RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **96** (3), 2002, pp. 315–320
Matemática Aplicada / Applied Mathematics
Comunicación Preliminar / Preliminary Communication

A dual data assimilation method for a layered quasi-geostrophic ocean model

Didier Auroux and Jacques Blum

Abstract. In this paper, we introduce the equations of a layered quasi-geostrophic ocean model, and the corresponding data assimilation problem. We first give the variational formulation. We then point out the linear theory of duality. Finally, we apply duality to our nonlinear model by describing an algorithm to solve the data assimilation problem, introducing a dual cost function and a simple way to compute its gradient.

Un modelo oceánico quasigeostrófico y el problema asociado de asimilaciones de dados

Resumen. En este trabajo, se consideran las ecuaciones de un modelo oceánico quasigeostrófico y el problema asociado de asimilaciones de datos. En la primera parte, se considera la formulación variacional. En la segunda parte, se aplica la dualidad a nuestro modelo nolineal y se describe un algoritmo para resolver el problema de asimilación de datos.

1. Introduction

Data assimilation is the problem of finding a good estimation of the trajectory of a system, given discrete (in space and time) observations. The best estimation can be given by the minimum of a least-square cost function, using a variational approach of the problem. For that, it is necessary to introduce the adjoint state and equations in order to be able to compute quickly the gradient of the cost function. The minimization is then done using a gradient method. But this approach is only valid if we suppose the model perfect. It is indeed not possible to take into account a model error in the minimization process without considerably increasing the size of the control vector. A dual theory has been presented a few years ago ([1]-[3]) to remove this difficulty.
We first present the model we are currently using for numerical experiments, a multi-layer quasi-geostrophic ocean model. We also formulate the corresponding data assimilation problem in a classical variational approach. We then review the theory of duality, which has been introduced in a linear case. Finally, we give a corresponding dual algorithm to solve our problem for the quasi-geostrophic model.

Presentado por Haïm Brezis.
Recibido: 13 de Octubre de 2002. *Aceptado*: 31 de Octubre de 2002.
Palabras clave / Keywords: data assimilation, duality, quasi-geostrophy, oceanography
Mathematics Subject Classifications: 49-XX, 86A10

2. Direct model

2.1. Equations of the model

We consider here a layered quasi-geostrophic ocean model. This model arises from the primitive equations (Navier-Stokes), assuming in particular that the rotational effect (Coriolis force) is much stronger than the inertial effect, and that the size of the ocean is small compared to the size of Earth and the depth of the basin is small compared to its width. The thermodynamic effects are neglected. 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. 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 jet and the Gulf Stream. Most large-scale geophysical flows are based on the geostrophic equilibrium between the rotational effect and the pressure gradient. The equations can be written as :

$$\frac{D_1\left(\theta_1(\Psi)+f\right)}{Dt} - \beta\Delta^2\Psi_1 = F_1 \tag{1}$$

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

$$\frac{D_k\left(\theta_k(\Psi)+f\right)}{Dt} - \beta\Delta^2\Psi_k = 0 \tag{2}$$

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

$$\frac{D_n\left(\theta_n(\Psi)+f\right)}{Dt} + \alpha\Delta\Psi_n - \beta\Delta^2\Psi_n = 0 \tag{3}$$

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

- n is the number of layers,
- Ψ_k is the stream function at layer k, Ψ is the vector $(\Psi_1,\dots,\Psi_n)^T$,
- θ_k is the dynamic and thermal vorticity at layer k. $\theta_k(\Psi) = \Delta\Psi_k - (W\Psi)_k$, with $-(W\Psi)_k = \dfrac{f_0^2\rho}{H_k g}\left(\dfrac{\Psi_{k+1}-\Psi_k}{\rho_{k+1}-\rho_k} - \dfrac{\Psi_k-\Psi_{k-1}}{\rho_k-\rho_{k-1}}\right)$. The matrix W can be diagonalized : $W = P.D.P^{-1}$, P being the transformation matrix.
- f is the Coriolis force (f_0 is the Coriolis force at the reference latitude of the ocean), g represents the gravity, ρ_k the fluid density at layer k (and ρ the fluid average 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.

