# SENSITIVITY ANALYSIS IN VARIATIONAL DATA ASSIMILATION.

*F. -X. LE DIMET, H. E. NGODOCK*
LMC - IMAG ( UJF ) B.P.53X 38041 Grenoble Cedex 9 ( France )
e-mail: ledimet@imag.fr ngodock@imag.fr

*I. M. NAVON*
Department of Mathematics and S.C.R.I.,Florida State University
Tallahassee,FL 32306-4052 (U.S.A.)
e-mail : navon@scri.fsu.edu

ABSTRACT

Optimal control theory is applied to a variational data assimilation problem in the context of the assimilation of altimeter data in a quasigeostrophic ocean model. Related to the issue of the minimization of the cost function, a sensitivity analysis is applied to the optimality system to derive the sensitivity of the retrieved control variable ( here the initial condition ) with respect to the observations. The derivation of the sensitivity of a response function in the case of data assimilation is reviewed and a new method of performing the derivation of this sensitivity is proposed .

# 1 Introduction

The basic equations used for numerical weather prediction are those of fluid dynamics, therefore numerical modelling in atmospheric or oceanic sciences is just a specialization of techniques used in fluid dynamics.

Nevertheless these geophysical fluids have a fundamental specificity: a particular situation cannot be reproduced. The classical methods of experimental sciences such as the duplication of an experiment cannot be used. in this instance.This fact should be taken into account in the methodology of modelling of geophysical flows.

The knowledge of initial conditions ( and sometimes even the boundary conditions ) in both oceanography and meteorology is a very difficult problem. One cannot perform a good prediction with a model without a good knowledge of its initial conditions ( assuming that the boundary conditions are known ). With a good model at our disposal, the first and most important problem to solve if we wish to carry out numerical prediction is the determination ( or the approximation ) of the initial conditions. But the problem cannot be solved unless we possess additional information. A natural additional piece of information needed is an observation of the phenomenon studied.

In terms of prediction, whatever the quality of a model, it cannot be used by itself to carry out a prediction. On the other hand, whatever the quality of a set of data, they cannot yield a prediction all by themselves.

As a consequence the important concept in numerical weather prediction is neither the model nor the data, but the couple consisting of the data and the model. Therefore the problem is to derive methods having the ability to take into account simultaneously the information contained in the equations of the model and in the data.

Optimal control methods ( Le Dimet,1982, Le Dimet and Talagrand, 1986 ) fulfill these requirements. They permit to retrieve atmospheric or oceanic fields in agreement with the model and the observations and to predict the evolution of the atmosphere or the ocean starting from a coherent situation.

In our point of view, this approach is not just a technical trick, but it creates a new and powerful perspective on numerical modelling in Earth Sciences.

As a consequence, the by-products of numerical modelling such as sensitivity analysis or predictability studies should be carried out in a slightly different way than used in classical modelling. The fundamental reason being that in variational analysis the result is obtained not only by solving the model, but by solving the optimality system.

Retrieved fields resulting from data assimilation processes on meteorological or oceanographical circulation models depend on data sets. Data assimilation in meteorology and in oceanography has improved considerably over the past two decades and the beginning of the current one. This is due to the fast improvement of both the techniques ( and means ) of data acquisition, and of the methods and algorithms for assimilation (Kalman Filter, Extended Kalman Filter, Optimal Interpolation and Variational Data assimilation). Some of the principal factors of this progress in oceanography are the adaptation of techniques used in meteorology ( where there is an extensive accumulated experience in data assimilation ), and the availability of new types of data: satellite altimetry, diffusiometry and acoustic tomography.

The data sets are characterized by configurations depending on both temporal and spatial sampling. In satellite altimetry, a data configuration is directly related to the orbital parameters ( inclination angle, repeat period, altitude, ... etc ) which are referred to here as *observation parameters.*

In the present work, we are interested in variational assimilation of altimeter data in an oceanic quasi-geostrophic general circulation model, and the sensitivity of retrieved fields with respect to the observation parameters.

The first part of this paper is devoted to theoretical results: we will show how to correctly derive sensitivity analysis in presence of data.

The second part of the paper will focus on sensitivity with respect to observations. An application to the estimation of optimal orbital characteristics of a network of satellites for observation of the ocean will be carried out.

# 2 General Sensitivity Analysis

Let us consider a general formulation of a model by the system:

$$F(X, K) = 0 \tag{1}$$

where $X$ may also be a function of time, $K$ is some parameter of the model. By parameter we mean any input of the model: classical parameters, initial or boundary conditions, or a combination of all of them.

For most geophysical problems $F$ is a nonlinear partial differential operator. In computational applications it becomes a mapping in a finite dimensional space.

For the sake of simplicity, we will assume that when $K$ is given then (1) has a unique solution $X(K)$.

Let $G$ be a given real function of $X$. It will be called *the response function.* ($G$ can also be a real function of both $X$ and $K$). By definition, the *sensitivity* of $G$ with respect to $K$ is the gradient of $G$ with respect to $X$. $G$ is an implicit function of $K$ trough $X$. Therefore the problem is : *how to estimate the gradient ?*

In order to estimate the gradient, we have to go trough the following steps:

*i)*The Gateaux derivative $\hat{F}$ of $F$ in a given direction $k$ is computed:

$$\hat{F}(X, K, k) = \lim_{\alpha \to 0} \frac{F(X, K + \alpha.k) - F(X, K)}{\alpha}$$

From (1) we get:

$$\frac{\partial F}{\partial X}.\hat{X} + \frac{\partial F}{\partial K}.k = 0 \tag{2}$$

$\hat{X}$ being the Gateaux derivative of $X$ in the direction $k$ and $\hat{G}$ the Gateaux derivative of $G$ is obtained by exhibiting the linear dependance of $\hat{G}$ with respect to $k$.

