E. Kazantsev

INRIA, MOISE team, Laboratoire Jean Kuntzmann. Grenoble

DATA ASSIMILATION AND SENSITIVITY
OF THE BLACK SEA MODEL TO PARAMETERS

An adjoint based technique is applied to a Shallow Water Model in order to estimate influence of the model's parameters on the solution. Among parameters the bottom topography, initial conditions, boundary conditions on rigid boundaries, viscosity coefficients and the amplitude of the wind stress tension are considered. Their influence is analyzed from different points of view.

Two configurations have been analyzed: an academic case of the model in a square box and a more realistic case simulating Black Sea currents. It is shown in both experiments that the boundary conditions near a rigid boundary influence the most the solution. This fact points out the necessity to identify optimal boundary approximation during a model development.

KEYWORDS: *Variational Data Assimilation, Sensitivity to parameters, Boundary conditions, Shallow water model.*

1. Introduction

Thirty years ago model and data were considered as independent one on another. Observational data were interpolatedon the model grid in order to provide the model with the initial conditions, forcings and all the other necessary parameters. However, since the pioneering work [1] of Edward Lorenz, we know that a geophysical fluid is extremely sensitive to initial conditions. A perturbation of initial state may grow exponentially in time limiting the validity of the forecast. This discovery leads to understanding that observational data can not be considered as independent of the model. We must perform a joint analysis of the model and data in order to choose the optimal initial point for the model.

This become possible by using variational data assimilation technique, first proposed in [2, 3], which is based on the optimal control methods [4] and perturbations theory [5]. This technique allows us to retrieve an optimal data for a given model from heterogeneous observational fields ensuring a better forecast.

However, even now, all other forcings and parameters of the model are obtained from data by more or less sophisticated interpolation and they can not be considered as optimal for a given model. In the same time, we may suppose, that their influence on the models solution is as strong as the influence of initial state. In this case, we should also analyze the possibility and utility to apply the data assimilation techniques to identify optimal values for all these parameters in order to improve the forecast.

The purpose of this paper is to analyze the sensitivity of a Shallow-Water model and, in particular, compare the influence of initial conditions with the influence of other parameters. Among these parameters, we consider the boundary conditions on the rigid boundaries, bottom topography, empirical coefficients like reduced gravity, forcing amplitude and dissipation.

2. Sensitivity and control of the boundary conditions

We shall focus our attention on the boundary conditions because (as we shall see later) they represent the most unusual control variable.

However, as it has been noted in [6], particular attention must be paid to the discretization process which must respect several rules because 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 them-self, it may be more useful to identify optimal discretization of differential operators in points adjacent to boundaries because this is more general case. Indeed, 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.

Boundary conditions are usually introduced into the model by a particular discretization of operators near the boundary. For example, taking into account the condition $u_0 = 0$ we can calculate the derivative at the point $x = h/2$ as

$$\left.\frac{\partial u}{\partial x}\right|_{1/2} = \frac{u_1}{h}.$$

In this paper, we shall write the approximation of the derivative in a general form

$$\left.\frac{\partial u}{\partial x}\right|_{1/2} = \frac{\alpha_0 + \alpha_1 u_1}{h}.$$

Coefficients α_0 and α_1 will be used as controls. That means we shall let them vary in the data assimilation procedure in order to find an optimal pair that realizes the minimum of the cost function.

2.1. Example: one-dimensional wave equation

In order to understand what happens when the data are assimilated to control the boundary conditions, we propose to take a look on a scholar example: one-dimensional wave equation written for $u = u(x,t)$ and $p = p(x,t)$ in the following way:

$$\frac{\partial u}{\partial t} - \frac{\partial p}{\partial x} = 0$$

$$\frac{\partial p}{\partial t} - \frac{\partial u}{\partial x} = 0$$
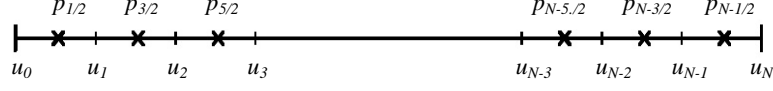
(1)

This equation is defined on the interval $0 < x < 1$ with boundary conditions prescribed for u only:

$$u(0,t) = u(1,t) = 0 \tag{2}$$

Initial conditions are prescribed for both u and p

$$u(x,0) = \overline{u} \; ; \; p(x,0) = \overline{p} \tag{3}$$

The equation is discretized on a regular grid that is somewhat similar to Arakawa C grid [7] in two dimensions:

Discrete derivatives of u and p are defined as follows

$$(\frac{\partial p}{\partial x})_i = \frac{p_{i+1/2} - p_{i-p/2}}{h}, \ (\frac{\partial u}{\partial x})_{i+1/2} = \frac{u_{i+1/2} - u_{i-p/2}}{h} \tag{4}$$

in all internal points i.e. $2 \leq i \leq N - 2$ for $(\frac{\partial p}{\partial x})_i$ and $1 \leq i \leq N - 2$ for $(\frac{\partial u}{\partial x})_{i+1/2}$.

Near the boundary, at points $1/2, 1, N - 1, N - 1/2$, we write the derivatives in a general form, like

$$(\frac{\partial p}{\partial x})_1 = \frac{\alpha_0^p p_{1/2} - \alpha_1^p p_{3/2}}{h}, \ (\frac{\partial u}{\partial x})_{1/2} = \frac{\alpha_0^u + \alpha_1^u u_1}{h} \tag{5}$$

considering α_0 and α_1 as the control coefficients. Leap-frog scheme was used for time stepping.

We introduce the simplest cost function that represents the distance between the model solution and observation at time t:

$$L(\alpha) = \int_0^T \int_0^1 (u(\alpha,x,t) - u^{obs}(x,t))^2 + (p(\alpha,x,t) - p^{obs}(x,t))^2 \, dx dt \tag{6}$$

and we calculate its gradient using the adjoint to the derivative of the solution with respect to control coefficients α^p, α^u:

$$\nabla L = 2 \int_0^T \left(\frac{\partial u(t), p(t)}{\partial \alpha} \right)^* \begin{pmatrix} u(\alpha,x,t) - u^{obs}(x,t) \\ p(\alpha,x,t) - p^{obs}(x,t) \end{pmatrix} dt \tag{7}$$

Once we prescribe the initial conditions for the equation

$$u(x,0) = \sin(k\pi x) \ p(x,0) = \cos(k\pi x),$$

we can calculate its exact solution:

$$u_{exact}(x,\text{t}) = \sqrt{2} \sin(k\pi t - \pi/4)\sin(k\pi x),$$
$$p_{exact}(x,\text{t}) = -\sqrt{2} \cos(k\pi t - \pi/4)\cos(k\pi x).$$

The exact solution is used as artificial observational data in this example. We perform the minimization of the cost function (6). The minimization procedure used here was developed by Jean Charles Gilbert and Claude Lemarechal, *INRIA* [8]. The procedure uses the limited memory quasi-Newton method.

