \end{pmatrix} \gg \tag{12}$$

Taking into account the results obtained in [20], we define the cost function as

$$\mathcal{I}(\alpha) = \int_0^T t\xi^2(\alpha, t)dt \tag{13}$$

that gives bigger importance to the difference ξ^2 at the end of assimilation interval.

10

It should be noted here, that this cost function can only be used in the case of assimilation of a perfect artificially generated data. When we assimilate some kind of real data that contains errors of measurements and is defined on a different grid, we should add some regularization term to the cost function (like the distance from the initial guess) and use some more appropriate norm instead of the Euclidean one (see, for example [27] for details).

To calculate the gradient of the cost function, we calculate first its variation:

$$
\begin{aligned}
\delta\mathcal{I} &= \mathcal{I}(\alpha + \delta\alpha) - \mathcal{I}(\alpha) = \int_0^T t(\xi^2(\alpha + \delta\alpha, t) - \xi^2(\alpha, t))dt = \\
&= \int_0^T t\Big(\ll\phi(\alpha + \delta\alpha, t) - \phi^{obs}(t), \phi(\alpha + \delta\alpha, t) - \phi^{obs}(t)\gg - \\
&- \ll\phi(\alpha, t) - \phi^{obs}(t), \phi(\alpha, t) - \phi^{obs}(t)\gg \Big)dt \sim 2\int_0^T t \ll\delta\phi(t), \phi(\alpha, t) - \phi^{obs}(t)\gg dt
\end{aligned}
$$

As it has been shown in [20], the scalar product $\ll\delta\phi(t), \phi(\alpha, t) - \phi^{obs}(t)\gg$ is equal to $\ll\delta\phi(0), \mathcal{A}^*(t, 0)(\phi(\alpha, t) - \phi^{obs}(t))\gg$ where $\mathcal{A}^*(t, 0)$ denotes the adjoint model (9) integrated from the state $\phi(\alpha, t) - \phi^{obs}(t)$ back in time from t to 0. So, the result of the adjoint model run, being scalarly multiplied by $\delta\phi(0)$ provides the variation of the cost function

$$
\delta\mathcal{I} = 2 \ll\delta\phi(0), \int_0^T t\mathcal{A}^*(t, 0)(\phi(\alpha, t) - \phi^{obs}(t))dt\gg \tag{15}
$$

As it has been mentioned above, vector $\delta\phi(0)$ is composed of 4 components: δu_0, δv_0, $\delta\eta_0$ and $\delta\alpha$. If we want to control the boundary scheme only, we put zero to the first three components of $\delta\phi(0)$. Only the fourth component of the variation of the cost function (and its gradient) is different from zero in this case and only this component is used in the control. On the other hand, if our purpose is to control initial state of the model, then vanishing $\delta\alpha$ must be imposed and the first three components of the gradient must be used. And, for the joint control of boundary and initial conditions of the model we use all four components of the gradient.

11

Thus, the gradient of the cost function writes

$$\nabla \mathcal{I} = 2 \int_0^T t \mathcal{A}^*(t, 0)(\phi(\alpha, t) - \phi^{obs}(t))dt. \tag{16}$$

This gradient is used in the minimization procedure that is implemented in order to find the minimum of the cost function:

$$\mathcal{I}(\bar{\alpha}) = \min_{\alpha} \mathcal{I}(u_0, v_0, \eta_0, \alpha) \tag{17}$$

if we control the discretization of operators near the boundary using only the fourth component of the gradient (16);

$$\mathcal{I}(\bar{u}_0, \bar{v}_0, \bar{\eta}_0) = \min_{u_0, v_0, \eta_0} \mathcal{I}(u_0, v_0, \eta_0, \alpha) \tag{18}$$

if we control only the initial state of the model using three components of the gradient;

$$\mathcal{I}(\bar{u}_0, \bar{v}_0, \bar{\eta}_0, \bar{\alpha}) = \min_{u_0, v_0, \eta_0, \alpha} \mathcal{I}(u_0, v_0, \eta_0, \alpha) \tag{19}$$

if we perform the joint control of both the initial state and the boundary discretization of the model using all four components of the gradient.

Coefficients $\bar{\alpha}$ are considered as coefficients achieving an optimal discretization of the model's operators in the boundary regions. We use the minimization procedure developed by Jean Charles Gilbert and Claude Lemarechal, INRIA [28]. The procedure uses the limited memory quasi-Newton method.

3   Model in a square box.

We start from the data assimilation in frames of the very well studied "academic" configuration. Several experiments have been performed with the model in a square box of side length $L = 2000$ km driven by a steady, zonal wind forcing with a classical sinusoidal profile

$$\tau_x = \tau_0 \cos \frac{2\pi(y - L/2)}{L}$$

that leads to the formation of a double gyre circulation [29]. The attractor of the model and the bifurcation diagram in a similar configuration has been described in [30]. Following their results, we intentionally chose the model's

parameters in order to ensure chaotic behavior. The maximal wind tension on the surface is taken to be $\tau_0 = 0.5\frac{dyne}{cm^2}$. The coefficient of Eckman dissipation and the lateral friction coefficient are chosen as $\sigma = 5 \times 10^{-8}s^{-1}$ and $\mu = 200\frac{m^2}{s}$ respectively.