*ii)* Let us introduce the adjoint variable $P$ of $X$, which will be defined later according to our convenience. If we take the inner product of (2) with $P$, then transpose, it comes:

$$\left(\hat{X}, \frac{\partial F^t}{\partial X} \cdot P\right) + \left(k, \frac{\partial F}{\partial K}.P\right) = 0$$

By definition

$$\hat{G} = (\nabla G, k) = \left(\frac{\partial F}{\partial X}, \hat{X}\right) \tag{3}$$

*iii)* If $P$ is defined as the solution of :

$$\frac{\partial F^t}{\partial X}.P = \frac{\partial G^t}{\partial X} \tag{4}$$

Then we obtain :

$$\nabla G = -\frac{\partial F^t}{\partial K}.P \tag{5}$$

Therefore the sensitivity analysis is carried out in the following way:

*i) The direct model is solved, according to (1), for the value of the parameter in the vicinity of which the sensitivity analysis is requested.*

*ii) The adjoint model (4) is solved giving $P$ the adjoint variable.*

*iii) From (5), the sensitivity with respect to $K$ is computed. The computation of the adjoint permits to evaluate the sensitivity of the model with respect to any perturbation. This method needs only one evaluation of the model and one evaluation of the adjoint model.*

# 3 Variational Data Assimilation

During the last few years, an abundant research literature has been devoted to this field. The problem is to retrieve the state of the ocean or atmosphere using both a model and a set of observations.

Let us assume, that between times 0 and T, the state of the ocean ( or the atmosphere ), represented by X is governed by the differential equation:

$$\frac{dX}{dt} = F(X) \tag{6}$$

During this period the ocean is observed. For convenience, we assume that this observation is continuous in time. A more realistic assumption would only make the notation more complicated but would not change the essence of the method.

With an initial condition:

$$X(0) = U$$

equation (6) has a unique solution on the interval $[0, T]$. Let $J$ be a *cost function* measuring the discrepancy between the observation and the solution of (6) associated to the initial condition $U$. $J$ may be given by:

$$J(U) = \frac{1}{2}\int_0^T \|C.X(t) - X_{obs}(t)\|^2 dt \tag{7}$$

Here $C$ is a mapping from the space of the states of the ocean to the space of observations. The best fit between model and observation is achieved by minimizing $J$. The problem can be stated as: *"Determine $U^*$ minimizing $J$ "*. The determination of $U^*$ is carried out in two steps:

*i) Characterization of $U^*$*

To minimize the cost function $U^*$ should be a stationary point of $J$, and a necessary condition for optimality is

$$\nabla J(U^*) = 0$$

where $\nabla J$ is the gradient of $J$ with respect to the initial condition. The derivation of the gradient is obtained by introducing the adjoint model. An analysis similar to these performed above can be carried out, pointing out that a transposition with respect to the time is nothing else than an integration by parts, it yields that $U^*$ is the solution of the following so-called *Optimality System*:

$$\frac{dX}{dt} = F(X) \tag{8}$$

$$X(0) = U \tag{9}$$

$$\frac{dP}{dt} + \left[\frac{\partial F}{\partial X}\right]^t \cdot P = C^t(C.X - X_{obs}) \tag{10}$$

$$P(T) = 0 \tag{11}$$

$$\nabla J(U^*) = -P(0) = 0 \tag{12}$$

The optimality system is nothing but the Euler-Lagrange system of equations of the optimization problem. It is worth mentioning at this point that all the available information (i.e model and data) is included in this set of equations.

*ii) Computation of the optimal initial condition.*

The algorithm for the determination of the optimal initial condition consists in starting from a first guess, the direct model being integrated forward then the adjoint model being integrated backward. The gradient of the cost-function is deduced from the adjoint variable by (12). Then the gradient is used in some unconstrained optimization algorithm ( see Fiacco and McCormick [9] ) to improve the guess.

# 4 Sensitivity in Presence of Data

Let $K$ be a large parameter of a model governed by a differential equation:

$$\begin{aligned} \frac{dX}{dt} &= F(X, K) \\ X(0) &= U \end{aligned}$$

The initial condition being issued from a variational procedure as described above.

The traditional way for estimating the sensitivity of a given response function $G$ with respect to K consists in integrating the model and its adjoint then in identifying

$$X \equiv X$$
$$P \equiv P$$

$$F(X, K) \equiv \frac{dX}{dt} - F(X, K) \tag{13}$$

and then the general analysis is used to obtain the gradient of $G$ with respect to $K$.

This procedure is not always correct: if some perturbation $k$ is applied to the parameter $K$, then the initial condition is no longer optimal; this change in the optimal initial condition should be taken into account in the subsequent sensitivity analysis.

Let $K_{ad}$ be the set of all the admissible values of the parameter $K$. When $K$ spans the set $K_{ad}$ then $U$ will span a set $U_{ad}$ of all the admissible initial conditions. A perturbation on $K$ will induce a perturbation on $U$. The problem can then be stated as:

" *What is F in variational data assimilation ?* "

A common error is to consider(13) and to use the adjoint of the meteorological or oceanographical model to derive the sensitivity. The state of atmosphere or ocean is not a direct solution of an equation, but the solution of a problem of optimization. In order to apply the general sensitivity analysis, the equation to be considered is only the equation of which the optimal state of the ocean (or the atmosphere) is a solution i.e. the solution of the optimality system. Therefore in order to take into account the dependance of the perturbations we should consider the following:

$$X = \begin{pmatrix} X \\ P \end{pmatrix}$$

$$F(X, K) = \begin{pmatrix} \dfrac{dX}{dt} - F(X, K) \\ \dfrac{dP}{dt} + \left[\dfrac{\partial F}{\partial X}\right]^t \cdot P - C^t(CX - X_{obs}) \end{pmatrix}$$

It means that in the context of variational data assimilation, the concept of model should be extended to the Optimality System. The observation $X_{obs}$ is considered as the parameter $K$ of the general theory.

