Data assimilation methods for an oceanographic problem

Didier Auroux[1] and Jacques Blum[2]

[1] Laboratoire J.A. Dieudonné, Université de Nice Sophia-Antipolis, Parc Valrose, F-06108 Nice Cedex 2, France. `auroux@math.unice.fr`

[2] Laboratoire J.A. Dieudonné, Université de Nice Sophia-Antipolis, Parc Valrose, F-06108 Nice Cedex 2, France. `jblum@math.unice.fr`

1 Introduction

The dynamics of the oceans play a major role in the knowledge of our environment and especially in the Earth's climate. Over the past twenty years, the new satellite techniques for observing the oceans, and especially the use of altimeter measurements, have greatly improved our knowledge of the oceans by allowing synoptic monitoring of the surface. The measurements of the sea-surface height have clearly demonstrated the feasibility and the usefulness of satellite altimetry. It was with the availability of Topex/Poseidon data since 1992, that the oceanographic community began intensive exploitation of this new observational source. It has already given incomparable information to study the general circulation of the ocean, to estimate the energy levels of the upper ocean, and to examine the local dynamics of different regions of particular interest, such as the Gulf Stream area, the Kuroshio extension, the Antarctic circumpolar current and the tropical oceans.

At the interface between the two major components of oceanographic science, i.e. observations and models, lies the domain of so-called data assimilation (DA). DA covers all the mathematical and numerical techniques which allow us to blend as optimally as possible all the sources of information coming from theory, models and other types of data. Clearly these techniques may not only apply in oceanography but also to other environmental disciplines. DA allows us to recreate the time-space structure of a system from a set of information which has, in general, a large disparity in nature, in space-time distribution and in accuracy. There are two main categories of DA methods: variational methods based on the optimal control theory [Lio68] and statistical methods based on the theory of optimal statistical estimation. The prototype of the first class which is actually of interest here is the optimal control method which was first introduced in meteorology (see [Lew85], [LeD86], [Tal87]) and more recently for the ocean (see [Tha88], [She90], [Moo91], [Sch93], [Nec94], [Luo98]). The prototype of statistical methods is the Kalman filter whose introduction

in oceanography dates back roughly a decade (see, for example, [Ghi89] and [Ghi91]). The Kalman filter was extended to nonlinear cases ([Jaz70], [Gel74]) but it has been mostly applied in oceanography to quasi-linear situations of the tropical oceans ([Gou92], [Fuk95], [Fuk93], [Can96], [Ver99]). We also refer to the recent book of Bennett [Ben02] on inverse methods, both for oceanography and meteorology.

All DA techniques encounter major difficulties in practice for computing reasons: memory size and computing costs. The full Kalman filter would, in principle, require the manipulation of $(N \times N)$ matrices where N is the state vector dimension which is typically 10^7 or 10^8 in an oceanic problem. The optimal control adjoint method often requires several hundred iterations of the minimization process to converge, thus implying an equivalent number of model runs.

In this paper, we first focus our interest on the use of the variational adjoint method in a relatively simple ocean model in order to try to reconstruct the four-dimensional ocean system from altimetric surface observations of the ocean. The variational method uses the strong constraint hypothesis, i.e. the ocean circulation model is assumed to be exact. The assimilation process is carried out by an identification of the initial state of the dynamical system which minimizes a cost function. This cost function is the mean-square difference between the observations and the corresponding model variables. The functional will be minimized using a numerical unconstrained optimization method such as the limited memory BFGS algorithm (see [Gil89]). The gradient vector is obtained analytically from the adjoint state, which can be interpreted as the Lagrange multiplier of the model equations. We then use a dual method, which consists in considering the model as a weak constraint. The use of an observation vector as a Lagrange multiplier for this constraint allows us to consider the minimization problem in a dual way. The dual cost function, measuring the difference between the data and the model state corresponding to a vector of the observation space, is minimized in the observation space, still using the BFGS algorithm.

In section 2, we introduce the physical model used for the theorical and numerical results. The primal and dual methods applied to our ocean model are introduced in sections 3 and 4 respectively. Some numerical results are given in section 5. A few conclusions will be given in section 6.

2 Physical model

2.1 Quasi-geostrophy

The system which governs the behaviour of the ocean is called the primitive equation system, constituted by the conservation laws of mass, momentum (Navier-Stokes equations), temperature and salinity. Most large-scale geophysical flows are based on the geostrophic equilibrium between the rotational

effect due to the Coriolis force and the horizontal pressure gradient.

We will use here a simplified quasi-geostrophic ocean model. This model arises from the primitive equations, assuming first that the rotational effect (Coriolis force) is much stronger than the inertial effect. This can be quantified by the fact that the ratio between the characteristic time of the rotation of the Earth and the inertial time is small. This ratio is called the Rossby number. The quasi-geostrophic model also assumes that the size of the ocean is small compared to the size of the Earth, and that this ratio is close to the Rossby number. Quasi-geostrophy finally assumes that the depth of the basin is small compared to its width (the ocean is supposed to be a thin layer of the Earth). In the case of the Atlantic Ocean, all these assumptions are not valid, but it has been shown that this approximate model reproduces quite well the ocean circulations at intermediate latitudes, such as the Gulf Stream.