The difference between the models solution and the exact one is shown in fig.1.
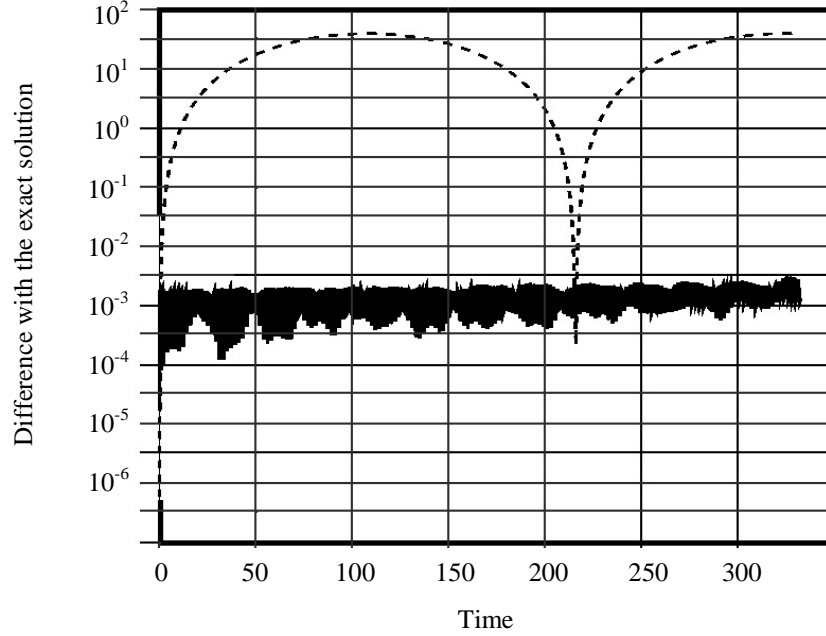
Fig. 1. Difference between the models solution and the exact one: classical
BC – dashed line (••••••••••), optimal *BC* – solid line (————).

We see that optimal discretization of derivatives near the boundary brings the solution much closer to the exact solution, but the set of optimal coefficients α does not approximate a derivative:

$$\left(\frac{\partial u}{\partial x}\right)_{1/2} = 1{,}048 \; \frac{u_1 - u_0}{h} \quad \left(\frac{\partial p}{\partial x}\right)_1 = \frac{3{,}014 p_{3/2} - 2{,}828 p_{1/2}}{h} \tag{8}$$

Neither expression for $\frac{\partial u}{\partial x}$, nor for $\frac{\partial p}{\partial x}$ has any reasonable order of approximation. The first one is of «0» order, the second is of «-1» order. Moreover, while we get always the same formula for $\frac{\partial u}{\partial x}$, approximation of the derivative of p varies in different assimilation experiments. Assimilations performed with different assimilation windows, for example, result in different coefficients for $\frac{\partial p}{\partial x}$. In fact, any combination α_0^p , α_1^p in (5) may be found as the result of assimilation under condition

$$\alpha_1^p = -1{,}104\alpha_0^p - 0{,}107 \, . \tag{9}$$

This linear relationship has been obtained experimentally performing assimilations with all assimilation windows in range from 600 to 2400 time steps (with the time step equal to 1/120 of the time unit).

164

To explain this strange result, we analyze the numerical solution of the equation. It is well known, the principal numerical error of the scheme is a wrong wave velocity. The wave speed, that must be equal to 1, is replaced by

$$\beta = \frac{h\sin(k\tau)}{2\tau\sin(kh/2)}$$

which depends on the time step τ and the grid step h. For the given parameters ($k = 3$, $h = 1/30$ and $\tau = 120$), error in the wave velocity is equal to $3,09 \times 10^{-3}$.

The data assimilation and control of the boundary derivatives can not modify numerical wave velocity. The only way for this control to get a better solution consists in modifying the length of the interval. A numerical wave with wrong velocity will propagate on the interval with wrong length. But the length of the interval is adapted by data assimilation in order to ensure the wave with numerical velocity propagates the modified interval in the same time that the exact wave propagates the exact interval. So far, the control can not correct the error in the wave velocity, it commits another error in length in order to compensate the first one as it is illustrated in fig. 2.
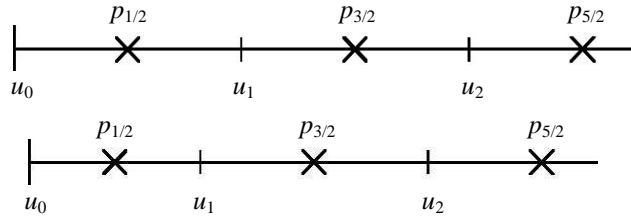

Fig. 2. Modification of the intervals length.

Non uniqueness of optimal α_1^p and α_0^p can be explained if we take into account that p has also a form of cosine of $3\pi x$. Hence, at any time $p_{1/2} = A(t)\cos(3\pi h/2)$ and $p_{3/2} = A(t)\cos(9\pi h/2)$ with some A depending on time. Their linear combination $\alpha_1^p p_{1/2} + \alpha_0^p p_{3/2}$ can vanish if

$$\alpha_1^p = -\frac{\alpha_0^p}{4\cos^2(k\pi h/2) - 3} \tag{10}$$

Consequently, all couples α_1^p, α_0^p belonging to the line that passes by the point $\alpha_0^p = -1,023$, $\alpha_1^p = 1,023$ with tangent $-\dfrac{\alpha_0^p}{4\cos^2(k\pi h/2) - 3} = -1,108$ produce the same derivative. This line coincides withing accuracy of computation with the set (9) obtained numerically. Any point on this line gives coeffcients α_p that theoretically provide the same value of the derivative and the same value of the cost function.