In order to estimate the sensitivity with respect to the parameter $K$, (i.e with respect to the observation ), let us apply a perturbation in a direction $k$ on this parameter. The Gateaux derivatives $\hat{X}$ of $X$ and $\hat{P}$ of $P$ are found as solution of the equations:

$$\frac{d\hat{X}}{dt} = \left[\frac{\partial F}{\partial X}\right] \cdot \hat{X} + \left[\frac{\partial F}{\partial K}\right] k \tag{14}$$

$$\hat{X} = 0 \tag{15}$$

$$\frac{d\hat{P}}{dt} + \left[\frac{\partial F}{\partial X}\right]^t \cdot \hat{P} + \left[\frac{\partial^2 F}{\partial X^2}\hat{X}\right]^t \cdot P = C^t C\hat{X} \tag{16}$$

$$\hat{P}(T) = 0 \tag{17}$$

$$\hat{P}(0) = 0 \tag{18}$$

Let $P = \begin{pmatrix} Q \\ R \end{pmatrix}$ be the adjoint variable to $X = \begin{pmatrix} X \\ P \end{pmatrix}$. Equation (14) is multiplied by $Q$ and (16) by $R$ (trough a suitable inner product), then they are integrated between times 0 and $T$. It comes:

$$\int_0^T \left[\left(\frac{d\hat{X}}{dt} - \left[\frac{\partial F}{\partial X}\right]\hat{X}, Q\right) + \left(\frac{d\hat{P}}{dt} + \left[\frac{\partial F}{\partial X}\right]^t \hat{P}, R\right)\right] dt + \int_0^T \left[\frac{\partial F}{\partial K}\right]$$

$$+ \int_0^T \left[\left(\left[\frac{\partial^2 F}{\partial X^2}\right]^t \cdot P, R\right) - \left(C^t C\hat{X}, R\right) + \left(C^t k, R\right)\right] dt = 0$$

After integrating by parts we obtain:

$$\left(\hat{X}(T), Q(T)\right) - \left(\hat{X}(0), Q(0)\right) - \int_0^T \left(\hat{X}, \frac{dQ}{dt} + \left[\frac{\partial F}{\partial X}\right]^t \cdot Q + \left[\frac{\partial^2 F}{\partial X^2}R\right]^t \cdot P - C^t CR\right) dt$$

$$+ \int_0^T \left(\left[\frac{\partial F}{\partial K}\right] k, Q\right) dt + \left(\hat{P}(T), R(T)\right) - \left(\hat{P}(0), R(0)\right) - \int_0^T \left(\hat{P}, \frac{dR}{dt} - \left[\frac{\partial F}{\partial X}\right] R\right) dt = 0$$

Let us assume that the response function is an integral of the form :

$$G = \int_0^T G(X)dt$$

$G$ being a scalar function. Therefore if $Q$ and $R$ are defined as the solution of:

$$\frac{dQ}{dt} + \left[\frac{\partial F}{\partial X}\right]^t \cdot Q + \left[\frac{\partial^2 F}{\partial X^2}P\right]^t \cdot R - C^t CR = \left[\frac{\partial G}{\partial X}\right] \tag{19}$$

$$\frac{dR}{dt} - \left[\frac{\partial F}{\partial X}\right] \cdot R = 0 \tag{20}$$

$$Q(0) = 0 \tag{21}$$

$$Q(T) = 0 \tag{22}$$

Then the gradient is given by:

$$\nabla G = -\left[\frac{\partial F}{\partial K}\right]^t \cdot Q \tag{23}$$

1.- If there is an observation at time $T$, the final condition for the adjoint of the direct model (11) becomes $P(T) = C^t\left(CX(T) - X_{obs}(T)\right)$ and the perturbed optimality system (17) becomes $\hat{P}(T) = C^t\left(C\hat{X}(T) - k\right)$ . Therefore, (22) is changed to $Q(T) + C^tCR(T) = 0$. However expression (23) yielding the gradient remains unchanged, although the values probably change as a consequence of the respective changes in $Q$ and $R$.

2.- We obtain a coupled system of two differential equations. The second one is nothing other but the so-called linear tangent model. The first equation has both initial and final conditions, the condition $Q(0) = 0$ being chosen to cancel the dependance with respect to the variation on the optimal initial condition. On the other hand, the second equation does not have any prescribed boundary condition.

3.- If $G \equiv J$ the cost function, it is clear that $R \equiv 0$ is a solution and we will find $Q \equiv P$ the adjoint variable. If $G \neq J$ then $R \equiv 0$ is a solution of the second equation. Formally we obtain an equation similar to the adjoint equation but with a different r.h.s.. It is clear that there is no reason to obtain $Q(0) = 0$ after the integration of the equation and therefore the sensitivity cannot be deduced from this "pseudo-adjoint" variable as is current practice in sensitivity analysis.

4-. The weakness of the method which consists to use this "pseudo-adjoint" is made evident when sensitivity with respect to the observation is sought. This is due to the fact that the observations are not explicit in the direct model. They are explicit in the optimality system and therefore from the formal point of view, they can be considered as any parameter in this *new model* which is the optimality system. For some response function $G$ depending on $X$ and therefore implicitly on $X_{obs}$, a similar analysis, as the one above, can be carried out and we find:

$$\nabla G = -C^t.R \tag{24}$$

In what follows, we propose a method ( an algorithm ) for solving this last system, given suitable hypothesis insuring existence and uniqueness of its solution, and hence the computation of the gradient defined by (23) or (24). We proceed in this way:

- Choose an initial condition $V$ for (20) : $R(0) = V$
- Integrate (20) forwards
- In the case of the preceding remark (1.), use the final coupling condition to deduce $Q(T)$, otherwise
- Integrate backwards (19)

Then we obtain $Q(0)$ as an implicit function of $V$ : $Q(0) = Q(0, V)$. The problem is then to choose a $V^*$ such that $Q(0, V^*) = 0$.

**Theorem 4.1** *If the Hessian of the cost function with respect to the initial condition of the direct model is positive definite, then there exists a unique $V^*$ such that*

$$Q(0, V^*) = 0$$

*The existence and uniqueness of $V^*$ prove the existence and uniqueness of the solution of our system.*

Proof

Let us fix a vector $V$ and set $R(0) = V$. We want to solve the system:

$$(I)\begin{cases} \frac{dQ}{dt} + \left[\frac{\partial F}{\partial X}\right]^t \cdot Q + \left[\frac{\partial^2 F}{\partial X^2}P\right]^t \cdot R - C^t C R & = & \left[\frac{\partial G}{\partial X}\right] \\ \frac{dR}{dt} - \left[\frac{\partial F}{\partial X}\right] \cdot R & = & 0 \\ Q(0) & = & 0 \\ Q(T) & = & 0 \end{cases}$$

and we seek a $V$ to satisfy $Q(0, V^*) = 0$. Let $(Q_0, R_0)$ be the solution of:

$$(II)\begin{cases} \frac{dQ_0}{dt} + \left[\frac{\partial F}{\partial X}\right]^t \cdot Q_0 + \left[\frac{\partial^2 F}{\partial X^2}P\right]^t \cdot R_0 - C^t C R_0 & = & 0 \\ \frac{dR_0}{dt} - \left[\frac{\partial F}{\partial X}\right] \cdot R_0 & = & 0 \\ R_0(0) & = & V \\ Q_0(T) & = & 0 \end{cases}$$