The thermodynamic effects are neglected, and we also assume that the forcing is due to the wind at the surface of the ocean and that the dissipation is essentially due to bottom and lateral friction.

2.2 Equations of the model

The ocean is supposed to be stratified in n layers, each of them having a constant fluid density [Hol78]. The quasi-geostrophic model is obtained by making a first order expansion of the Navier-Stokes equation with respect to the Rossby number [Ped79]. The model system is then composed of n coupled equations resulting from the conservation law of the potential vorticity. The equations can be written as :

$$\frac{D_1\left(\theta_1(\Psi) + f\right)}{Dt} - \beta \Delta^2 \Psi_1 = F_1 \qquad \text{in } \Omega \times]0, T[, \tag{1}$$

at the surface layer ($k = 1$),

$$\frac{D_k\left(\theta_k(\Psi) + f\right)}{Dt} - \beta \Delta^2 \Psi_k = 0 \qquad \text{in } \Omega \times]0, T[, \tag{2}$$

at intermediate layers ($k = 2, \ldots, n - 1$), and

$$\frac{D_n\left(\theta_n(\Psi) + f\right)}{Dt} + \alpha \Delta \Psi_n - \beta \Delta^2 \Psi_n = 0 \qquad \text{in } \Omega \times]0, T[, \tag{3}$$

at the bottom layer ($k = n$), where

- $\Omega \subset \mathbb{R}^2$ is the circulation basin and $]0, T[$ is the time interval,

- n is the number of layers,

- Ψ_k is the stream function at layer k, Ψ is the vector $(\Psi_1, \ldots, \Psi_n)^T$,

- θ_k is the sum of the dynamical and thermal vorticity at layer k :

$$\theta_k(\Psi) = \Delta\Psi_k - (W\Psi)_k,$$

with $-(W\Psi)_k = \dfrac{f_0^2\rho}{H_k g}\left(\dfrac{\Psi_{k+1} - \Psi_k}{\rho_{k+1} - \rho_k} - \dfrac{\Psi_k - \Psi_{k-1}}{\rho_k - \rho_{k-1}}\right).$

- f is the Coriolis force (f_0 is the Coriolis force at the reference latitude of the ocean).
  In the β-plane approximation, the Coriolis force varies linearly with respect to the latitude.
- g represents the gravity, ρ_k the fluid density at layer k (and ρ the average fluid density), and H_k the depth of the layer k,

- $\dfrac{D_k}{Dt}$ is the Lagrangian particular derivative : $\dfrac{D_k}{Dt} = \dfrac{\partial}{\partial t} + J(\Psi_k, .)$,

  where J is the Jacobian operator $J(f, g) = \dfrac{\partial f}{\partial x}\dfrac{\partial g}{\partial y} - \dfrac{\partial f}{\partial y}\dfrac{\partial g}{\partial x}$,

- $\Delta\Psi_n$ represents the bottom friction dissipation, $\Delta^2\Psi_k$ represents the lateral friction dissipation,

- and F_1 is the forcing term, the wind stress applied to the ocean surface.

2.3 Boundary conditions

The tridiagonal matrix W (used to couple the stream functions at different layers) can be diagonalized :

$$W = P.\mathrm{diag}(\lambda_1, \ldots, \lambda_n).P^{-1}, \qquad (4)$$

where $0 = \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$ are the eigenvalues, and P is the transformation matrix. We can then define the mode vector of the stream functions $\Phi = (\Phi_1, \ldots, \Phi_n)^T$:

$$\Phi = P^{-1}\Psi.$$

The first mode Φ_1 corresponds to the eigenvalue 0 and is called the barotropic mode. The next modes ares the baroclinic modes. The boundary conditions result from the mass conservation law (Holland 1978), and can then be written as :

$$\Phi_1 = 0 \qquad \text{in } \partial\Omega\times]0, T[,$$

$$\int_\Omega \Phi_k(t)d\sigma = 0 \qquad \forall t \in [0, T], \quad \forall k \geq 2, \qquad (5)$$

and

$$\Delta\Psi_k(t) = 0 \qquad \text{in } \partial\Omega\times]0, T[, \quad \forall k. \qquad (6)$$

The initial conditions $\Psi_k(0)$ complete the equations of the direct model.

3 Primal variational method

We suppose that the data we want to assimilate come from satellite measurements of the sea-surface height, which is directly related to the upper layer stream function Ψ_1 by $h = \dfrac{f_0}{g}\Psi_1$. Thus, we assume that we have an observational stream function Ψ_1^{obs}. These observations are only available at times t_i, $i = 1 \ldots N$, over the data assimilation period $[0, T]$, and are also discrete in space. We consider then that the vector $\Psi_1^{obs}(t_i)$ represents the observations of the ocean surface available at time t_i.