As it has been already noted, the Coriolis parameter is a linear function in y with $f_0 = 7 \times 10^{-5}s^{-1}$ and $\beta = 2 \times 10^{-11}(ms)^{-1}$. The reduced gravity and the depth are respectively equal to $g = 0.02\frac{m}{s^2}$, $H_0 = 1000m$.

All operators in the model are approximated with the second order accuracy both in the interior of the domain and near its boundary. That means the expression (4), that is used to interpolate functions and to calculate their derivatives near boundary, is written with $\alpha_1^D = -1/h$, $\alpha_2^D = 1/h$ for all derivative operators and $\alpha_1^S = \alpha_2^S = 1/2$ for all interpolations. That gives, for example, the value of the derivative of u at the point $i = 1/2$ as $(D_x u)_{1/2,j-1/2} = \frac{u_{1,j-1/2} - u_{0,j-1/2}}{h}$.

The solution of the model in this configuration possesses a boundary layer near the Western boundary. This is a well known Munk layer [31] that is characterized by the local equilibrium between the β-effect and the lateral dissipation. It's width is expressed by the formula $\frac{2\pi}{\sqrt{3}}(\mu/\beta)^{1/3} = 52.3$ km for the given parameters.

As it has been discussed in [32], the model must resolve this layer with at least one grid point (optimally, more than one grid point) in order to maintain numerical stability. The work of [33] emphasized the importance of having at least two grid points in the Munk layer in order to minimize the level of spurious oscillations visible in the velocity fields as well in the field of the sea surface elevation.

The resolution of the model in this section is intentionally chosen to be too coarse to resolve the Munk layer. The model's grid is composed of 30 nodes in each direction, that means the grid-step is equal to 67 km, that is more than the Munk layer's width. As it can be seen in fig.1, there is only one grid node in the layer for variables v and η and no nodes at all for u variable.

Artificial "observational" data are generated by the same model with all the same parameters but with 9 times finer resolution (7.6 km grid step). The fine resolution model, having 7 nodes in the Munk layer, resolves explicitly the layer and must have no spurious oscillations. All nodes of the coarse grid belong to the fine grid, consequently, we do not need to interpolate
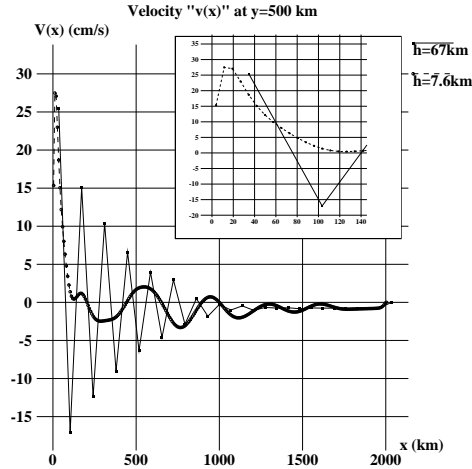
13

Figure 2: Profiles of the velocity $v(x, y)$ at $y = 500$ km obtained on the coarse grid with $h = 67 km$ (solid line), and fine grid with $h = 7.6$ km (dashed line)

"observational" data to the coarse grid. We just take values in nodes of the high resolution grid that correspond to nodes on the coarse grid. A profile of the meridional velocity component v is shown in fig.2 in order to illustrate effects of resolved and under-resolved Munk layer. These effects are the most visible in the field of this variable. However, similar oscillations are also present in the fields of the velocity u and the sea surface height (SSH) η. As it is shown in fig.1, the point on the boundary ($x = 0$) does not belong to the grid where v variable is defined. However, boundary conditions are prescribed at this point imposing vanishing v.

If we compare profiles of the velocity $v(x, y)$ plotted in fig.2 at $y = 500$ km (one quarter of the side length), we see that the profile obtained using the model with fine resolution is smooth, while the profile on the coarse resolution (solid line in fig.2) shows strong spurious oscillations due to the unresolved Munk layer. As expected, one point in the layer is not sufficient to suppress the numerical mode in the velocity field. We shall assimilate data of the fine resolution model in order to bring the solution of the coarse resolution's model closer to fine resolution's one.

The model on the fine grid has been spun up from the rest state during 3 years and integrated for the subsequent 3 years. From the result of this integration we have extracted values of all three variables at all grid points

14

that belong to the coarse grid (as it has been noted, the grids have been chosen so, that all grid points of the coarse grid belong to the fine grid). This set is used as artificial observations in the following experiments.

So far the model is nonlinear with intrinsicly unstable solution, there is no hope to obtain close solutions in long time model runs because any difference (even infinitesimal) between two models grows exponentially on these time scales. Consequently, we have to confine our study to the analysis of relatively stable properties of the solution. So, we perform two experiments. The first one describes a short time evolution of the model's solution simulating the forecasting properties of the model. The second one addresses the climatic averages of the model's solution which should also represent more stable structures than particular trajectories (see [34], [35] for example).