using the linearity of $(I)$, we can write :

$$Q = Q_0 + Q_1 \tag{25}$$
$$R = R_0 + R_1 \tag{26}$$

where $(Q_1, R_1)$ are defined by:

$$(III)\begin{cases} \frac{\partial Q_1}{\partial t} + \left[\frac{\partial F}{\partial X}\right]^t .Q_1 & = & \frac{\partial G}{\partial X} \\ \frac{\partial R_1}{\partial t} - \left[\frac{\partial F}{\partial X}\right] .R_1 & = & 0 \\ Q_1(T) & = & 0 \\ R_1(0) & = & 0 \end{cases}$$

where $R_1$ is identically null. $Q_0, Q_1, R_0, R_1$ are uniquely defined and:

$$Q(0) = Q_0(0) + Q_1(0) \tag{27}$$

System $(II)$ is nothing but the second order adjoint (SOA) described by Wang and al.[1]. And in the same paper the authors show that

$$Q_0(0) = H.V \tag{28}$$

where $H$ is the hessian of the cost function with respect to initial condition of the direct model. Thus

$$Q(0) = H.V + Q_1(0) \tag{29}$$

Therefore, if $H$ is positive definite ( by definition $H$ is symmetric ), then there exists a unique $V^*$ such that $Q(0) = 0$.

1.- The computation of the full $H$ and its inversion are computationally prohibitive, but the product $H.V$ is available trough the (SOA) for any $V$. Therefore, $V^*$ can also be characterized as the solution of:

$$f(V^*) = \text{Min}\{f(V), V \in \text{E}_{ini}\} \tag{30}$$

where

$$f(V) = \frac{1}{2}V^t.H.V + Q_1(0).V \tag{31}$$

and then an efficient conjugate gradient descent method can be used for this minimization.

2.- $V^*$ is nothing but the sensitivity of the optimal initial condition with respect to a perturbation on the parameter ( or on the observations ).

3.- The hypothesis of the theorem above is neither strong nor restrictive, but it is a natural necessary condition for the existence of the solution of the optimization problem.

# 5 Example

Let us consider a model given by the ordinary differential equation:

$$\frac{dX}{dt} = \alpha X$$

With the initial condition:

$$X(0) = U$$

X is a scalar variable and $\alpha$ is a constant, on the interval $[0, 1]$ and we assume that the observation $X_{obs}$ of $X$ is $\beta$ which is constant on $[0, 1]$.

The cost function $J$ is:

$$J(U) = \frac{1}{2}\int_0^1 (X(t) - \beta)^2 dt$$

We consider the response function:

$$G(\alpha) = \int_0^1 X(t)dt$$

$G$ is an implicit function of $\alpha$ trough $X$. Let us compute the sensitivity of $G$ to $\alpha$. The solution of the differential equation is:

$$X(t) = Ue^{\alpha t}$$

$J$ is explicitly computed as a function of $U$:

$$J(U) = \frac{1}{2}\int_0^1 \left(Ue^{\alpha t} - \beta\right)^2 dt$$

Setting the value of the derivative of $J$ with respect to $U$ equal to 0 yields the optimal value of $U$ namely:

$$U_{opt} = \frac{2\beta}{e^{\alpha} + 1}$$

leading to the explicit estimate of $G$ as a function of $\alpha$:

$$G = \frac{2\beta}{\alpha}\left(\frac{e^{\alpha} - 1}{e^{\alpha} + 1}\right)$$

From which the sensitivity is deduced by taking the derivative of $G$ with respect to $\alpha$:

$$\nabla G = \frac{2\beta}{\alpha(e^{\alpha} + 1)}\left[\frac{2e^{\alpha}}{e^{\alpha} + 1} - \frac{e^{\alpha} - 1}{\alpha}\right]$$

On the other hand, if the adjoint were used to estimate the sensitivity we would have to consider the differential equation:

$$\frac{dP}{dt} + \alpha P = 1$$

with the condition:

$$P(1) = 0$$

The solution of this equation being:

$$P(t) = \frac{1}{\alpha}\left(1 - e^{\alpha}e^{-\alpha t}\right)$$

then the sensitivity of $G$ would be given by:

$$\nabla G = -\int_0^1 X(t).P(t)dt$$

Using the explicit solutions of both the direct model and the adjoint model would give:

$$\nabla G = \frac{U}{\alpha}\left[e^{\alpha} - \frac{e^{\alpha} - 1}{\alpha}\right]$$

Writing $U$ explicitly yields ::

$$\nabla G = \frac{2\beta}{\alpha(e^{\alpha} + 1)}\left[e^{\alpha} - \frac{e^{\alpha} - 1}{\alpha}\right]$$

The two expressions for the sensitivity are different. Figure (1) shows the variations of both estimated sensitivities as function of $\alpha$. The second sensitivity estimate is erroneous since it does not take into account the derivative of $U$ with respect to $\alpha$.

Let us now apply the new sensitivity analysis as developed above to this problem. The optimality system corresponding to the minimization of $J$ is:

$$(A)\left\{\begin{array}{rcl} \frac{dX}{dt} & = & \alpha X \\ X(0) & = & U_{opt} \\ \frac{dP}{dt} + \alpha P & = & X - \beta \\ P(1) & = & 0 \end{array}\right.$$