The control vector u (which has to be determined) is the initial state of the stream functions at all layers $(\Psi_k(0))_{k=1 \ldots n}$.

3.1 Cost function

We can define a cost function

$$
\begin{aligned}
\mathcal{J}(u) = \frac{1}{2}\sum_{i=1}^{N}\langle R_i^{-1}\left(H_i\Psi_1(t_i) - \Psi_1^{obs}(t_i)\right), H_i\Psi_1(t_i) - \Psi_1^{obs}(t_i)\rangle \\
+ \frac{1}{2}\langle P_0^{-1}u, u\rangle,
\end{aligned}
\tag{7}
$$

where P_0 and R_i are covariance matrices, H_i are (linear) observation operators connecting observations Ψ_1^{obs} and model solutions Ψ_1, and $\langle \, . \, , \, . \, \rangle$ is the canonical real scalar product.

The first part of the cost function quantifies the difference between the observations and the state function, and the second part is a regularisation term. The inverse problem which consists in the minimization of $\mathcal{J}$ is then well-posed.

3.2 Adjoint model

In order to minimize the cost function, we need its gradient $\nabla\mathcal{J}$. Because of the large dimension of the model state vector (more than 10^6), it is not possible to compute directly the gradient by using finite difference methods. The gradient vector of the functional is then obtained by solving backwards in time the adjoint model ([LeD86]). The quasi-geostrophic adjoint equations are :

$$
\begin{aligned}
\frac{\partial \theta_1^T(\Lambda)}{\partial t} - \Delta J(\Psi_1, \Lambda_1) - (W^T J(\Psi, \Lambda))_1 - J(\Lambda_1, \theta_1(\Psi) + f) \\
- \beta \Delta^2 \Lambda_1 = E_1
\end{aligned}
\tag{8}
$$

at the surface layer,

$$\frac{\partial \theta_k^T(\Lambda)}{\partial t} - \Delta J(\Psi_k, \Lambda_k) - (W^T J(\Psi, \Lambda))_k - J(\Lambda_k, \theta_k(\Psi) + f) \\ -\beta \Delta^2 \Lambda_k = 0 \qquad (9)$$

at the intermediate layers, and

$$\frac{\partial \theta_n^T(\Lambda)}{\partial t} - \Delta J(\Psi_n, \Lambda_n) - (W^T J(\Psi, \Lambda))_n - J(\Lambda_n, \theta_n(\Psi) + f) \\ +\alpha \Delta \Lambda_n - \beta \Delta^2 \Lambda_k = 0 \qquad (10)$$

at the bottom layer, in $\Omega \times]0, T[$, where

- $\Lambda_1, \ldots, \Lambda_n$ is the adjoint vector,

- $\theta_k^T(\Lambda) = -\Delta \Lambda_k + (W^T \Lambda)_k$ is the *vorticity* corresponding to the adjoint state,

- and E_1 is the derivative of $\mathcal{J}$ with respect to Ψ_k :

$$E_1(t) = \sum_{i=1}^{N} R_i^{-1}(H_i \Psi_1(t) - \Psi_1^{obs}(t)) \, \delta(t - t_i).$$

If we denote by $\chi = (\chi_1, \ldots, \chi_n)^T$ the modal adjoint vector :

$$\chi = P^T \Lambda,$$

the space boundary conditions satisfied by the adjoint state Λ are :

$$\chi_1 = 0 \qquad \text{in } \partial \Omega \times]0, T[,$$
$$\int_{\Omega} \chi_k(t) d\sigma = 0 \qquad \forall t \in [0, T], \quad \forall k \geq 2, \qquad (11)$$

and

$$\Delta \Lambda_k(t) = 0 \qquad \text{in } \partial \Omega \times]0, T[, \quad \forall k. \qquad (12)$$

The gradient of the first part of $\mathcal{J}$ is obtained by solving equations (8-12) with a final condition of nullity of the adjoint state. The gradient of the second part of $\mathcal{J}$ is obtained directly by deriving it with respect to u, and we obtain :

$$\nabla \mathcal{J} = H(-\Delta + W)H^{-1} \begin{pmatrix} \Lambda_1(0) \\ \vdots \\ \Lambda_n(0) \end{pmatrix} + P_0^{-1} u \qquad (13)$$

where H is the diagonal matrix with the layers' depths H_k on the diagonal.

3.3 Minimization process

The numerical minimization of the cost function $\mathcal{J}$ can be realized using a quasi-Newton method. The Newton class of minimization algorithms is based on an iterative process, using at iteration k the descent direction $d_k = -H_k^{-1}.\nabla\mathcal{J}(x_k)$, where $H_k = \nabla^2\mathcal{J}(x_k)$ is the Hessian matrix of the cost function. The direct computation of the Hessian matrix is impossible (its dimension being too large), but it is possible to evaluate it, using the second order adjoint equations. However, the inversion of H is nearly impossible. The quasi-Newton algorithms consist in replacing the inverse of the Hessian matrix by a succession of matrices $(W_k)_{k\in\mathbb{N}}$ which are symmetric positive definite approximations to H^{-1}.