Of course, in this simple example we can avoid the ambiguity in the solution: it is sufficient to control only one coefficient α^p rather than two. But, in more complex problems, it may be difficult to locate and avoid the presence of kernels.

165

Consequently we can say that the data assimilation allows to place the boundary in the optimal position resulting in a solution closer to the exact one. Boundary control allows to compensate numerical errors committed in the interior of the domain, but it may be difficult to understand the physical meaning of optimal coefficients α and non-null kernels may exist leading to non unique result.

More details of this study can be found in [9]

2.2. Shallow Water Model

In this paper we consider a shallow-water model written in a conservative form:

$$\frac{\partial hu}{\partial t} = fhu - \frac{\partial}{\partial x}\left(hu^2 + gh(h-H) - \mu h \frac{\partial u}{\partial x} \right) -$$
$$- \frac{\partial}{\partial y}\left(huv - \mu h \frac{\partial u}{\partial y} \right) - \sigma hu + \tau_0 \tau_x \ ,$$

$$\frac{\partial hv}{\partial t} = -fhu - \frac{\partial}{\partial x}\left(huv - \mu h \frac{\partial u}{\partial x} \right) - \tag{11}$$
$$- \frac{\partial}{\partial y}\left(hv^2 + \mu gh(h-H) - \mu h \frac{\partial v}{\partial y} \right) - \sigma hv + \tau_0 \tau_x \ ,$$

$$\frac{\partial h}{\partial t} = -\frac{\partial hu}{\partial x} - \frac{\partial hv}{\partial y} \ ,$$

where $hu(x,y,t)$ and $hv(x,y,t)$ are two flux components that represent the product of the velocity by the ocean depth, $h(x,y,t)$, that corresponds to the distance from the sea surface to the bottom of the ocean. The sea surface elevation is represented by the difference $h(x,y,t) - H(x,y)$, where $H(x,y)$ is the bottom topography. The model is driven by the surface wind stress with components $\tau_x(x,y,t)$ and $\tau_y(x,y,t)$ normalized by τ_0 and subjected to the bottom drag that is parameterized by linear terms σhu and σhv. Horizontal eddy diffusion is represented by harmonic operators $div(\mu h \nabla u)$ and $div(\mu h \nabla u)$. Coriolis parameter is represented by the variable $f(y)$ that is equal to $f_0 + \beta y$ assuming β-plane approximation. Parameter g is the reduced gravity. The system is defined in some domain Ω with characteristic size L requiring that both hu and hv vanish on the whole boundary of Ω. No boundary conditions is prescribed for h. Initial conditions are defined for all variables: hu, hv and h.

As usual, initial conditions are considered as the control parameter of the model in this paper. We study the sensitivity of the model to its initial point and assimilate data to find its optimal value. However, in addition to initial conditions, all other parameters of the model, and namely the discretization of operators near the boundary, its bottom topography $H(x, y)$, scalar coefficients μ, σ, g and τ_0, are also considered as control variables. All of them are allowed to vary in the data assimilation procedure in order to bring them to their optimal values.

We discretize all variables of this equation on the regular Arakawa's C-grid [10] with constant grid step $\delta x = \frac{L}{N}$ in both x and y directions. Discretizing the

system (11), 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 are defined. The discretized system (11) writes

$$\frac{\partial hu}{\partial t} - fS_x S_y hu + D_x\left(\frac{(S_x hu)^2}{h} + gh(h - H(x,y)) - \mu h D_x \frac{hu}{S_x h}\right) +$$

$$+ D_y\left(\frac{(S_x hu \; S_y hu)}{S_x S_y h} - \mu(S_x S_y h)D_y \frac{hu}{S_x h}\right) = -\sigma hu + \tau_0\tau_x \; ;$$

$$\frac{\partial hu}{\partial t} + fS_x S_y hu + D_x\left(\frac{(S_x hu \; S_y hu)}{S_x S_y h} - \mu(S_x S_y h)D_y \frac{hv}{S_y h}\right) + \qquad (12)$$

$$+ D_y\left(\frac{(S_x hu)^2}{h} + gh(h - H(x,y)) - \mu h D_y \frac{hu}{S_y h}\right) = -\sigma hu + \tau_0\tau_y \; ;$$

$$\frac{\partial h}{\partial t} = -D_x hu - D_y hv \; .$$

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

$$(D_x hu)_{i-1/2,j-1/2} = \frac{hu_{i,j-1/2} - hu_{i-1,j-1/2}}{dx} \qquad \text{for } i = 2, \dots N-1,$$

$$(S_x hu)_{i-1/2,j-1/2} = \frac{hu_{i,j-1/2} - hu_{i-1,j-1/2}}{2} \qquad \text{for } i = 2, \dots N-1. \qquad (13)$$

Discretization of operators in the directly adjacent to the boundary nodes are different from (13) and represent the control variables in this study. In order to obtain their optimal values assimilating external data, we suppose nothing about derivatives and interpolations near the boundary and write them in a general form

$$(D_x hu)_{1/2,j-1/2} = \sigma_0^{D_x^{hu}} + \frac{\sigma_1^{D_x^{hu}} hu_{0,j-1/2} - \sigma_2^{D_x^{hu}} hu_{1,j-1/2}}{dx} \; ;$$

$$(S_x hu)_{1/2,j-1/2} = \sigma_0^{S_x^{hu}} + \frac{\sigma_1^{S_x^{hu}} hu_{0,j-1/2} - \sigma_2^{SD_x^{hu}} hu_{1,j-1/2}}{2} \; . \qquad (14)$$

This formula represents a linear combination of values of hu at two points adjacent to the boundary with coefficients α. The constant α_0 maybe added in some cases to simulate non-uniform boundary conditions like $hu(0,y) = \alpha_0 \neq 0$.

We distinguish α for different variables and different operators allowing different controls of derivatives because of the different nature of these variables

167

and different boundary conditions prescribed for them. It is obvious, for example, that the approximation of the derivative D_x in the first equation may differ from the approximation of D_x in the third one. Although both operators represent a derivative, boundary conditions for hu and h 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 $\alpha^{D_x^{hu}}$ and $\alpha^{D_x^h}$.

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

As well as before, the approximation of the derivative introduced by (13) and (14) depends on variables α. These variables are added to the set of control variables enumerated above. 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 all control variables 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.