2.2. Cost function

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. Thus, we assume that we have an observational stream function Ψ_1^{obs}. By means of spatial interpolation and statistical considerations, we suppose that Ψ_1^{obs} is defined all over the ocean surface (that we will denote by Ω). On the other hand, these observations are only available at times t_i, $i = 1\dots N$, over the data assimilation period $[0,T]$.

The control vector u (which has to be determined) is the initial state of the stream functions at all layers $(\Psi_k(0))_{k=1\dots n}$. Then, we can define a cost function

$$\mathcal{J}(u) = \frac{1}{2}\sum_{i=1}^{N}\int_{\Omega}\left(\Psi_1(t_i) - \Psi_1^{obs}(t_i)\right)^2 dx + \frac{\varepsilon}{2}\|u\|^2 \tag{4}$$

where $\Psi_1(t_i)$ are deduced from the initial conditions u and the evolution equations (1-3), and $\|u\|^2$ is a Tikhonov regularisation term. The inverse problem which consists of the minimization of $\mathcal{J}$ is ill-posed, and thus a Tikhonov term is necessary. One can use the norm of the potential vorticity : $\|u\|^2 = \sum_{k=1}^{n} H_k \int_{\Omega} [\theta_k(\Psi)(t=0)]^2\, dx$. The choice of ε is difficult and depends on the error on observations. If ε is chosen too small, the minimum of $\mathcal{J}$ will give an excessively noisy solution Ψ. If ε is too large, the fields obtained are on the contrary too smooth. ε is generally chosen using the cross validation method.

3. Adjoint model

3.1. Adjoint equations

The adjoint equations (see [4] for further details) are :

$$\frac{\partial\theta_1^T(\Lambda)}{\partial t} - \Delta J(\Psi_1,\Lambda_1) - (W^T J(\Psi,\Lambda))_1 - J\left(\Lambda_1, \theta_1(\Psi) + f\right) - \beta\Delta^2\Lambda_1 = E_1 \tag{5}$$

at the surface layer,

$$\frac{\partial\theta_k^T(\Lambda)}{\partial t} - \Delta J(\Psi_k,\Lambda_k) - (W^T J(\Psi,\Lambda))_k - J\left(\Lambda_k, \theta_k(\Psi) + f\right) - \beta\Delta^2\Lambda_k = 0 \tag{6}$$

at the intermediate layers, and

$$\frac{\partial\theta_n^T(\Lambda)}{\partial t} - \Delta J(\Psi_n,\Lambda_n) - (W^T J(\Psi,\Lambda))_n - J\left(\Lambda_n, \theta_n(\Psi) + f\right) + \alpha\Delta\Lambda_n - \beta\Delta^2\Lambda_k = 0 \tag{7}$$

at the bottom layer, where

- $\Lambda_1, \dots, \Lambda_n$ 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 :

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

3.2. Gradient computation

The gradient of the first part $\mathcal{J}_0$ of $\mathcal{J}$ is obtained by solving equations (5-7) with a final condition of nullity of the adjoint state :

$$\nabla\mathcal{J}_0 = H(-\Delta + W)H^{-1}\begin{pmatrix}\Lambda_1(0)\\ \vdots \\ \Lambda_n(0)\end{pmatrix} \tag{8}$$

where H is the diagonal matrix with the layers' depths H_k on the diagonal. The gradient of the second part of $\mathcal{J}$ is obtained directly by deriving it with respect to u.
The numerical minimization of the cost function $\mathcal{J}$ can then be realized, using a quasi-Newton method.