In the first experiment we analyze the data assimilation that was used to identify optimal initial conditions of the model $\bar{u}_0, \bar{v}_0, \bar{\eta}_0$, optimal discretization of its operators near the boundary $\bar{\alpha}$ or both applying minimization procedure and controlling different parameters as it is shown in (17),(18) and (19). The final point of spin-up of the high-resolution model was used as initial guess in experiments that control the initial state of the low-resolution model. Classical second order discretization of all operators near the boundary $\alpha_0 = 0$, $\alpha_1^D = -1/h$, $\alpha_2^D = 1/h$, $\alpha_1^S = \alpha_2^S = 1/2$ was assumed as initial guess in all experiments that control the discretization.

In the experiment that analyses the forecasting properties of the model, we assimilate "observational data" during 5 days and we examine the evolution of the difference between the model's solution and observations during next 15 days.

One can see in fig.3 that when we control initial conditions of the model, the control moves the initial state (dashed line) far from initial guess. The initial difference with observations becomes $\xi(0) = 0.1$. However, at the end of the assimilation window the distance from observations is reduced to $\xi(5 \text{ days}) = 0.07$. After the end of assimilation the distance from observations increases rapidly. The line approaches the upper solid line that corresponds to the solution of the model with no assimilation at all. This fact can easily be understood. The model's dynamics has not been modified by the assimilations procedure. The model remains the same, possessing the same deficiencies. Consequently, it is not surprising that beyond assimilation window the solution tends to the attractor of the coarse resolution model. The model starts to develop spurious oscillations shown in fig.2 moving away from the attractor of the high resolution model where these oscillations are
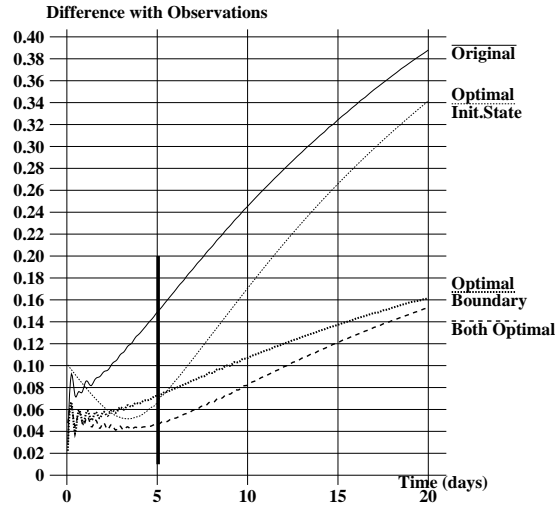
15

Figure 3: Evolution of the distance $\xi(t)$ (11) during and after assimilation.

absent.

Controlling the discretization of model's operators, we modify the model. Assimilation, in this case, brings the model dynamics towards the dynamics of the high resolution model (used to produce artificial observations). This control does not modify the initial point ($\xi(0) = 0$) but the distance $\xi(5 \text{ days})$ at the end of assimilation time (dotted line in fig.3) is almost equal to the distance obtained controlling the initial state of the model. However, beyond the window, we see that the modified model's dynamics allows the solution to remain closer to observations. The difference ξ on the twentieth day is more than two times lower. If we suppose that we assimilate data in the past 5 days in order to deliver a forecast for 15 days in the future, we see that controlling coefficients α provides two times better result.

If we control both initial state and boundary discretization, we obtain dashed line which is similar to the dotted one. Indeed, the assimilation procedure was particularly concentrated on the control of coefficients α rather than on the control of the initial state, that remains rather close to the initial guess for this state ($\xi(0) = 0.01$) . In this experiment we have lower distance ξ at the end of the assimilation window, but almost the same ξ at the end of the forecasting time.

16

Thus, we see in this experiment that if the model's dynamics suffers from low resolution and other numerical errors, better forecast is achieved by controlling the model's operators rather than initial conditions.

If we consider a long time behavior of the model, we should analyze climatic averages of the solution rather than the difference between particular trajectories. Due to intrinsic instability any trajectories diverge and the value of ξ becomes determined by the attractor size. If we assimilate data in order to control initial state of the model, we can not hope to improve its climatic averages because we do not modify the model's dynamics. No matter from which state the model starts, the same dynamics determines the same attractor and the same climatic (calculates over sufficiently long time interval) averages. In the same time, data assimilation performed with the purpose to control the discretization of operators near the boundary, does modify the dynamics. In this case, together with the short time behavior, we can hope to improve the model's climate.

To study the modification of the climatic averages, we perform another experiment starting from the same initial guess, controlling also both initial state and parameters α, but with assimilation window $T = 100$ days. Such a large window is necessary to collect an observational information about a number of physical processes that determine long-time model behavior. Optimal initial point $\bar{u}_0, \bar{v}_0, \bar{\eta}_0$ and optimal $\bar{\alpha}$ found in the experiment are used in the 1000 days model run that means 10 times the assimilation window.