To define the cost function we introduce dimensionless state vector ϕ that is composed of three variables of the model $\phi = \{w_{hu}hu, w_{hv}hv, w_h h\}^t$ weighted by coeffcients w. These weights are used to normalize values of the flux components by $w_{hu} = w_{hv} = (H_0\sqrt{gH_0})^{-1}$ and the Sea surface elevation by $w_h = (H_0)^{-1}$. The distance between the model solution and observations is defined as the Euclidean norm of the difference

$$\xi^2 = \xi^2(\phi(p,t)) = \sum_k (\phi_k - \phi_k^{obs})^2 = \tag{15}$$

$$= w_{hu}^2 \sum_{t,j} (hu_{i,j} - hu_{i,j}^{obs})^2 + w_{hu}^2 \sum_{t,j} (hv_{i,j} - hv_{i,j}^{obs})^2 + w_h^2 \sum_{t,j} (h_{i,j} - h_{i,j}^{obs})^2. \tag{16}$$

In this expression, we emphasize implicit dependence of ξ on time and on the set of the control parameters p that is composed of:

– the set of initial conditions of the model $\phi_0 = \{hu|_{t=0}, \; hv|_{t=0}, \; h|_{t=0}\}$;

– the set of the coefficients α that controls the discretizations of operators near the boundary;

– the bottom topography $H(x, y)$;

– four scalar parameters σ, μ, g, τ_0.

Taking into account the results obtained in [11], we define the cost function as

$$\mathrm{L}(p) = \int_0^T t\xi^2(\phi(p,t))dt$$

$$\tag{17}$$

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

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 [12] for details).

The n-th component of the gradient of the cost function can be calculated as the Gateaux derivative of an implicit function:

$$(\nabla \mathsf{L})_n = \frac{\partial \mathsf{L}}{\partial p_n} = \int\limits_0^T t \left(\frac{\partial \xi^2}{\partial p_n} \right) dt = \int\limits_0^T t \left(\sum_k \frac{\partial \xi^2}{\partial \phi_k} \frac{\partial \phi_k}{\partial p_n} \right) dt =$$
$$= 2\int\limits_0^T t \left(\sum_k (\phi_k - \phi_k^{obs}) \frac{\partial \phi_k}{\partial p_n} \right) dt \tag{18}$$

because the derivative $\dfrac{\partial \xi^2}{\partial \phi_k}$ can easily be calculated from (15): $\dfrac{\partial \xi^2}{\partial \phi_k} = 2(\phi_k - \phi_k^{obs})$.

The second term in (18), $\dfrac{\partial \phi_k}{\partial p_n}$, represents the matrix of the tangent linear model that relates the perturbation of the parameter p_n and the perturbation of kth component of the model state vector ϕ_k. This relationship, of course, is assumed in the linear approach, that means it is only valid for infinitesimal perturbations.

In the classical case, when initial conditions are considered as the only control variable, the derivative $\dfrac{\partial \phi(t)}{\partial p} = \dfrac{\partial \phi(t)}{\partial \phi_0}$ is the classical tangent model that describes the temporal evolution of a small error in the initial model state. The matrix is a square matrix that is widely studied in numerous sensitivity analyses. Its singular values at infinite time limit are related to well known Lyapunov exponents that determine the model behavior (chaotic or regular) and the dimension of it's attractor.

In our case, the matrix $\dfrac{\partial \phi(t)}{\partial \phi_0}$ is rectangular. It describes the evolution of an infinitesimal error in any parameter (including initial state). However, we can study it's properties in the similar way as we do with the classical tangent linear model. Its structure and composition is described in [11] for the case of using coefficients α as control parameters and in [13] for the case when the bottom topography is used to control the model's solution.

The product $\sum_k \left(\phi_k - \phi_k^{obs} \right) \dfrac{\partial \phi_k}{\partial p_n}$ in (18) represents an unusual vector-matrix

169

product. To calculate this product directly we would have to evaluate all the elements of the matrix. This would require as many tangent model runs as the size of the state vector is. So, instead of the tangent model, we shall use the adjoint one that allows us to get the result by one run of the model. Backward in time adjoint model integration that starts from $(\phi - \phi^{obs})$ provides immediately the product

$\left(\dfrac{\partial \phi}{\partial p}\right)^* (\phi - \phi^{obs})$ which is exactly equal to $(\phi - \phi^{obs})\dfrac{\partial \phi}{\partial p}$ in (18).

Using these notations, we write

$$\nabla \mathsf{L} = 2 \int_0^T t \left(\frac{\partial \phi(t)}{\partial p}\right)^* (\phi(t) - \phi^{obs}(t)) \; dt \qquad (19)$$

where the expression in the integral is the result of the adjoint model run from t to 0 starting from the vector $(\phi(t) - \phi^{obs}(t))$.

Tangent and adjoint models have been automatically generated by the Tapenade software [14, 15] developed by the *TROPICS* team in *INRIA*. This software analyzes the source code of the nonlinear model and produces codes of it's derivative $\dfrac{\partial \phi}{\partial p}$ and of the adjoint $\left(\dfrac{\partial \phi}{\partial p}\right)^*$.

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

$$\mathsf{L}(p) = \min_p \mathsf{L}(p) \qquad (20)$$

Coefficients *p* are considered as coefficients achieving an optimal parameters for the model. As it has been already noted, the set of parameters p is composed of the set of initial conditions of the model ϕ_0, the set of the coefficients α that controls the discretization of operators near the boundary, the bottom topography $H(x, y)$ and four scalar parameters σ, μ, g, τ_0. We shall minimize the cost function controlling either the total set of available parameters p or any possible subset, comparing the efficiency of the minimization.

We use the minimization procedure developed by Jean Charles Gilbert and Claude Lemarechal, *INRIA* [8]. The procedure uses the limited memory quasi-Newton method.