The BFGS algorithm ([Bro69]) uses the following update formula :

$$W_{k+1} = U(W_k, s_k, \eta_k) := \left(I - \frac{s_k \otimes \eta_k}{\langle \eta_k, s_k \rangle}\right) W_k \left(I - \frac{\eta_k \otimes s_k}{\langle \eta_k, s_k \rangle}\right) + \frac{s_k \otimes s_k}{\langle \eta_k, s_k \rangle}$$

with $s_k = x_{k+1} - x_k$, $\eta_k = \nabla\mathcal{J}(x_{k+1}) - \nabla\mathcal{J}(x_k)$ and $a \otimes b : c \mapsto \langle b, c \rangle a$. The disadvantage of this formula is the need to store all pairs (s_k, η_k).

The L-BFGS algorithm ([Liu89]) is a limited memory version of the previous algorithm. Only the last M pairs are stored, M being often equal to 5. The update formula is then :

$$W_k = U(W_{k-1}, s_{k-1}, \eta_{k-1}), \quad 1 \le k \le M,$$

and

$$\begin{cases} W_k^0 = D_k, \\ W_k^{i+1} = U(W_k^i, s_{k-M+i}, \eta_{k-M+i}), \quad i = 0 \ldots M-1, \\ W_k = W_k^M, \end{cases}$$

for $k \ge M+1$, where D_k is a diagonal matrix. The update formula for D_k is :

$$D_{k+1}^{(i)} = \left(\frac{1}{D_k^{(i)}} + \frac{\eta_k^{(i)\,2}}{\langle \eta_k, s_k \rangle} - \frac{s_k^{(i)\,2}}{(D_k^{(i)})^2 \langle D_k^{-1} s_k, s_k \rangle} \right)^{-1}.$$

4 Dual method

4.1 General description

The primal method has many disadvantages. First, the minimization process is often stopped before convergence to the minimum, because of the size of the state vector. Moreover, it is also impossible to take into account a model error : in the previous section, we have supposed that the model and the equations were perfect. This is obviously not the case (for example, not all parameters are well known). The only solution to incorporate the model error

into the minimization process is to add corrective terms to the model, consider them as part of the control vector, and add a third term to the cost function. This is not computationally realistic because the size of the control vector would be multiplied by the number of time steps. Therefore, it is not possible to take into account in a straightforward way the model error in the primal variational approach.

A new approach to data assimilation problems has been recently introduced ([Amo95], [Ben92], [Cou97]). Rather than minimizing a cost function on the state space, the dual method consists in working in the observation space (which is smaller than the state space).

4.2 Dual algorithm

Instead of solving first the direct equations and then the adjoint equations in the primal variational approach, the dual method consists in solving first the adjoint equations in order to use the information contained in the observation vector, and then the direct equations in order to reconstruct a trajectory. The dual algorithm for the quasi-geostrophic model can be constructed as follows :

- Let m be an observation vector that can be directly related to Ψ_1 (assume that m is a vector containing an observation of a part of the ocean surface at different times t_i),
- Solve the adjoint equations (with a final condition equal to zero) :

$$
\frac{\partial \theta_1^T(\Lambda)}{\partial t} - \Delta J(\Psi_1, \Lambda_1) \ - (W^T J(\Psi, \Lambda))_1 - J\left(\Lambda_1, \theta_1(\Psi) + f\right)
$$
$$
- \beta \Delta^2 \Lambda_1 = \tilde{E}_1(m),
$$

$$
\frac{\partial \theta_k^T(\Lambda)}{\partial t} - \Delta J(\Psi_k, \Lambda_k) - (W^T J(\Psi, \Lambda))_k - J\left(\Lambda_k, \theta_k(\Psi) + f\right)
$$
$$
- \beta \Delta^2 \Lambda_k = 0, \quad 1 < k < n, \tag{14}
$$

$$
\frac{\partial \theta_n^T(\Lambda)}{\partial t} - \Delta J(\Psi_n, \Lambda_n) - (W^T J(\Psi, \Lambda))_n - J\left(\Lambda_n, \theta_n(\Psi) + f\right)
$$
$$
+ \alpha \Delta \Lambda_n - \beta \Delta^2 \Lambda_k = 0,
$$

where

$$
\tilde{E}_1(m)(t) = \sum_{i=1}^{N} H_i^T R_i^{-1}(m(t) - \Psi_1^{obs}(t)) \ \delta(t - t_i).
$$

- Solve the direct equations

$$\frac{D_1\left(\theta_1(\Psi)+f\right)}{Dt} - \beta\Delta^2\Psi_1 = F_1 + (Q\Lambda)_1,$$

$$\frac{D_k\left(\theta_k(\Psi)+f\right)}{Dt} - \beta\Delta^2\Psi_k = (Q\Lambda)_k, \qquad\qquad 1 < k < n \qquad (15)$$

$$\frac{D_n\left(\theta_n(\Psi)+f\right)}{Dt} + \alpha\Delta\Psi_n - \beta\Delta^2\Psi_n = (Q\Lambda)_n,$$

with the initial conditions

$$\Psi_k(0) = \Psi_k^e(0) + (P_0\Lambda(0))_k,$$

where Q and P_0 are statistical preconditioning matrices, and $\Psi_k^e(0)$ is an *a priori* estimation of $\Psi_k(0)$.