Averages of the original coarse resolution model suffer a lot from the numerical effects that are present due to insufficient resolution. All fields contain spurious oscillations (see fig.2) due to unresolved Munk layer. The length of the jet-stream in the middle of the square is about 400 km (800-1000 km in the high resolution model) and the variability of the model's solution is very low. Eddy kinetic and eddy potential energies of the high resolution model show approximately 20 times bigger amplitudes. These fields are not shown in the paper.

Optimal discretization of operators in the boundary region allows us to correct these numerical errors. Average sea surface elevation $\bar{\eta}$ and eddy potential energy $\overline{\text{EPE}}$

$$\bar{\eta}(x,y) = \frac{1}{T}\int_0^T \eta(x,y,t)dt, \quad \overline{\text{EPE}} = \frac{1}{T}\int_0^T (\eta(x,y,t) - \bar{\eta}(x,y))^2 dt \quad (20)$$

obtained in the "observational" run of the high resolution model and in the run of the low resolution model with optimal boundary discretization are

17

presented in fig.4 and fig.5. One can see in that both averages and variability are very similar. That means, optimal discretization modifies the climate of the coarse resolution model bringing it closer to the reference model.
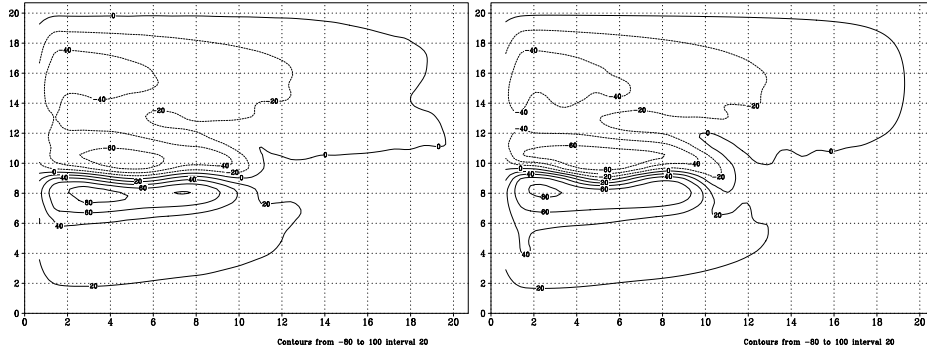

Figure 4: Thousand days average of the Sea Surface Elevation obtained with the high resolution model (left) and with the low resolution model and optimal discretization (right).
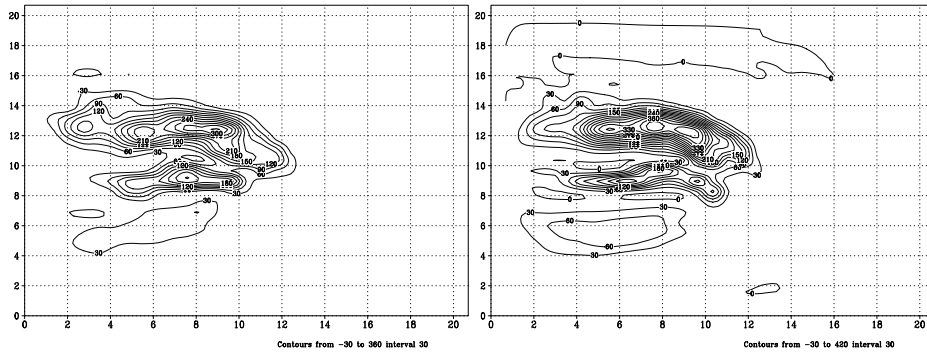

Figure 5: Thousand days average Eddy Potential Energy obtained with the high resolution model (left) and with the low resolution model and optimal discretization (right).

4   Model of the Black Sea.

In this section we use the same model, but all the parameters are defined to describe the upper layer circulation of the Black sea. Configuration of the model and observational data have been kindly provided by Gennady Korotaev from the Marine Hydrophysical Institute, National Academy of Sciences of Ukraine, Sevastopol, Ukraine. This configuration is described in [36].

The model grid counts 141×88 nodes that corresponds to the grid box of dimension 7860 m and 6950 m in x an y directions respectively. 15 minutes time step is used for integration of the model. The Coriolis parameter is equal to $f_0 = 10^{-4}s^{-1}$ and $\beta = 2 \times 10^{-11}(ms)^{-1}$. Horizontal viscosity is taken as $\mu = 200m^2s^{-1}$. Using a typical density difference, $\rho_0 - \rho_1$ of $3.2kg/m^3$, and unperturbed layer thickness of $H_0 = 150m$, the Rossby radius of deformation is estimated at about 22 km. The grid therefore resolves the mesoscale processes reasonably well.

The model has been forced by the ECMWF wind stress data, available as daily averages for the years 1988 through 1999. Dynamical sea level reconstructed in [37] was used as observational data in this section. These data have been collected in ERS-1 and TOPEX/Poseidon missions and preprocessed by the NASA Ocean Altimeter Pathfinder Project, Goddard Space Flight Center. Observational data are available from the 1st May 1992 until 1999. These data have been linearly interpolated to the model grid.

So far the sea surface elevation is the only observational variable available in this experiment, we put $w_u = w_v = 0$ in (11). Consequently, the difference between the model's solution and observations is calculated taking into account the variable η only.