In addition to the data assimilation, we perform also the sensitivity study of the model solution to parameters enumerated above. We are looking for a perturbation in the model's parameters δp that, for a given small norm, maximizes the norm of the perturbation of the solution at time *t*.

$$\lambda(t) = \max_p \frac{\|\delta \phi(t)\|}{\|\delta p\|} \qquad (21)$$

170

We can note that we already have all the necessary software to estimate $\lambda(t)$. Tangent linear model $\left(\dfrac{\partial \phi(t)}{\partial p}\right)$ allows us to calculate $\delta\phi(t) = \left(\dfrac{\partial \phi(t)}{\partial p}\right)\delta p$. Using the scalar product that corresponds to the norm in the definition of the distance ξ (16), we can write $\gg$

$$\lambda(t) = \max \frac{\ll \delta\phi(t), \delta\phi(t) \gg}{\ll \delta p, \delta p \gg} = \max \frac{\ll \left(\dfrac{\partial \phi(t)}{\partial p}\right)\delta p, \left(\dfrac{\partial \phi(t)}{\partial p}\right)\delta p \gg}{\ll \delta p, \delta p \gg} = $$

$$= \max \frac{\ll \left(\dfrac{\partial \phi(t)}{\partial p}\right)^{*}, \left(\dfrac{\partial \phi(t)}{\partial p}\right)\delta p, \delta p \gg}{\ll \delta p, \delta p \gg} \tag{22}$$

This expression is a well known Rayleigh-Ritz ratio which is equal to the largest eigenvalue of the problem

$$\left(\frac{\partial \phi(t)}{\partial p}\right)^{*}\left(\frac{\partial \phi(t)}{\partial p}\right)\vartheta = \lambda(t)\vartheta \tag{23}$$

So far, we need just the maximal eigenvalue and the matrix of the problem is a self-adjoint positive definite matrix, we can solve the problem (23) by the power method performing successive iterations

$$\vartheta_{n+1} = \frac{\left(\dfrac{\partial \phi(t)}{\partial p}\right)^{*}\left(\dfrac{\partial \phi(t)}{\partial p}\right)\vartheta_{n}}{\left\|\left(\dfrac{\partial \phi(t)}{\partial p}\right)^{*}\left(\dfrac{\partial \phi(t)}{\partial p}\right)\vartheta_{n}\right\|}, \quad \vartheta_{0} = \text{random vector.}$$

In the limit, the denominator of the right-hand-side tends to the largest eigenvalue and ϑ_{n} – to the corresponding eigenvector of the matrix. The principal advantage of this method consists in the fact that we do not need to calculate the matrix itself, we just need a matrix-vector product. So far, we have both codes for the tangentand adjoint models, we can successively run these models and get the left-hand side of (23).

We should note here that when the initial conditions of the model are used as the control parameters (i.e. $\delta p = \delta\phi(0)$), the sensitivity characteristics $\lambda(t)$ are all close to one when $t \to 0$. It is evident because the perturbation has no time to betransformed by the model's dynamics and we get $\delta\phi(t)|_{t\to 0} = \delta\phi(0) = \delta p$.

When any other model parameter is used as the control and the error growing time is small, all $\lambda(t)$ are vanishing. This is also clear: the model's dynamics has no time to transmit the perturbations from the parameters to the solution. The perturbation of the solution remains, consequently, close to zero as well as the value of $\lambda(t)|_{t\to 0} = 0$.

In order to make the behavior of the sensitivity characteristics uniform with different parameters, we shall use $\lambda(t) \to 1$ every time when the initial model's state is considered as the control parameter.

3. Configurations

3.1. 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 [16]. The attractor of the model and the bifurcation diagram in a similar configuration has been described in [17]. Following their results, we intentionally chose the model's parameters to ensure chaotic behavior. The maximal wind tension on the surface is taken to be $\tau_0 = 0.5 \; dyne/cm^2$. The coefficient of Eckman dissipation and the lateral friction coefficient are chosen as $\sigma = 5 \times 10^{-8}$ (ms)$^{-1}$ and $\mu = 200$ 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 $g = 0.02$ m/s^2, H$_0 = 1000$ m.

The resolution of the model in this section is intentionally chosen to be too coarse to resolve the Munk layer [18] that is characterized by the local equilibrium between the β-effect and the lateral dissipation. Its characteristic width is determined by the Munk parameter $d = 2\left(\dfrac{\mu}{\beta}\right)^{1/3}$ which is equal to 42 km in the present case. 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 parameter. Thus, there is only one grid node in the layer and the solution exhibits spurious oscillations near the western boundary due to unresolved boundary layer.

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 «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.

The model on the fine grid has been spun up from the rest state during 3 years. The end of spin up was used as the initial state for the further integration of the model. From the result of this integration we have extracted values of all three variables at all grid points 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 instable solution, there is no hope to obtain close solutions in long time model runs because any difference

(even infinitesimal) between two models grows exponentially in time. Consequently, we have to confine our study to the analysis of a short time evolution of the model's solution simulating the forecasting properties of the model.

As the initial guess for the initial conditions we use the state vector of the high resolution model reduced on the coarse grid. This state is also used as the initial conditions in all other assimilation experiments with other control parameters. Noted above values of the model's parameters (flat bottom topography, linear in y Coriolis parameter and scalar parameters (μ, σ, τ_0, g) are used as the initial guess in the experiments that control these parameter, otherwise we simply use these parameters in the model.

3.2. Model of the Black Sea

In this section we use the same model, but all the parameters are defined to describe the upper layers 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 [19].

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 $f_0 = 10^{-4} \text{s}^{-1}$ and $\beta = 2 \times 10^{-11} (\text{ms})^{-1}$. Horizontal viscosity is taken as $\mu = 50 \text{ m}^2\text{s}^{-1}$. Using a typical density difference between upper and underlying layers of 3,1 kg/m^3, and unperturbed layer thickness of $H_0 = 150$ m, the Rossby radius of deformation is estimated at about 22 km and the reduced gravity value $g = 0,031$ m/s^2. 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 [20] 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 $w_{hu} = w_{hv} = 0$ in (16). Consequently, the difference between the model's solution and observations is calculated taking into account the variable h only.

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 h_u and h_v we add the distance from the initial guess to the cost function (17). 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:

$$\mathsf{L}_{smooth} = \sum_{i,j} \left(\frac{\partial(hv_{i,j} - hv_{i,j}^0)}{\partial x} - \frac{\partial(hu_{i,j} - hu_{i,j}^0)}{\partial y} \right)^2 \qquad (24)$$

where hu^0, hv^0 denote flux components of the initial guess of the minimization-procedure.

Moreover, using real observational data requires to add at least one another term to the cost function. One can see in the fig. 2 in [20], 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

$$\mathsf{L}_{mass} = \int_0^T \left(\sum_{i,j} (h_{i,j}(t) - h_{i,j}(0)) \right)^2 dt . \qquad (25)$$

Similarly to (24), this term also ensures the regularity of the solution. 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.