- Define the operator $\mathcal{D}$: $(\mathcal{D}m)(t) = \sum_{i=1}^{N} H_i\Psi_1(t_i)\ \delta(t - t_i).$

We can then define the dual cost function as follows :

$$\mathcal{J}_{\mathcal{D}}(m) = \frac{1}{2}\langle\mathcal{D}m, m\rangle - \langle\Psi_1^{obs}, m\rangle. \qquad (16)$$

$\mathcal{J}_{\mathcal{D}}$ measures the difference between $\mathcal{D}m$ and Ψ_1^{obs}, i.e. between the trace (in the observation space) of a solution of the direct model and the observation vector.

As $\mathcal{D}$ is a linear symmetric positive definite operator, the gradient is obviously given by

$$\nabla\mathcal{J}_{\mathcal{D}}(m) = \mathcal{D}m - \Psi_1^{obs}. \qquad (17)$$

It is therefore easy to perform the minimization of $\mathcal{J}_{\mathcal{D}}$, given its gradient, simply by using a quasi-Newton method such as a BFGS algorithm. Once the minimum has been found, it is easy to reconstruct the corresponding trajectory in the state space by solving (14-15).

We can observe that the minimization of the dual cost function takes place over a smaller space than the minimization of the primary one. Moreover, this method also takes into account the model error, which was numerically impossible in the classical approach.

5 Numerical results

5.1 Model parameters

The numerical experiments are performed for a square three-layered ocean. The basin has horizontal dimensions of 4000 km × 4000 km and its depth is 5 km. The layers' depths are 300 meters for the surface layer, 700 meters for the intermediate layer, and 4000 meters for the bottom layer. The ocean is

discretized by a Cartesian mesh of $200 \times 200 \times 3$ grid zones. The time step is 1.5 hour. The initial conditions are chosen equal to zero for a six-year ocean spin-up phase, the final state of which being then the initial state for the data assimilation period. Then the assimilation period starts (time $t = 0$) with this initial condition $(\Psi_k(0))$, and lasts 5 days (time $t = T$), i.e. 80 time steps. The numerical method used to minimize the cost functions is a limited memory BFGS quasi-Newton method. The M1QN3 code by Gilbert and Lemaréchal ([Gil89]) is used for our experiments.

The experimental approach consists in performing *twin experiments* with simulated data. First, a reference experiment is run and the corresponding data are extracted. This reference trajectory will be further called the exact solution. Experimental surface data are supposed to be obtained on every fifth gridpoint of the model, with a time sampling of 7.5 hours (every 5 time steps). Simulated surface data are then noised with a blank Gaussian distribution, and provided as observations for the cost function. The first guess of the assimilation experiments is chosen as the reference state of the ocean one year before the assimilation period. The results of the identification process are then compared to the reference experiment.

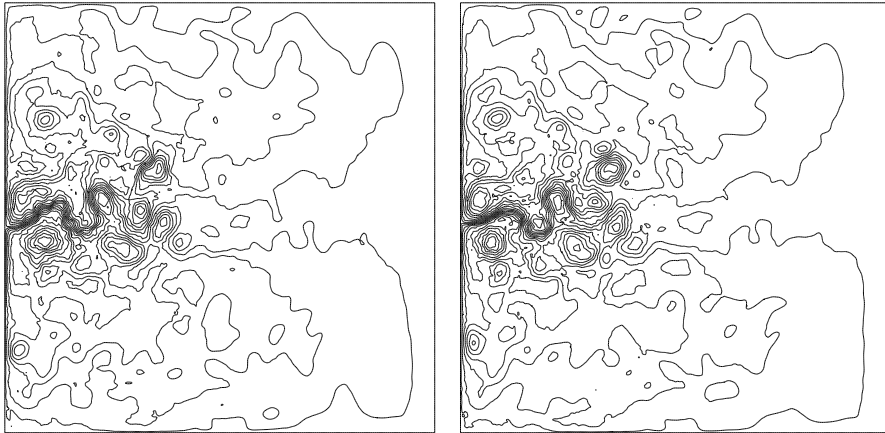## 5.2 Exact solution, noised observations

Fig. 1. Exact solution at the beginning (a), resp. the end (b), of the assimilation period.