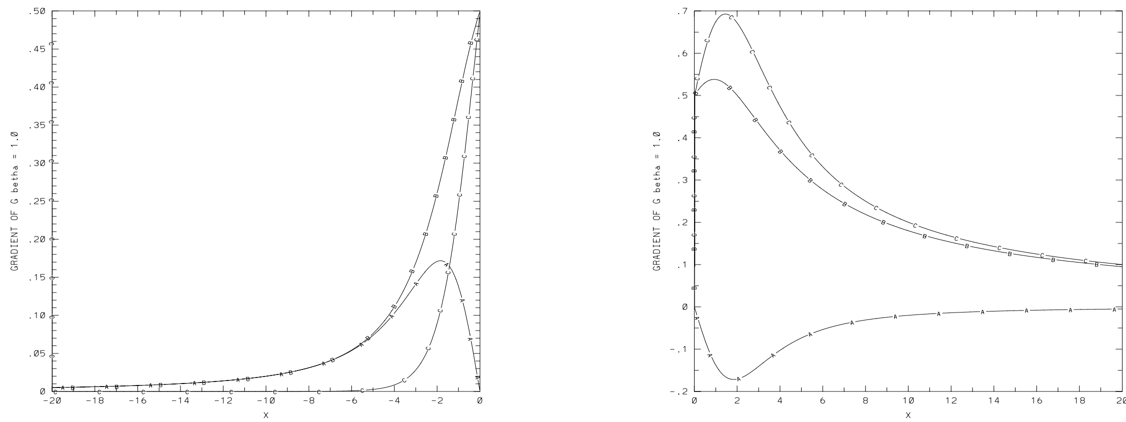


Figure 1: Gradients computed: A:direct analysis(exact result), B: using the adjoint equations, C: absolute difference of A and B. Left and right curves correspond respectively to negative and positive values of $\alpha$ represented on the x-axis

with solution:

$$X(t) = U_{opt} e^{\alpha t}$$

$$P(t) = \frac{U_{opt}}{2\alpha}\left(e^{\alpha t} - e^{2\alpha} e^{-\alpha t}\right) + \frac{\beta}{\alpha}\left(e^{\alpha} e^{-\alpha t} - 1\right)$$

The gradient of $J$ with respect to $U$ is:

$$\nabla J(U) = -P(0) = -\frac{U_{opt}}{2\alpha}\left(1 - e^{2\alpha}\right) - \frac{\beta}{\alpha}\left(e^{\alpha} - 1\right)$$

and equating $\nabla J(U)$ to zero leads to

$$U_{opt} = \frac{2\beta}{e^{\alpha} + 1}$$

which is equal to the one obtained in the direct analysis.

A perturbation $\tilde{\alpha}$ is applied to $\alpha$, and we derive the second order adjoint equations as described above to compute the sensitivity of $G$. The system to be solved is:

$$(B)\begin{cases} \dfrac{dQ}{dt} + \alpha Q &= 1 - R \\ \dfrac{dR}{dt} &= \alpha R \\ Q(1) &= 0 \\ Q(0) &= 0 \end{cases}$$

And the sensitivity is given by:

$$\nabla G(\alpha) \;=\; \int_0^1 [R(t)P(t) - Q(t)X(t)]dt \qquad (32)$$

System (B) is solved in the way proposed in the general theory (section 4). We set $R(0) = V$ and look for $V^*$ such that $Q(0) = 0$ and we find

$$V^* = \frac{2}{e^\alpha + 1}$$

. Therefore the solution of system (B) is:

$$R(t) = \frac{2e^{\alpha t}}{e^\alpha + 1}$$

$$Q(t) = \frac{1}{\alpha}\left[\frac{1}{e^\alpha + 1}\left(e^{2\alpha}e^{-\alpha t} - e^{\alpha t}\right) + \left(e^\alpha e^{-\alpha t} - 1\right)\right]$$

and if we substitute in (32), the expressions of $X, P, R, Q$ we obtain:

$$\nabla G(\alpha) = \frac{2\beta}{\alpha(e^\alpha + 1)}\left[\frac{2e^\alpha}{e^\alpha + 1} - \frac{e^\alpha - 1}{\alpha}\right]$$

which is the expected result ( the one obtained by direct computation in this simple example).

We may also consider $G$ as a function of $\beta$ since the optimal state $X$ depends on $\beta$. This is of great importance in data assimilation where the retrieved fields depend on data. The gradient of $G$ with respect to $\beta$ is therefore its sensitivity with respect to a perturbation on the data.

$$G(\beta) = \int_0^1 X(t)dt = \frac{2\beta(e^\alpha - 1)}{\alpha(e^\alpha + 1)}$$

and its gradient:

$$G\prime(\beta) = \frac{2(e^\alpha - 1)}{\alpha(e^\alpha + 1)}$$

Using the second order technique above, system (B) is unchanged ( and hence its solution), but the gradient in this case is given by the formula:

$$\nabla G(\beta) = \int_0^1 R(t)dt = \frac{2(e^\alpha - 1)}{\alpha(e^\alpha + 1)}$$

If we use the cost function at the optimum as a function of $\beta$, i.e $G = J$ then solving system (B) yields: $R \equiv 0$ and $Q \equiv P$. The sensitivity computed directly is ( also in this case ) equal to the one obtained by the (SOA).

# 6 Application

We apply the method (of computing the sensitivity) described above for the assimilation of observations in oceanography. We consider a quasigeostrophic general circulation model of the North Atlantic ocean. The model is governed by a system of partial differential equations ( derived from the primitive Navier-Stokes equations of fluid dynamics : see Pedlosky 1986 ), for $k = 1, ...., N$, $N$ being the number of layers into which the ocean has been subdivided .

$$\frac{D.}{Dt}\left[\Delta\Psi_k + f + \frac{f_0}{H}\left(h_{k+\frac{1}{2}} - h_{k-\frac{1}{2}}\right)\right] = F_k - T_k \tag{33}$$

$$\frac{D\phi}{Dt} = \frac{\partial\phi}{\partial t} + J(\Psi_k, \phi) \tag{34}$$

$$J(\Psi_k, \phi) = \frac{\partial\Psi_k}{\partial x}\frac{\partial\phi}{\partial y} - \frac{\partial\Psi_k}{\partial y}\frac{\partial\phi}{\partial x} \tag{35}$$

$\Psi_k$ is the stream function in the layer $k$, $f = f_0 + \beta y$ is the Coriolis force, $F$ is the wind stress, $T$ is the dissipation term, $h_{k+\frac{1}{2}} = \dfrac{f_0}{g\prime_{k+\frac{1}{2}}}(\Psi_{k+1} - \Psi_k)$ and $h_{N+\frac{1}{2}}$ is the bottom topography height; $g\prime_{k+\frac{1}{2}} = \dfrac{g(\rho_{k+1} - \rho_k)}{\rho_0}$