The total cost function in this section is composed of three parts: (17), (24) and (25)

$$\mathsf{L}_{total} = \mathsf{L} + \gamma_1 \mathsf{L}_{smooth} + \gamma_2 \mathsf{L}_{maas} \qquad (26)$$

Coefficients γ are introduced to weight the information that comes from observational data (with L) 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:

$$\nabla \mathsf{L}_{total} = \nabla \mathsf{L} + 2\gamma_1 \left(D_y^* D_y (hu - hu_0) + D_x^* D(hv - hv_0) \right) + 2\gamma_2 \sum_{i,j} \left(\eta_{i,j} - \eta_{i,j}(0) \right) \quad (27)$$

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 controls the initial conditions ϕ_0 only 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.

The behavior of the model solution is not chaotic in this configuration. Variability of 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. Thus, we can hope that assimilating data in a relatively short window allows us to bring the model's solution closer to observation for a long integration period.

The minimization of the cost function has been accompanied by the mass preserving correction (25) with $\gamma_2 = 0{,}01$.

4. Sensitivity analysis

The flexibility of the model is illustrated in fig. 3. We perform the data assimilation experiment in two configurations using parameters described above as initial guess. Due to high *CPU* time of the data assimilation, we limit the number of iterations of the minimization procedure by 20. Thus, we have similar and reasonable computational cost in each experiment.
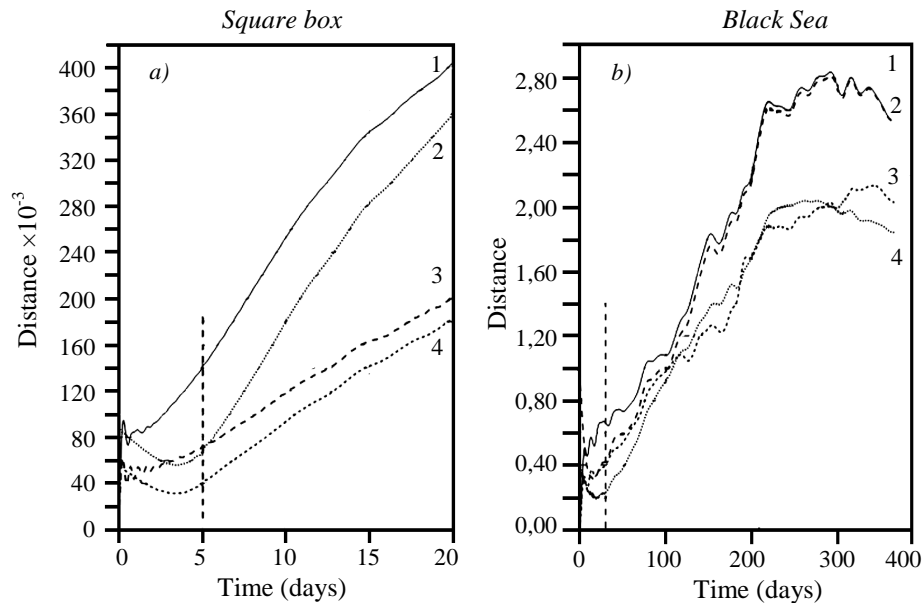

Fig. 3: Distance between the model solution and observations for the model in the square box (*a*) and the model of the upper layer of Black Sea (*b*). 1 – No Assimilation; 2 – Optimal Init Cond.; 3 – Optimal Boundary; 4 – All Optimal.

In both configurations we examine the evolution of the distance «model–observations» $\xi(t)$ during assimilation and after the end of assimilation. Assimilation window has been chosen as 5 days in the square box configuration and $T = 30$ days for the Black Sea model. The distance is examined over longer intervals: 20 days in the first case and 1 year in the second one.

Analyzing the difference between the model solution and observations shown in fig. 3, we see that in the assimilation window the model is almost equally flexible with respect to both initial and boundary conditions. Data assimilation allows us to reduce the distance between the model solution and observations at the end of the window approximately twice in both configurations. The only difference that can be seen in the assimilation window is that non-optimal initial point leads to the spurious oscillations of the solution. These oscillations occur in both configurations and show us the necessity to identify the optimal initial point.

However, the influence of parameters is significantly different beyond the window. While the solution with optimal initial point tends towards the solution obtained without any data assimilation, optimal set of boundary conditions ensures a new solution that is much closer to observational data. That means the control of boundary conditions allows us to improve a long-range forecasting quality of the model.
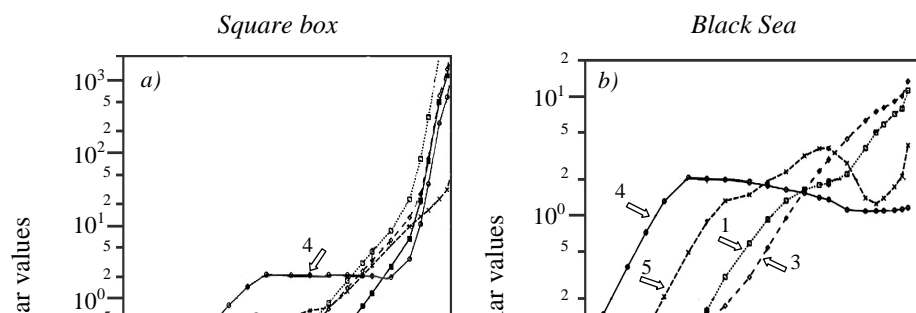
The third way of the sensitivity analysis consists in solving of the eigenvalue problem (23) and analyzing $\lambda(t)$ on different scales of error growing time from about 10 minutes (10^{-3} day) to approximately one year. As it has been already noted, $\lambda(t)-1$ is plotted in the case when initial conditions are considered as the parameter.

Analyzing the figure (see fig. 4), we can see that three time scales can be clearlydistinguished for the sensitivity characteristics of the model in both configurations. The first, short time scales, approximately from 0 up to 2-3 hours is characterized by the linear growth of $\lambda(t)$. Indeed, the model behaves as a linear model on this scales, the model's solution can be well approximated by just one step of the numerical time scheme.

The second time scale that can be distinguished in the figure (see fig. 4) corresponds to error growing times from 2-3 hours to 10 days. On these time scales we see slower growth of the sensitivity characteristics $\lambda(t)$ and, sometimes, no growth at all. These time scales are characterized by the modification of the stable-instable subspaces of the model. Instable space on short time scale is not the same as for long time evolution. Short time instabilities are usually localized in space, while long time eigenvectors of (23) possesses a global structure.