The behavior of the model solution is not chaotic in this configuration. Variability in the model is generated directly by the variability of the wind stress on the surface. Consequently, we can compare particular trajectories of the model on any time interval because their evolution is stable without exponential divergence.

As it has been already noted, absence of observational data for the velocity fields brings us to modify the cost function. We have to add the background term in the cost function in order to require the velocity field to be sufficiently smooth. Otherwise, lack of information about velocity components in observational data would result in a spuriously irregular fields obtained in assimilation. To ensure necessary regularity of u and v we add the distance

from the initial guess to the cost function (13). In order to emphasize the requirement of smoothness, this distance is measured as an enstrophy of the difference between the initial guess and current state:

$$\mathcal{I}_{smooth} = \sum_{i,j} \left(\frac{\partial(v_{i,j} - v_{i,j}^0)}{\partial x} - \frac{\partial(u_{i,j} - u_{i,j}^0)}{\partial y} \right)^2 \qquad (21)$$

where u^0, v^0 denote velocity components of the initial guess of the minimization procedure. Of course, this term is only taken into account in the identification of the initial state of the model.

Moreover, using real observational data requires to add at least one another term to the cost function. One can see in the Figure 2 in [37], spatially averaged sea surface elevation of the Black sea exhibits a well distinguished seasonal cycle. That means the mass is not constant during a year, it decreases in autumn and increases in spring. Consequently, if we assimilate data during a short time (a season or less), we assimilate also the information about the mass flux specific for this season. This flux can not be corrected later by the model because the discretization of operators near the boundary (that controls the mass evolution) is obtained once for all seasons. The mass variation of the Black sea reaches 25 centimeters of the sea surface elevation. Assimilating data within one season may, consequently, result in a persisting increasing or decreasing of the seal level of order of 50 cm per year. To avoid this spurious change of the total mass, we must either take the assimilation window of at least one year, or prescribe the mass conservation to the model's scheme. One year assimilation window is computationally expensive and is not justified by the model's physics. On the other hand, prescribed mass conservation removes just the sinusoidal seasonal variation, allowing us to keep all other processes and to choose any assimilation window we need.

To correct the mass flux of the model, we add the following term to the cost function

$$\mathcal{I}_{mass} = \int_0^T \left(\sum_{i,j} (\eta_{i,j}(t) - \eta_{i,j}(0)) \right)^2 dt \qquad (22)$$

Similarly to (21), this term also ensures the regularity of the solution, but it is taken into account when the assimilation is performed for identification of the boundary parametrization. It can be noted here that other terms may be added to the cost function in order to make a numerical scheme energy and/or enstrophy conserving, but we do not use them in this paper.

20

The total cost function in this section is composed of three parts: (13), (21) and (22):

$$\mathcal{I}_{total} = \mathcal{I} + \gamma_1 \mathcal{I}_{smooth} + \gamma_2 \mathcal{I}_{mass} \qquad (23)$$

Coefficients γ are introduced to weight the information that comes from observational data (with $\mathcal{I}$) and an a priori knowledge about mass conservation and regularity of the solution.

This modification of the cost function results, of course, in additional terms in the gradient:

$$\nabla \mathcal{I}_{total} = \nabla \mathcal{I} + 2\gamma_1 \Big(D_y^* D_y (u - u_0) + D_x^* D_x (v - v_0) \Big) + 2\gamma_2 \sum_{i,j} \Big(\eta_{i,j}(t) - \eta_{i,j}(0) \Big).$$
$$(24)$$

The model is spun up from the beginning of 1988 to May 1992 using the wind tension data on the surface. The state corresponding to the 1st of May 1992 12h GMT is used as the initial guess in the data assimilation procedure controlling initial conditions of the model. The assimilation is performed following the procedure (17) with the assimilation window $T = 1$ day and the regularization parameter $\gamma_1 = 0.04$. Such a short window was chosen in order to get almost instantaneous state of the model to be used in further experiment as an initial state.

This assimilation provides sufficiently smooth velocity fields (see fig.6 left) that contain such a specific features of the Black Sea circulation as Western and Eastern gyres, Batumi gyre, filament formation along the Caucasus coast and a formation of the Sebastopol eddy (see [36] for discussion of these features). Obtained sea surface elevation field (see fig.6 right) is close to observational data that was used in the assimilation. The value of the difference ξ (11) is less than 0.1, that means the actual average difference between observed and reconstructed fields is approximately equal to 30 cm.