Fig. 1 represents the stream function Ψ_1 at the surface layer, at the beginning and at the end of the assimilation period. These fields will be useful to measure the identification of the initial state, and also the reconstruction of

the stream function at the final time. One can observe the turbulent structure of the ocean, with a main current simulating a Gulf Stream type configuration.
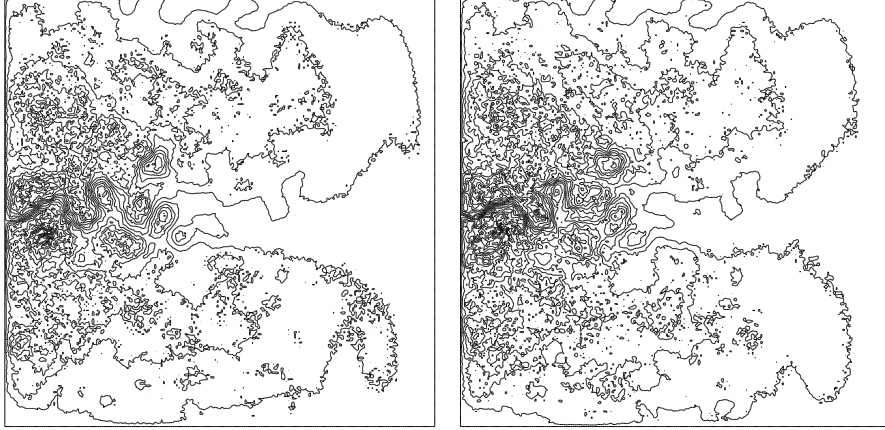

Fig. 2. Noised extracted data at the surface layer (a) and corresponding state at the end of the assimilation period (b)

The first part of fig. 2 represents the noised data extracted from the reference run, still at the surface layer. The second part of this figure is the corresponding state after a model run using the noised data as initial condition. This experiment clearly shows the importance of data assimilation. The model will indeed not smooth the trajectory, and it is not possible to obtain good predictions by simply integrating the model with observation data as initial conditions.

5.3 Primal method

The initial estimated vector to start the minimization process is chosen to be the reference state of the ocean one year before the assimilation period. The minimization process is stopped after 40 iterations, each iteration consisting of one integration of the forward direct model (in order to compute $\mathcal{J}$) and one integration of the backward adjoint model (in order to compute $\nabla \mathcal{J}$). The result of the minimization is shown on fig. 3-a. The direct model is then integrated over the assimilation period, using the computed minimizer as initial condition, and the corresponding state of the ocean at the end of the assimilation period is shown on fig. 3-b.
We can notice that the stream function of the solution at time $t = 0$ at the surface layer is comparable to the exact solution at the same time, but to a
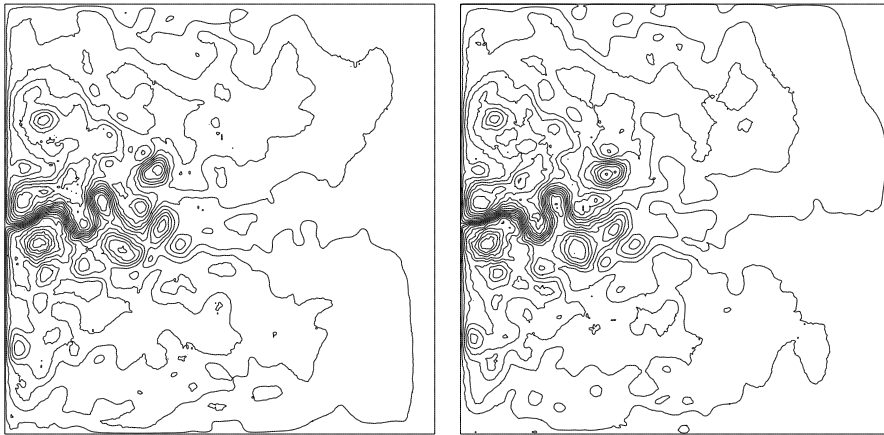
Fig. 3. Result of the minimization of the primal cost function. Solution at the beginning (a) and the end (b) of the assimilation period

lesser extent at time $t = T$. This can be explained by the fact that the primal algorithm gives more importance to the state at $t = 0$ than to any other time, as it is the control vector.
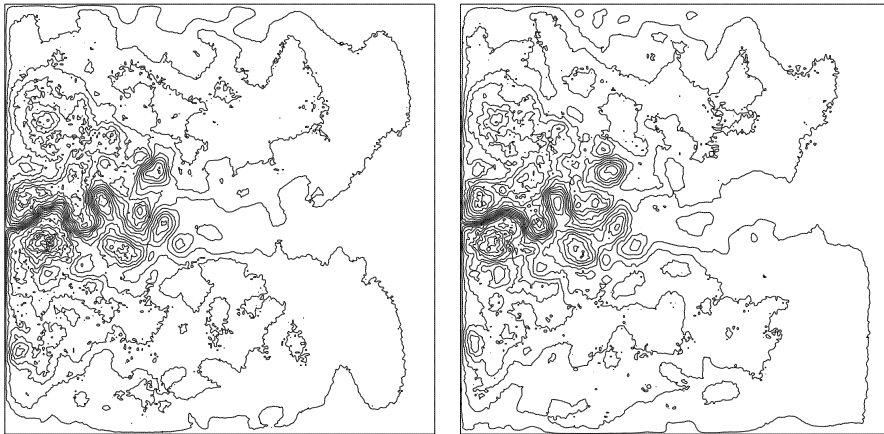
5.4 Dual method

Fig. 4. Result of the minimization of the dual cost function. Solution at the beginning (a) and the end (b) of the assimilation period

The initial estimated vector is the same. The minimization process is still stopped after 40 iterations, each iteration consisting now of one integration of the backward adjoint equations and one integration of the direct equations (in order to compute $\mathcal{J}_\mathcal{D}$ and $\nabla\mathcal{J}_\mathcal{D}$). The result of the minimization process is shown on fig. 4-a at time $t = 0$ and fig. 4-b at time $t = T$.