The domain in which we integrate these equations is a square box of $4000kms$ by size and $5kms$ of depth at mid-latitude. The discretization is carried out by using homogeneous finite differences in space for both zonal and meridional directions $\delta x = \delta y = 20kms$ , and a leap-frog time differencing scheme is used with $\delta t = 90min$. The observations are simulated from integrations of the model, and are made available only along satellite ground tracks that have also been simulated. The cost function we seek to minimize assumes the form :

$$\mathrm{J(U)} = \frac{1}{2}\sum_{i=1}^{N_T}\sum_{k=1}^{N_i}|\Psi_1^k(t_i) - \Psi_{1,obs}^k(t_i)|^2\delta t\delta x\delta y + \frac{\lambda}{2}\sum_{n=1}^{N}\sum_{m=1}^{N}|(\Delta\mathrm{U})_{nm}|^2\delta x\delta y$$

where: $N_T$ : is the number of time steps of integration of the model: we take $N_T = 464$ corresponding to 30 days.
$N_i$ : is the number of observed points at time $t_i$ $\Psi_1^k(t_i)$ :is the value of the field at point $k$ and time $t_i$
$U$ : is the initial condition of the model $\lambda$ : is the regularization (penalty) parameter which plays a very important role in variational data assimilation. In sensitivity analysis with respect to the observations, we may assign small values to $\lambda$ so that the discrepancy between the model solution and the data always dominates the regularization (penalty) term during the minimization process.

We carried out the validation of our assimilation by means of a twin experiment. We ran the model and stored a reference state, then used a fully decorrelated first-guess . The unconstrained minimization algorithm is of a quasi-Newton limited memory type with stopping criterion either on the number of iterations or on the norm of the gradient of the cost function.

We have tested the correctness of our computations for both first and second order adjoint by the Taylor's formula. The correctness of the gradient of the cost function and the product $H.V$ where $H$ is the Hessian matrix of the cost function are obtained by:

$$\zeta_1(\mu) = \frac{J(U + \mu V) - J(U)}{\mu\nabla J(U).V}$$

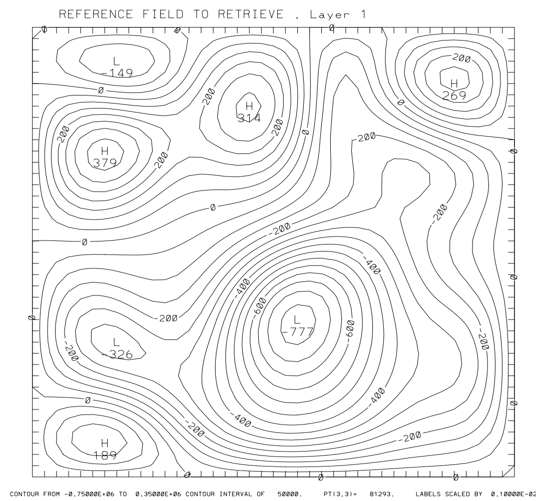


Figure 2: The initial condition of the reference state. It is the field to retrieve

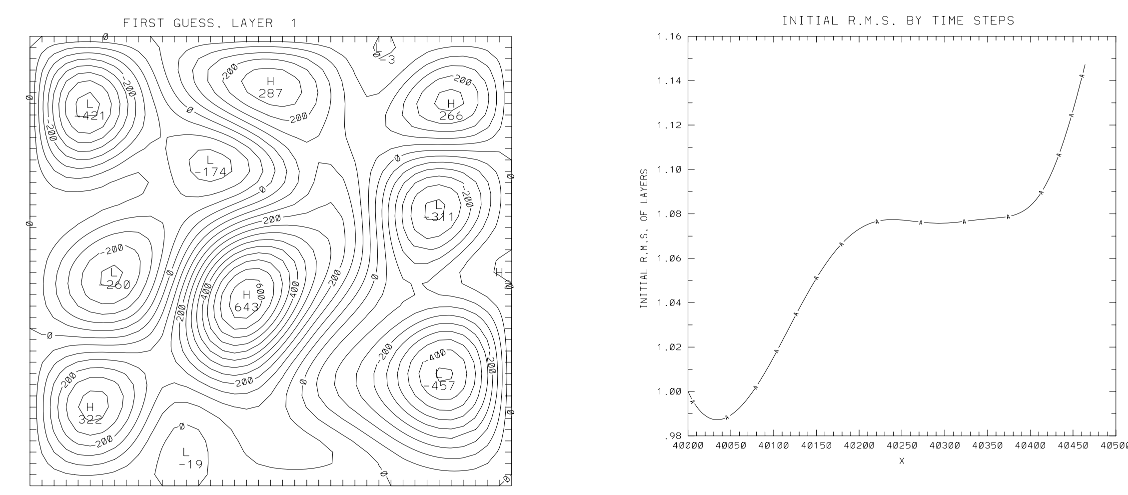


Figure 3: The fully decorrelated first guess and the initial r.m.s

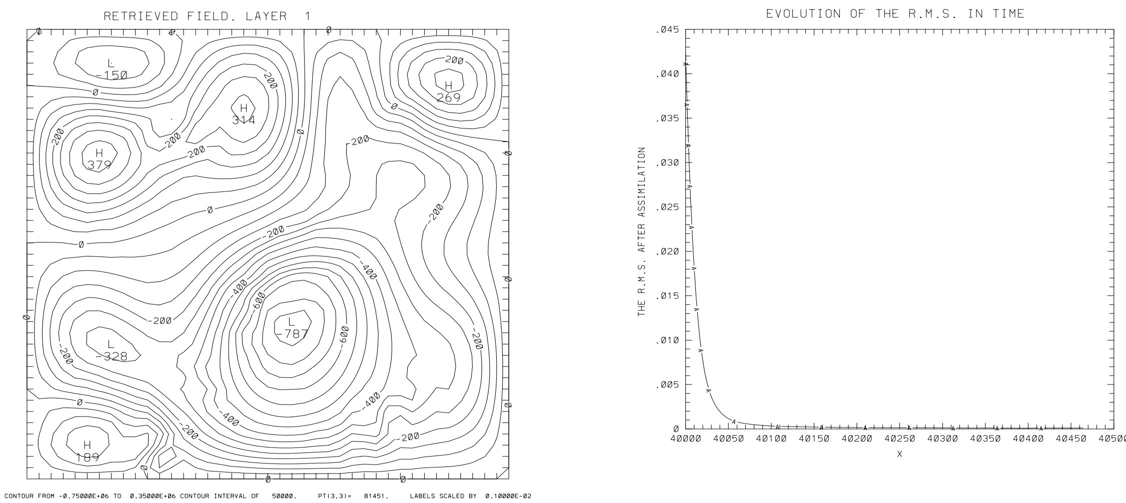


Figure 4: The initial condition retrieved and the rms between retrieved and reference states for $\lambda = 10^{-8}$