The third time scale corresponds to the error growing times more than 100 days. On these scales the model exhibits either non-linear chaotic behavior with exponential growth of all $\lambda(t)$ (as it is the case in the square box), or stable behavior when a perturbation of initial state decreases with time (as it is the case in the Black Sea model).

In order to zoom these time scales, we plot the same data in the Log-Log and Log-Linear coordinates in fig.4 on the left and on the right respectively. One can see the error growth in the square box on this time scale is purely exponential with the same exponent $\lambda(t) = A \cdot \exp(0{,}027t)$. The multiplier A is particular for each parameter, but the exponent is always the same. This confirms the remark made in [13], [11]: no matter how the perturbation was introduced into the model, it's long-time growth is determined by the model's dynamics.

Square box   *Black Sea*

Comparing the evolution of the sensitivity of the model to different parameters,we see that on small scales it is the bottom topography that the model is the most sensitive to (thin solid line in fig. 4). An error in the topography produces 13 times bigger perturbation in the model state than a similar error in the model's initial conditions (thick solid line in fig. 4). However, $\lambda(t)$ does not grow at all on medium scales due to significant changes in the eigenvectors pattern. This leads to the fact that on long scales, the sensitivity of the model to the bottom topogra-

phy is about 2 times lower than the sensitivity to initial conditions.

On the other hand, the sensitivity of the model to the discretization of operators near the boundary exhibits the opposite behavior. On short scales, corresponding λ is 2 times lower than λ obtained for perturbations of ϕ_0, but there is no stagnation of the growth on the middle scales. As a result, we see that the model is 4 times more sensitive to α than to ϕ_0 for long error growing times. Moreover, small perturbation of initial conditions decreases in the Black sea configuration, while the perturbation of α results in an increasing perturbation of the solution.

5. Modification of the boundary conditions

As it has been noted, it is useless to analyze the set of obtained coefficients α to understand the modification of the boundary conditions. Instead of this, we shall see the difference between velocity fields with classical and with optimal coefficients α similarly to [21]. This difference has been averaged in time over 200 days time interval in order to reveal persistent modifications of the flow produced by the optimal discretization.

This average difference of the velocity together with the original velocity are presented in fig. 5. We zoom the Southern part of the Black sea because it is in this region the difference shows the biggest values reaching 5 cm/s while in the middle of the sea it rarely exceeds 1 cm/s .
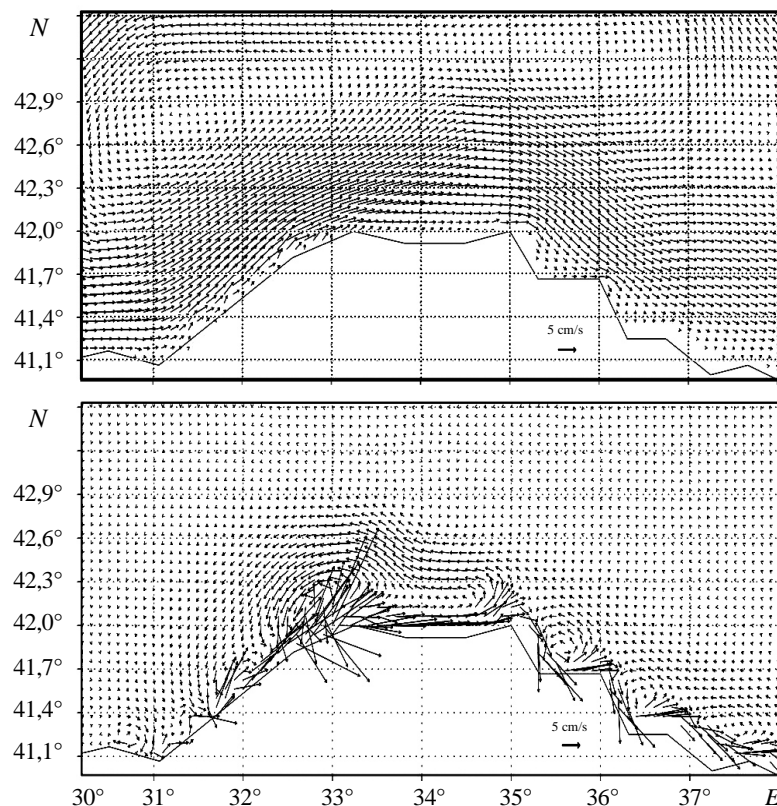

Fig. 5: Original velocity field (*a*) and its modification (*b*) near the boundary.

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. That means the flow is moved towards the boundary, allowing more optimal representation of a thin current on a coarse grid that brings the model solution towards observations.

6. Conclusion

The comparative study presented in this paper shows the influence of different model parameters on the solution. The study is confined to the analysis of a low resolution model with a rather limited physics. Consequently we must acknowledge the results may be valid only in the described case. Additional physical processes (baroclinic dynamics, variable density due to heat and salinity fluxes, etc.) may modify results of this study revealing other parameters the model may be sensitive to.

The main conclusion we can made from this comparison is the important role played by the boundary conditions on rigid boundaries. Almost all experiments show that the model is the most flexible with respect to control of coefficients α, this control allows us to bring the model's solution closer to the solution of the high-resolution model or to the observed data.

Optimal α found in the assimilation window remain optimal long time after the end of assimilation improving the forecasting ability of the model. We could see that the fourth order model in the square box allows us to divide by two the forecast error of the 20 days forecast. Optimal α obtained in one month assimilation remains optimal even for a one year run of the Black sea model.

Finally, the long time sensitivity of the model's solution to α exceeds the sensitivity to almost all other parameters including the sensitivity to initial conditions. A perturbation of α of a given small norm results in a bigger perturbation of the model's solution than a perturbation of some other parameter of an equal norm.

However, we could see that the influence of boundary conditions is only important on long time scales, i.e. time scales that exceeds the characteristic time of the domain. In both experiments presented above the characteristic time was approximately equal to 5 days and in both experiments the sensitivity to α becomes important on scale longer than 5 days. On the other hand, on short scales, it is the bottom topography that influences the most the model's solution. Both in the Black sea and in the square box the sensitivity to topography is approximately 40 times

more important than the sensitivity to α.