3.3. Model error

In previous sections, we have supposed that the model and the equations were perfect. This is obviously not the case : for example, all parameters are not well known. This gives rises to a problem. How can we incorporate the model error in the minimization process ? The only solution is to add corrective terms in the model, to consider them as part of the control vector, and add a third term in the cost function $\mathcal{J}$. This is not computationally realistic because the size of the control vector (nearly 10^7) would be multiplied by the number of time steps (at least 80). Therefore, it is not possible to take into account in a straight forward way the model error in this variational data assimilation method.

4. Dual formulation

4.1. General description

A new approach of data assimilation problems has been recently introduced by Bennett [2], Amodei [1] and Courtier [3]. In a linear case, assuming that the model isn't perfect, this new approach is strictly equivalent to the classical approach, but the model error is taken into account in an inherent way in the dual formulation.
Let us consider a linear model

$$\frac{dy}{dt} + A(t)y = f + v, \quad y(0) = y_0 + u \tag{9}$$

where A is linear, v the unknown model error and $y_0 + u$ the unknown initial condition. Let us assume that we have some observations z_i of the system at different times t_i. We can then define a cost function

$$\mathcal{J}(u,v) = \frac{1}{2}\sum_{i=1}^{N}\langle R_i^{-1}\left(H_i y(t_i) - z_i\right), H_i y(t_i) - z_i\rangle + \frac{1}{2}\langle P_0^{-1}u, u\rangle + \frac{1}{2}\int_0^T \langle Q^{-1}v(t), v(t)\rangle dt \tag{10}$$

where P_0, Q and R_i are covariance matrices, and H_i are (linear) observation operators in order to connect observations and model solutions. We can then introduce the adjoint vector and equation, and we obtain the optimality system

$$\begin{aligned} &\frac{d\hat{y}}{dt} + A(t)\hat{y} = f + Q\hat{p}, && \hat{y}(0) = y_0 + P_0\hat{p}(0), \\ &-\frac{d\hat{p}}{dt} + A(t)^T\hat{p} = \sum_{i=1}^{N} H_i^T R_i^{-1}(z_i - H_i\hat{y}(t_i))\delta(t - t_i), && \hat{p}(T) = 0. \end{aligned} \tag{11}$$

It is then possible to define an operator $\mathcal{D}$ on the observation space as follows :

- Take $m = \begin{pmatrix} m_1 \\ \vdots \\ m_N \end{pmatrix}$ in the observation space (m_i is an observation of the system at time t_i),
- Solve the adjoint equation $-\frac{dp_m}{dt} + A(t)^T p_m = \sum_{i=1}^{N} H_i^T m_i \delta(t - t_i)$ with the *initial* condition $p_m(T) = 0$,
- Deduce $y_m(0) = P_0 p_m(0)$,
- Solve the direct equation $\frac{dy_m}{dt} + A(t)y_m = Qp_m$,

- Define $\mathcal{D}m = Hy_m$, which belongs to the observation space.

If we denote by $\tilde{y}$ the solution of (9) with $u = 0$ and $v = 0$ (the solution of the perfect model, with no error on the initial condition), and $d = z - H\tilde{y}$ the innovation vector, we obtain the following :

Proposition 1 *$(\mathcal{D} + R)\hat{m} = d$, where $\hat{m} = R^{-1}(z - H\hat{y})$.*

See [1] for the proof. □

It is then possible to define a dual cost function as follows :

$$\mathcal{J}_{\mathcal{D}}(m) = \frac{1}{2}\langle(\mathcal{D} + R)\, m, m\rangle - \langle d, m\rangle. \tag{12}$$

It can be proven that the optimal solution $\hat{y}$ of (11) is the sum of the *perfect* solution $\tilde{y}$ and the solution corresponding to $\hat{m}$, $y_{\hat{m}}$. It can also be proven that the minimum of $\mathcal{J}$ is related to the minimum of $\mathcal{J}_{\mathcal{D}}$, and the minima correspond to the same solution.