The stream function appears to be less smooth than in the primal case. This is due to the fact that the observations are noised and the dual algorithm works over the observation space. The corresponding state at the end of the assimilation period is closer to the exact solution than in the primal case. The dual algorithm looks indeed for a global acceptable solution : the control vector is a set of *observations* all over the assimilation period.

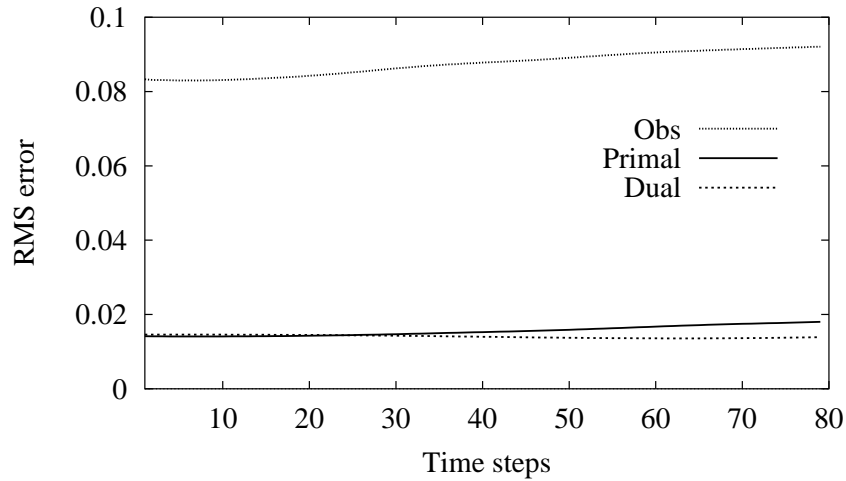## 5.5 Comparison between the two methods

Fig. 5. RMS errors of the different methods versus number of time steps over the assimilation period, using as initial conditions: the first noised observation (thin line), the primal algorithm (bold line) and the dual algorithm (dot line).

Fig. 5 represents the root mean square (RMS) error over the entire assimilation period between the exact solution and an identified solution, using either the observations, the primal solution, or the dual one. The RMS error at time t for one of these solutions is :

$$rms(t) = \frac{\int_\Omega \left[\Psi_1^{\mathrm{sol}}(t) - \Psi_1^{\mathrm{exact}}(t)\right]^2 d\sigma}{\int_\Omega \left[\Psi_1^{\mathrm{exact}}(t)\right]^2 d\sigma}.$$

The error reaches nearly 10% in the case of the trajectory resulting from the observations, and increases in time. This is due to the inherent non-linearities of the model.

In the case of the two data assimilation methods, the RMS error is clearly smaller (by a factor of more than 5). This proves once again the usefulness of data assimilation, which makes it possible to reconstruct a trajectory with less than 2% RMS error using noised observations with a 10% RMS error. The RMS error of the primal method tends to increase in time. This is in agreement with the remark concerning the resemblance between the stream function of the primal solution and that of the exact solution at initial time, and the loss of this similarity at final time. The RMS error of the dual solution is a little larger, but it tends to remain constant (and even decrease a little bit sometimes) in time. Moreover, as the dimension of the control vector u in the primal variational approach is 121203 ($201 \times 201 \times 3$), versus 28577 in the dual one ($41 \times 41 \times 17$), the minimization of the dual cost function is faster than for the primal approach.

6 Conclusion

As a matter of fact, the oceanic circulation model is governed by complex equations and behave as certain typical characteristics of the turbulent flow. Besides, in practice, the observation data are of various nature and should be combined together in the same functional to be minimized. In the present work, synthetic data are sampled using the whole surface layer. Generally, in the framework of the realistic oceanic data assimilation, the data are available only along ground tracks for time intervals corresponding to the satellite repeat period. Therefore, the optimal initial state would not be as well estimated because of the relatively small number of observations and their heterogeneous spatial distribution. Also, we notice that the performances of these methods have been assessed with a quasi-geostrophic model. It is necessary to apply them to a more complicated model such as the primitive equation model.

The dual data assimilation method is promising taking into account the computing time which is smaller than the primal optimal control method because of the smaller dimension of the observation space. Moreover, it enables to introduce an error in the model and not to consider the equations of the fluid as a strong constraint. The non linear character of the equations remains a problem for the proof of convergence of the dual method.

It remains a promising step towards operational oceanography.