In addition to that, we should note that usually prescribed boundary conditions (impermeability and no-slip conditions have been used here as the initial guess for the minimization of the cost function) seem not to be optimal for the model. As we can see in fig. 3, modifying α we can bring the model much closer to the high resolution model or to the observational data. But, the numerical scheme is strongly modified in the assimilation process violating even impermeability condition.

Taking into account an important influence of the numerical scheme that introduces boundary conditions into the model, it is reasonable to think about identification of the optimal scheme by data assimilation process instead of prescribing classical boundary conditions.

REFERENCES

1. *Lorenz E.N.* Deterministic non periodic flow // J. Atmos. Sci. – 1963. – vol. 20. – P. 130-141.
2. *Le Dimet F.-X.* A general formalism of variational analysis: Tech. Rep. OK 73091: CIMMS report, Normann, 1982.
3. *Le Dimet F.-X., Talagrand O.* Variational algorithm for analysis and assimilation of meteorological observations. Theoretical aspects // Tellus. – 1986. – vol. 38A. – P. 97-110.
4. *Lions J.-L.* Contrôle optimal de systèmes gouvernés pas des équations aux dérivées parielles. – Dunod, 1968.
5. *Marchuk G.I.* Formulation of theory of perturbations for complicated models //Appl. Math. Optimization. – 1975. – vol. 2. – P. 1-33.
6. *Leredde Y., Lellouche J. M., Devenon J. L., Dekeyser I.* On initial, boundary conditions and viscosity coefficient control for Burgers' equation // Int. J. for Numerical Methods in Fluids. – 1998. – vol. 28, № 1. – P. 113-128.
7. *Arakawa A., Lamb V.* Computational design of the basic dynamical processes of the ucla general circulation model. // Methods in Computational Physics. – 1977. – vol. 17. – P. 174-267.
8. *Gilbert J.C., Lemarechal C.* Some numerical experiments with variable storage quasi-Newton algorithms // Mathematical programming. – 1989. – vol. 45. – P. 407-435.
9. *Kazantsev E.* Identification of an Optimal Boundary Approximation by Variational Data Assimilation. // J. Comp. Phys. – 2010. – vol. 229. – P. 256-275.
10. *Arakawa Akio, Lamb Vivian.* Computational Design of the Basic Dynamical Processes of the UCLA General Circulation Model. – Academic Press, 1977. – vol. 17. – P. 174-267.
11. *Kazantsev E.* Optimal Boundary Discretisation by Variational Data Assimilation // Int. J. for Numerical Methods in Fluids. – 2011. – vol. 65, № 10. – P. 1231-1259. doi: 10.1002/fld.2240. http://hal.inria.fr/inria-00388862/fr/.
12. *Ide K., Courtier P., Ghil M., Lorenc A.C.* Unified Notation for data Assimilation: Operational, Sequential and Variational // J. of the Met. Soc. of Japan. – 1997. – vol. 75, № 1B. – P. 181-189.

13. *Kazantsev E.* Identification of Optimal Topography by Variational Data Assimilation. // J. Phys. Math. – 2009. – vol. 1. – P. 1-23. (http://www.dinahgroup.com/volume1 2009 JPM.htm).

14. *Hascoët L., Pascual V.* TAPENADE 2.1 user's guide: Technical Report 0300: INRIA, 2004. http://www.inria.fr/rrrt/rt-0300.html.

15. *Tber M-H., L. Hascoët L., Vidard A., Dauvergne B.* Building the Tangent and Adjoint codes of the Ocean General Circulation Model OPA with the Automatic Differentiation tool TAPENADE: Research Report 6372: INRIA, 2007. (https://hal.inria.fr/inria-00192415).

16. *Le Provost C., Verron J.* Wind-driven ocean circulation transition to barotropic instability // Dyn.Atmos.Oceans. – 1987. – vol. 11. – P. 175-201.

17. *Simmonet Eric, Ghil Michael, Ide Kayo et al.* Low-Frequency Variability in Shallow-Water Models of the Wind-Driven Ocean Circulation // J. Phys. Oc. – 2003. – vol. 33. – P. 729-752.

18. *Munk W.H.* On the wind-driven ocean circulation // Journal of Meteorology. – 1950. – vol. 7. – P. 3-29.

19. *Korotaev G., Oguz T., Nikiforov A., Koblinsky C.* Seasonal, interannual, and mesoscale variability of the Black Sea upper layer circulation derived from altimeter data // J. Geoph. Res. – 2003. – vol. 108, № C4 – P. 3122-3137. (doi:10.1029/2002JC001508).

20. *Korotaev G.K., Saenko O.A., Koblinsky C.J.* Satellite altimetry observations of the Black Sea level // J. Geoph. Res. – 2001. – vol. 106, № C1. – P. 917-933.

21. *Kazantsev E.* Boundary conditions control for a Shallow-Water model // Int. J. for Numerical Methods in Fluids. – 2011. – publised online, (doi: 10.1002/fld.2526).

АНОТАЦІЯ   Методика, заснована на сполучених рівняннях, була застосована до моделі «дрібної води» для оцінки чутливості її рішення до параметрів моделі. Розглядалися такі параметри як рельєф дна, початкові і граничні умови, коефіцієнти в'язкості і напруга тертя вітру на поверхні. Їх вплив аналізувалося з різних точок зору.

Задача була розглянута в двох конфігураціях: академічний варіант моделі в квадратному басейні і більш реалістична версія, що моделює течії в Чорному морі. Було показано, що граничні умови на твердих кордонах мають найбільший вплив на рішення в обох випадках. цей факт вказує на необхідність пошуку оптимальних граничних умов при розробці моделі.

АННОТАЦИЯ   Методика, основанная на сопряженных уравнениях, была применена к модели «мелкой воды» для оценки чувствительности ее решения к параметрам модели. Рассматривались такие параметры как рельеф дна, начальные и граничные условия, коэффициенты вязкости и напряжение трения ветра на поверхности. Их влияние анализировалось с различных точек зрения.

Задача была рассмотрена в двух конфигурациях: академический вариант модели в квадратном бассейне и более реалистичная версия, моделирующая течения в Черном море. Было показано, что граничные условия на твердых границах оказывают наибольшее влияние на решение в обоих случаях. Этот факт указывает на необходимость поиска оптимальных граничных условий при разработке модели.