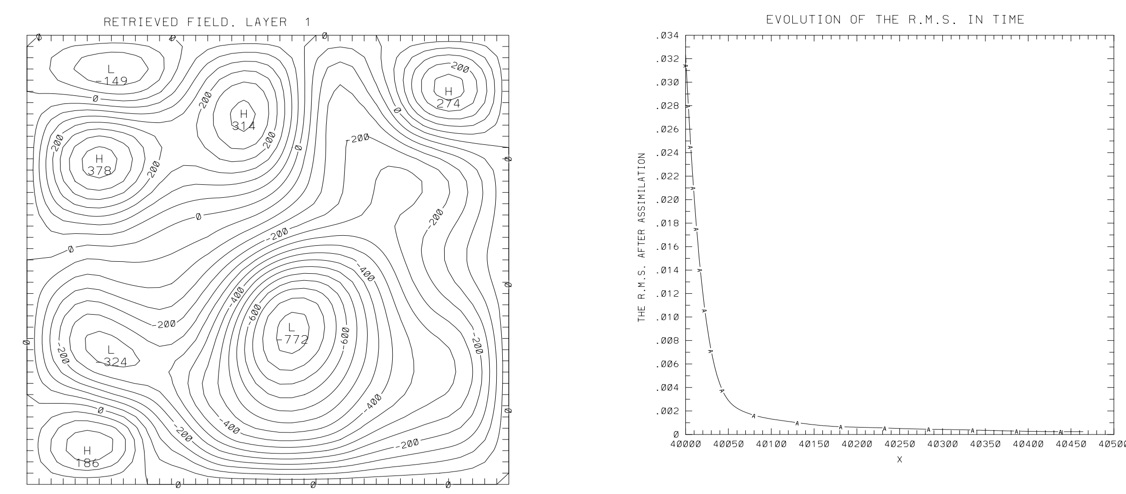


Figure 5: The initial condition retrieved and the rms between retrieved and reference states for $\lambda = 10^{-6}$

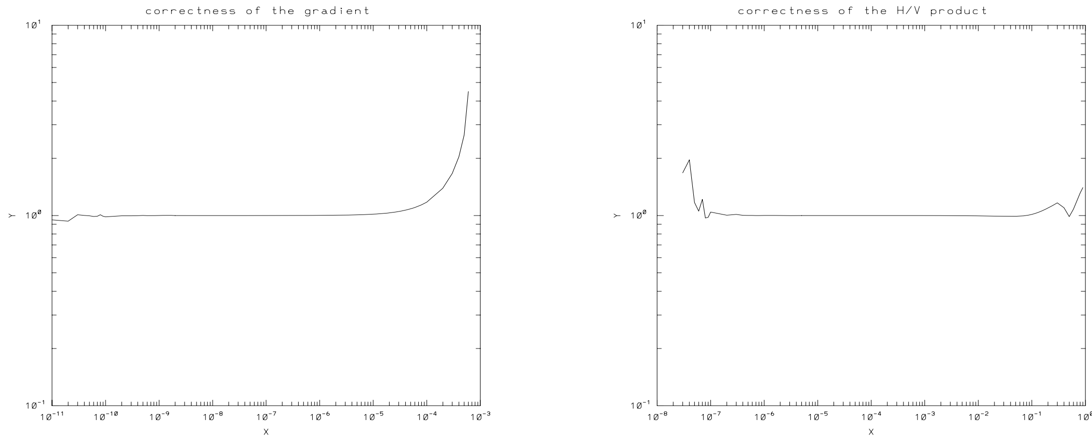


Figure 6: The correctness of the gradient of the cost function (left) and of the Hessian/Vector product (right)

$$\zeta_2(\mu) = \frac{J(U + \mu V) - J(U) - \mu \nabla J(U) \cdot V}{\frac{1}{2}\mu V^t \cdot H \cdot V}$$

where

$$\lim_{\mu \to 0} \zeta_1(\mu) = 1 \tag{36}$$

$$\lim_{\mu \to 0} \zeta_2(\mu) = 1 \tag{37}$$

We applied a 1% perturbation on the data in order to carry out the sensitivity analysis on the optimality system, and we also computed the r.m.s between the retrieved fields related to the perturbed data and the reference state.

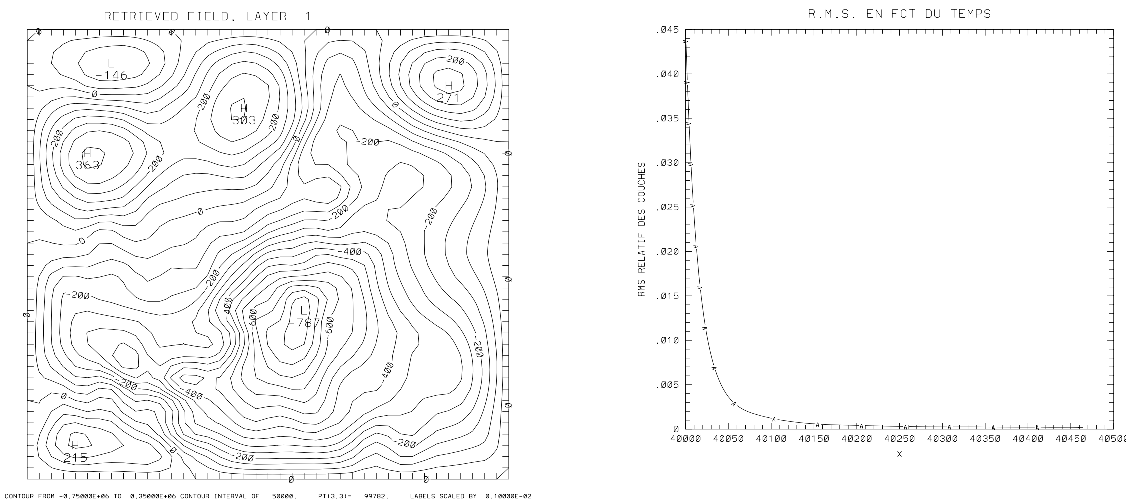


Figure 7: The initial condition retrieved with 1% perturbation on the data and the rms between retrieved and reference states for $\lambda = 10^{-6}$

This figure shows the sensitivity of the retrieved initial condition with respect to the 1% perturbation of the observations. We observe that the more points are closed to the tracks, the less they are sensitive to the perturbation.