This initial state is used in data assimilation experiments that control the discretization of the model's operators near boundary. Of course, short assimilation window is not appropriate to identify internal model's parameters. Observational data must contain information, and especially about long time model behavior and must be capable to identify optimal discretization which is supposed to be constant in time. We have already shown on the example of the total mass evolution that short assimilation windows may lead to a wrong model behavior. However, choosing a long assimilation window require much computer time.
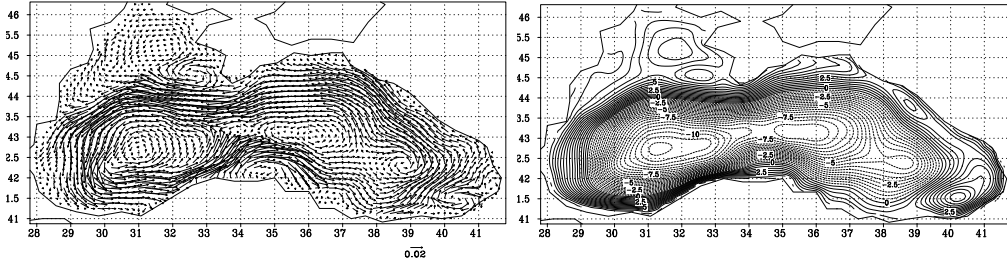
Figure 6: The model initial state corresponding to the 1st of May 1992 obtained by data assimilation: velocity (left), sea surface elevation (right).

In this paper we chose $T = 50$ days window which is longer than synoptic time scales. The minimization of the cost function has been accompanied by the mass preserving correction (22) with $\gamma_2 = 0.01$.

As well as in the previous section, we perform 3 experiments controlling initial conditions, boundary parametrization and both of them. However, due to internal stability of the model solution, we can compare particular trajectories on long time intervals. Evolution of the difference ξ between the model's solution and observations is shown in fig.7. As we can see, optimal initial point allows the solution to remain close to observations in the assimilation window but not beyond the window. As soon as the assimilation ends, the solution goes rapidly toward the solution of the original model and becomes indistinguishable from it after 200 days. On the other hand, the solution of the model with optimally discretized operators remains always closer to the observational data than the solution of the original model. That means modified model's dynamics allows the solution to approach observations.

Comparing the computational cost of the data assimilation, we must acknowledge that controlling the boundary is more expensive. First, the adjoint model's run requires approximately double computer time comparing with the adjoint model for the identification of the initial state. And second, the minimization process requires more iterations. In the experiment shown in fig.7, 5 iterations are necessary to minimize the cost function controlling the initial state of the model, but we need at least 22 iterations to reach the same stopping criterion controlling α.

As it has been discussed in [19], it is difficult to distinguish principal modifications that have been made in the numerical scheme by the data assimilation. Various kernels are present in the space of α even in a very
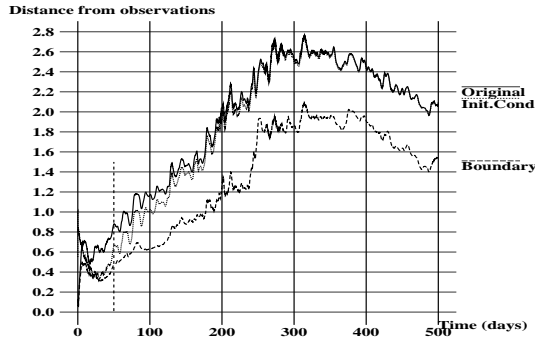
22

Figure 7: Evolution of the distance $\xi(t)$ (11) during and after assimilation.

simple model like one-dimensional wave equation. These kernels make difficult the analysis of the assimilation results because optimal coefficients are not unique. Obtained set of α represent just one point in the kernel, while the same model behavior can be obtained with any other kernel point. Consequently, in two similar assimilation experiments we can obtain very different sets of optimal coefficients but almost the same model behavior.

In the present non-linear model it is also difficult to analyze coefficients α directly. We have performed several experiments assimilating observational data during the same seasons with the same assimilation windows but in different years (1992, 1993, 1995). Assimilation results (not shown here) reveal very close values of the cost function obtained in the minimization procedure, very similar evolution of the difference ξ in the assimilation window and beyond it (like shown in fig.7), but very different sets of coefficients α. As well as in experiments with simpler models ([19],[20]), this non-uniqueness of the optimal discretization coefficients has also its roots in the kernel that exists in the space of α. Different points in the kernel correspond to different discretizations of the model's operators, that result in almost the same model's solutions.

Instead of analyzing the set of obtained coefficients α, we shall see the modification of the solution that this set generates, and namely the difference between the velocity of the model with classical boundary (the distance of this solution from observations is plotted by the solid line in fig.7) and the velocity with optimal boundary discretization (dashed line in fig.7). This difference has been averaged in time over 200 days time interval in order to reveal

23

persistent modifications of the flow produced by the optimal discretization.