Therefore, we can see two main advantages in favour of the dual method :

- the inherent consideration of the model error,
- and the minimization of the dual cost function on the observation space, which is smaller than the state space.

4.2. Dual algorithm for the quasi-geostrophic model

Let us see now a dual algorithm for the quasi-geostrophic model. It is no longer possible to use an incremental approach $\hat{y} = \tilde{y} + y_{\hat{m}}$ because of the highly nonlinear characteristics of the model. We have then to construct a new dual algorithm for our model. This can be done 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 the entire ocean surface at different times t_i),
- Solve the adjoint equations (with a final condition equal to zero) :

$$\begin{aligned}
&\frac{\partial\theta_1^T(\Lambda)}{\partial t} - \Delta J(\Psi_1, \Lambda_1) - (W^T J(\Psi, \Lambda))_1 - J\left(\Lambda_1, \theta_1(\Psi) + f\right) - \beta\Delta^2\Lambda_1 = \tilde{E}_1(m),\\
&\frac{\partial\theta_k^T(\Lambda)}{\partial t} - \Delta J(\Psi_k, \Lambda_k) - (W^T J(\Psi, \Lambda))_k - J\left(\Lambda_k, \theta_k(\Psi) + f\right) - \beta\Delta^2\Lambda_k = 0, \quad 1 < k < n,\\
&\frac{\partial\theta_n^T(\Lambda)}{\partial t} - \Delta J(\Psi_n, \Lambda_n) - (W^T J(\Psi, \Lambda))_n - J\left(\Lambda_n, \theta_n(\Psi) + f\right) + \alpha\Delta\Lambda_n - \beta\Delta^2\Lambda_k = 0,
\end{aligned} \tag{13}$$

  where

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

- Solve the direct equations

$$\begin{aligned}
&\frac{D_1\left(\theta_1(\Psi) + f\right)}{Dt} - \beta\Delta^2\Psi_1 = F_1 + (Q\Lambda)_1,\\
&\frac{D_k\left(\theta_k(\Psi) + f\right)}{Dt} - \beta\Delta^2\Psi_k = (Q\Lambda)_k, \qquad 1 < k < n\\
&\frac{D_n\left(\theta_n(\Psi) + f\right)}{Dt} + \alpha\Delta\Psi_n - \beta\Delta^2\Psi_n = (Q\Lambda)_n,
\end{aligned} \tag{14}$$

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

We can define the dual cost function as follows :

$$\mathcal{J}_{\mathcal{D}}(m) = \frac{1}{2}\langle \mathcal{D}m, m\rangle - \langle \Psi_1^{obs}, m\rangle. \tag{15}$$

The gradient is obviously given by

$$\nabla \mathcal{J}_{\mathcal{D}}(m) = \mathcal{D}m - \Psi_1^{obs}. \tag{16}$$

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

5. Conclusion

A duality between the classical variational approach and the dual "observation vectors" method has been shown on a layered quasi-geostrophic ocean model. There are *a priori* many ways to extend the linear duality theory to a nonlinear model, but the one presented in this paper seems quite reasonable. Numerical experiments are under development and the first results show that, assuming the model is perfect, the dual method gives faster results that are comparable to those of the primary method.

References

[1] Amodei L. (1995). Approximate solution for a meteorological data assimilation with model error. *C. R. Acad. Sci. Paris Serie II a*, **321**, 1087–1094.

[2] Bennett A. F. (1992). *Inverse Methods in Physical Oceanography*, Cambridge University Press.

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

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

Didier AUROUX
Laboratoire J.A. Dieudonné
Université de Nice Sophia-Antipolis
Parc Valrose
F-06108 Nice Cedex 2
France
auroux@math.unice.fr

Jacques BLUM
Laboratoire J.A. Dieudonné
Université de Nice Sophia-Antipolis
Parc Valrose
F-06108 Nice Cedex 2
France
jblum@math.unice.fr