References

[Amo95] Amodei, L.: Solution approchée pour un problème d'assimilation de données météorologiques avec prise en compte de l'erreur modèle. C.R. Acad. Sci. Paris, **321 (II)**, 1087–1094 (1995)

[Ben92] Bennett, A.F.: Inverse methods in physical oceanography. Cambridge University Press, Cambridge (1992)

[Ben02] Bennett, A.F.: Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press, Cambridge (2002)

[Bro69] Broyden, C.G.: A new double-rank minimization algorithm. Notices American Math. Soc., **16**, 670 (1969)

[Can96] Cane, M.A., Kaplan, A., Miller, R.N., Tang, B., Hackert, E.C., Busalacchi, A.J.: Mapping tropical Pacific sea level: data assimilation via a reduced state Kalman filter. J. Geophys. Res., **101**(C10), 22599–22617 (1996)

[Cou97] Courtier, P.: Dual formulation of four-dimensional variational assimilation. Q. J. R. Meteorol. Soc., **123**, 2449–2461 (1997)

[Fuk93] Fukumori, I., Benveniste, J., Wunsch, C., Haidvogel, D.B.: Assimilation of sea surface topography into an ocean circulation model using a steady state smoother. J. Phys. Oceanogr., **23**, 1831–1855 (1993)

[Fuk95] Fukumori, I.: Assimilation of Topex sea level measurements with a reduced-gravity, shallow water model of the tropical Pacific ocean. J. Geophys. Res., **100**(C12), 25027–25039 (1995)

[Gel74] Gelb, A.: Applied Optimal Estimation. MA: MIT Press, Cambridge (1974)

[Ghi89] Ghil, M.: Meteorological data assimilation for oceanographers. Part I: description and theoretical framework. Dyn. Atmos. Oceans, **13**, 171–218 (1989)

[Ghi91] Ghil, M., Manalotte-Rizzoli, P.: Data assimilation in meteorology and oceanography. Adv. Geophys., **23**, 141–265 (1991)

[Gil89] Gilbert, J.-Ch., Lemaréchal, C.: Some numerical experiments with variable storage quasi-Newton algorithms. Math. Prog., **45**, 407–435 (1989)

[Gou92] Gourdeau, L., Arnault, S., Ménard, Y., Merle, J.: Geosat sea-level assimilation in a tropical Atlantic model using Kalman filter. Ocean. Acta, **15**, 567–574 (1992)

[Hol78] Holland, W.R.: The role of mesoscale eddies in the general circulation of the ocean. J. Phys. Ocean., **8-3**, 363–392 (1978)

[Jaz70] Jazwinski, A.H.: Stochastic Processes and Filtering Theory. Academic, New York (1970)

[LeD86] Le Dimet, F.-X., Talagrand, O.: Variational algorithms for analysis and assimilation of meteorogical observations : theoretical aspects. Tellus, **38A**, 97–110 (1986)

[Lew85] Lewis, J.M., Derber, J.C.: The use of adjoint equations to solve a variational adjustment problem with convective constraints. Tellus, **37A**, 309–322 (1985)

[Lio68] Lions, J.L.: Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles. Dunod (1968)

[Liu89] Liu, D.C., Nocedal, J.: On the limited memory BFGS method for large scale optimization. Math. Prog., **45**, 503–528 (1989)

[Luo98] Luong, B., Blum, J., Verron, J.: A variational method for the resolution of a data assimilation problem in oceanography. Inverse Problems, **14**, 979–997 (1998)

[Moo91] Moore, A.M.: Data assimilation in a quasigeostrophic open-ocean model of the Gulf-Stream region using the adjoint model. J. Phys. Oceanogr., **21**, 398–427 (1991)

[Nec94] Nechaev, V., Yaremchuk, M.I.: Application of the adjoint technique to processing of a standard section data set: world ocean circulation experiment section S4 along 67S in the Pacific ocean. J. Geophys. Res., **100**(C1), 865–879 (1994)

[Ped79] Pedlosky, J.: Geophysical fluid dynamics. Springer-Verlag, New-York (1979)

[Sch93] Schröter, J., Seiler, U., Wenzel, M.: Variational assimilation of Geosat data into an eddy-resolving model of the Gulf Stream area. J. Phys. Oceanogr., **23**, 925–953 (1993)

[She90] Sheinbaum, J., Anderson, D.L.T.: Variational assimilation of XBT data. Part I. J. Phys. Oceanogr., **20**, 672–688 (1990)

[Tal87] Talagrand, O., Courtier, P.: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. Q. J. R. Meteorol. Soc., **113**, 1311–1328 (1987)

[Tha88] Thacker, W.C., Long, R.B.: Fitting dynamics to data. J. Geophys. Res., **93**, 1227–1240 (1988)

[Ver99] Verron, J., Gourdeau, L., Pham, D.T., Murtugudde, R., Busalacchi, A.J.: An extended Kalman filter to assimilate satellite altimeter data into a nonlinear numerical model of the tropical Pacific Ocean: method and validation. J. Geophys. Res., **104**, 5441–5458 (1999)