This average difference of the velocity is presented in fig.8. We zoom the Southern part of the Black sea because it is in this region the difference shows the biggest values reaching 15 $\frac{cm}{s}$ while in the middle of the sea it rarely exceeds 1 $\frac{cm}{s}$.
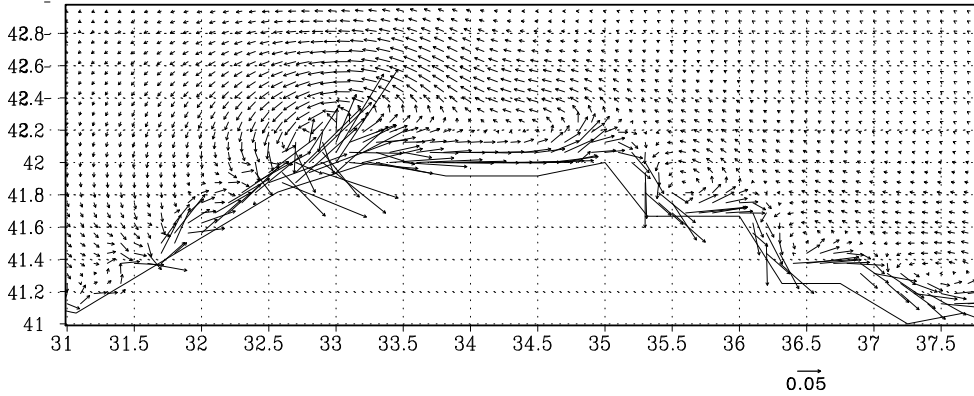

Figure 8: Difference in the velocity field of solutions with classical and optimal discretizations.

We can note several principal features of the flow that have been modified by boundary conditions. First, we can see a strong current on the boundary. The slip condition (vanishing tangential velocity) has been replaced by a permanent current along the boundary. Moreover, impermeability condition has also been modified. The flow is now allowed to leave the domain ensuring, however, the global mass balance. One can see a strong persistent vortex centered at 42.2°N, 32.8°E which southern part crosses the boundary resulting in not only tangential but normal flux also. Similar vortices with lower amplitude can also be seen in places where the boundary changes direction. Optimal discretization allows the flow to cross the boundary in places where the direction change is not smooth.

Tangential velocity component is amplified in the direct vicinity of the boundary. In these nodes we see a strong eastward flow that was forbidden by the boundary conditions in the classical formulation of the model. On the other hand, the eastward velocity is lower at nodes distanced by several grid cells from the boundary. At these nodes we see westward flow in the difference of the optimally discretized and classical models. In fact, the control of the discretization of operators near the boundary results in the same phenomenon

as we have seen above in the experiments with the square box showing the Munk boundary layer fig.2. A strong tangential current in the boundary layer is moved on the boundary allowing more optimal representation of a thin current on a coarse grid that brings the model solution towards observations.

5   Conclusion

This paper is an extension of the study presented in [20]. Both papers discuss the variational data assimilation procedure applied for identification of the optimal parametrization of derivatives and interpolation operators near the boundary, but now we work with nonlinear models assimilating both artificial and real observational data.

Contrary to linear models, even assimilating artificially generated data in twin experiments, we can not obtain a solution that is indistinguishable from assimilated data. Due to intrinsic instability, even initially close solutions of a non-linear chaotic model diverge. Hence, any difference between models in twin experiments grows with time and prevent the solutions to remain close. Thus, dealing with nonlinear models, we can control the discretization near the boundary. This allows us to correct the numerical errors due to insufficient resolution (see fig.2), to bring the model solution closer to observations (see fig.3, fig.7) and to improve statistical averages of a chaotic model's solution (see fig.4, fig.5).

Controlling the discretization of operators of the model possesses another advantage. The solution with optimal discretization remains closer to observational data after the assimilation end than the solution with optimal initial conditions. This fact can be observed both in the square box and in the Black sea. Starting from the optimal initial point, the trajectory remains close to observations in the assimilation window, but the distance with observations increases rapidly beyond the window. The solution with optimal discretization remains closer to observations even after the end of assimilation time. That means, observational data in the past are more efficiently used for forecasting if we assimilate them to control internal model's parameters than to control the initial point.

References

[1] Le Dimet FX. A general formalism of variational analysis. *Technical Report OK 73091*, CIMMS report, Normann 1982.

[2] Le Dimet FX, Talagrand O. Variational algorithm for analysis and assimilation of meteorological observations. theoretical aspects. *Tellus* 1986; **38A**:97–110.

[3] Lions JL. *Contrôle optimal de systèmes gouvernés pas des équations aux dérivées parielles.* Dunod, 1968.

[4] Marchuk G. Formulation of theory of perturbations for complicated models. *Appl. Math. Optimization* 1975; **2**:1–33.

[5] Losch M, Wunsch C. Bottom topography as a control variable in an ocean model. *J. Atmospheric and Oceanic Technology* 2003; **20**:1685–1696.

[6] Heemink AW, Mouthaana EEA, Roesta MRT, Vollebregta EAH, Robaczewskab KB, Verlaanb M. Inverse 3d shallow water flow modelling of the continental shelf. *Continental Shelf Research* 2002; **22**:465–484.

[7] Kazantsev E. Identification of optimal topography by variational data assimilation. *J. Phys. Math.* 2009; **1**:1–23.

[8] Shulman I. Local data assimilation in specification of open boundary conditions. *J. of Atmospheric and Oceanic Technology* 1997; **14**:1409–1419.

[9] Shulman I, Lewis JK, Blumberg AF, Kim BN. Optimized boundary conditions and data assimilation with application to the m2 tide in the yellow sea. *J. of Atmospheric and Oceanic Technology* 1998; **15**(4):1066–1071.