# 7 Summary and conclusions

In this paper we have used optimal control theory of partial differential equations and applied it to a variational data assimilation problem in a one-layered quasi-geostrophic ocean model where only the initial conditions served as control variables. The Limited-Memory quasi-Newton method of J.C Gilbert and C. Lemaréchal [14] was applied to the problem of minimizing the cost function consisting of the weighted sum of squares between the model computed field and known observations, plus some regularization (penalty) term. The observations were created from model integration (i.e. a twin type experiment). The variational assimilation process shows its ability to retrieve the perfect initial conditions.

We have carried out a sensitivity analysis and applied its theoretical results to the optimality system to derive the sensitivity of a suitable response function, and also ( having carried out a perturbation on the observations ), the sensitivity of the retrieved initial condition to the data. The use of simulated altimeter data shows that grid points in the computational domain in the proximity of satellite ground tracks are less sensitive to the noise of the data than grid points further away from the satellite ground tracks.

This paper shows clearly that in the context of variational data assimilation, the sensitivity of a response function cannot be obtained by the use of adjoint equations of the direct model, but rather by using the adjoint equations of the optimality system.

In order to compute the sensitivity of the retrieved initial condition to the data,

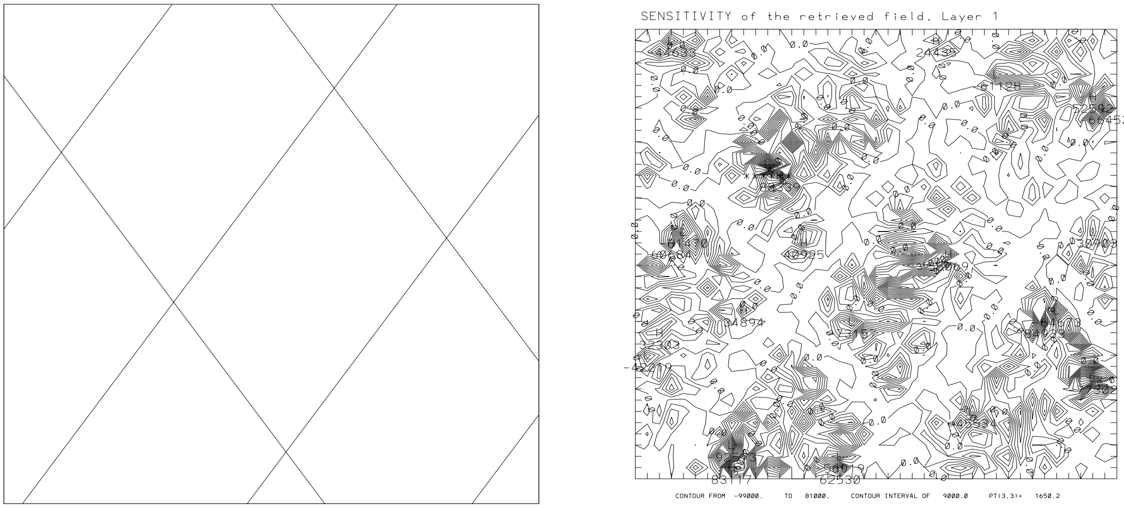


Figure 8: Ground tracks at the initial day and the sensitivity of the initial condition.

we use a conjugate gradient algorithm for the large-scale unconstrained minimization of a quadratic functional whose matrix is the hessian of the cost function with respect to the control variables. This Hessian matrix is ill-conditioned in the case of the model used in this study and so, the convergence rate of the minimization is very slow unless a regularization (penalty) term is included in the cost function and convexifies it.

# References

[1] X.L. Zou, I.M.Navon, F.X. Le Dimet: *Incomplete observations and control of gravity waves in variational data assimilation. Tellus,* **44A**, *273-296. 1992*

[2] Z. Wang , I.M. Navon, F.X. Le Dimet and X.L. Zou : *The second order adjoint analysis: theory and applications. Meteorology and Atmospheric Physics,***50**, *3-20 (1992)*

[3] D.G. Cacuci *Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses. J. Math. Phy.* **22***(12),2803-2812. 1981*

[4] F.X. Le Dimet *A general formalism of variational analysis. CIMMS report, Norman. OK 73091. 22. 1-34pp 1986*

[5] F.X. Le Dimet, O. Talagrand *Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. Tellus* **38A**, *97-110. 1986*

[6] F.X. Le Dimet, A. Nouailler *In Sasaki,Y.K. (ed.) Assimilation of Dynamics in Geosciences. Amsterdam: Elsevier, 1986, 181-198.*

[7] P. Courtier, O. Talagrand *Variational assimilation of meteorological observations with the adjoint equations Part 2. Numerical results.Quart.J. Roy. Meteor. Soc.* **113**, *1329-1347, 1987*

[8] D.G. Cacuci, M.C.G. Hall, M.E. Schlesinger *Sensitivity analysis of a radiative-convective model by the adjoint method. J. Atmos. Sci.,***39**, *2038-2050, 1982.*

[9] A.V. Fiacco, G.P. McCormick *Nonlinear programming: Sequential unconstrained minimization techniques, John Wiley and Sons, New York and Toronto, 210pp, 1968.*

[10] F.X. Le Dimet, *Determination o of the adjoint of a numerical weather prediction model, Tech. Rep. . FSU-SCRI-88-79, Florida State University, Tallahassee, Florida 32306-4052, 22pp. 1988.*

[11] G. Kontarev *The adjoint equation technique applied to meteorological problems, ECMWF Tech. Rep.. 21, Reading, U.K., 1-21. 1980.*

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

[13] J. Pedlosky : *The equation for geostrophic motion in the ocean. (1984), J. Phys. Oceanogr. N<sup>o</sup>14, (448-455).*

[14] J.Ch. Gilbert C.Lemarechal : *Some numerical experiments with variable storage quasi-Newton algorithms. (1989), Mathematical programming.*

[15] E. Blayo, J. Verron J.M. Molines : *Assimilation of Topex/Poseidon altimeter data into a circulation model of the N, Journal of Geophysical research, in press. (1994).*