[10] Taillandier V, Echevin V, Mortier L, Devenon JL. Controlling boundary conditions with a four-dimensional variational data-assimilation method in a non-stratified open coastal model. *Ocean Dynamics* 2004; **54**(2):284–298.

[11] ten Brummelhuis PGJ, Heemink AW, Van Den Boogaard HFP. Identification of shallow sea models. *International Journal for Numerical Methods in Fluids* 1993; **17**:637–665.

[12] Zou X, Navon I, Le Dimet FX. An optimal nudging data assimilation scheme using parameter estimation. *Q. J. of Roy. Met. Soc.* 1992; **118**:1163–1186.

[13] Panchang V, O'Brien J. *On the Determination of Hydraulic Model Parameters Using the Strong Constraint Formulation Modeling Marine Systems*, vol. 1, chap. 1. CRC Press Inc, 1988; 5–18.

[14] Chertok DL, Lardner RW. Variational data assimilation for a nonlinear hydraulic model. *Applied mathematical modelling* 1996; **20**(9):675–682.

[15] Verron J, Blayo E. The no-slip condition and separation of western boundary currents. *J. Phys. Oc.* 1996; **26**(9):1938–1951.

[16] Chen H, Lin S, Wang H, Fang LC. Estimation of two-sided boundary conditions for twodimensional inverse heat conduction problems. *Int. J. Heat Mass Transfer* 2002; **45**:15–43.

[17] Gillijns S, Moor BD. Joint state and boundary condition estimation in linear data assimilation using basis function expansion. *Modelling, Identification, and Control*, Bruzzone L (ed.), 2007.

[18] Le Dimet FX, Ouberdous M. Retrieval of balanced fields: an optimal control method. *Tellus A* 1993; **45**(5):449–461.

[19] Kazantsev E. Identification of an optimal boundary approximation by variational data assimilation. *J.Comp.Phys.* 2010; **229**:256–275.

[20] Kazantsev E. Optimal boundary discretisation by variational data assimilation. *Int. J. for Numerical Methods in Fluids* 2010; n/a. doi: 10.1002/fld.2240.

[21] Leredde Y, Lellouche JM, Devenon JL, Dekeyser I. On initial, boundary conditions and viscosity coefficient control for burgers' equation. *International Journal for Numerical Methods in Fluids* 1998; **28**(1):113–128.

[22] Lellouche J, Devenon J, Dekeyser I. Boundary control on burgers equation. a numerical approach. *Comput. Math. Appl.* 1994; **28**:33–44.

[23] Gill A. *Atmosphere-Ocean Dynamics*. Academic Press, 1982.

[24] Pedlosky J. *Geophysical Fluid Dynamics*. 2d ed., Springer-Verlag, 1987.

[25] Arakawa A, Lamb V. *Computational Design of the Basic Dynamical Processes of the UCLA General Circulation Model*, vol. 17, chap. Methods in Computational Physics. Academic Press, 1977; 174–267.

[26] Arakawa A, Lamb V. A potential enstrophy and energy conserving scheme for the shallow water equations. *Mon. Wea. Rev.* 1981; **109**:18–36.

[27] Ide K, Courtier P, Ghil M, Lorenc A. Unified notation for data assimilation: Operational, sequential and variational. *J. of the Met. Soc. of Japan* 1997; **75**(1B):181–189.

[28] Gilbert J, Lemarechal C. Some numerical experiments with variable storage quasi-newton algorithms. *Mathematical programming* 1989; **45**:407–435.

[29] Le Provost C, Verron J. Wind-driven ocean circulation transition to barotropic instability. *Dyn.Atmos.Oceans* 1987; **11**:175–201.

[30] Simmonet E, Ghil M, Ide K, Temam R, Wang S. Low-frequency variability in shallow-water models of the wind-driven ocean circulation. *J. Phys. Oc.* 2003; **33**:729–752.

[31] Munk WH. On the wind-driven ocean circulation. *Journal of Meteorology* 1950; **7**:3–29.

[32] Bryan K, Manabe S, Pacanowski R. A global ocean-atmosphere climate model. part ii. the oceanic circulation. *J. Phys. Oc.* 1975; **5**(1):30–46.

[33] Griffies SM, Pacanowski RC, Hallberg RW. Spurious diapycnal mixing associated with advection in a z-coordinate ocean model. *Mon. Wea. Rev.* 2000; **128**(3):538–564.

[34] Kazantsev E. Sensitivity of the attractor of the barotropic ocean model to external influences: Approach by unstable periodic orbits. *Nonlinear processes in Geophysics* 2001; **8**:281–300.

[35] Dymnikov VP. Modelling of the climatic system response to small external forcing. *Russ. J. Numer. Anal. Math. Modelling* 2000; **15**(1):15–28.

[36] Korotaev G, Oguz T, Nikiforov A, Koblinsky C. Seasonal, interannual, and mesoscale variability of the black sea upper layer circulation derived from altimeter data. *JGR* 2003; **108**:3122.

[37] Korotaev GK, Saenko OA, Koblinsky CJ. Satellite altimetry observations of the black sea level. *JGR* 2001; **106**:917–934.