Revista Brasileira de Meteorologia, **v.26**, **n.3**, 432 - 442, 2011

VARIATIONAL DATA ASSIMILATION: MATHEMATICAL FOUNDATIONS AND ONE APPLICATION

PAULO S. D. DA SILVA[1], GUDÉLIA MORALES[2], PRÍSCILA H. G. OLIVEIRA[2]

[1]Laboratório de Ciências Matemáticas, Universidade Estadual do Norte Fluminense, RJ, Brazil
[2]Laboratório de Engenharia de Produção, Universidade Estadual do Norte Fluminense, RJ, Brazil

paulosd@uenf.br, gudelia@uenf.br, phgoliveira@ibest.com.br

Received September – Accepted January 2011

ABSTRACT

Variational Data Assimilation and Adjoint Equation Method are presented here as a general methodology designed to improve the quality of computational simulation when are given the dynamics and a set of observation of the system under study. The mathematical foundations the procedures to obtain the adjoint of a given computational program, a fundamental task in order to apply the methodology, are carefully examined.
Keywords: Variational Data Assimilation, Adjoint Equations, Adjust of Trajectories.

RESUMO: ASSIMILAÇÃO VARIACIONAL DE DADOS: FUNDAMENTOS MATEMÁTICOS E UMA APLICAÇÃO
A Assimilação Variacional de Dados e o método das Equações Adjuntas são aqui apresentados como uma metodologia geral, desenvolvida para melhorar a qualidade das simulações computacionais, quando a dinâmica e um conjunto de observações do sistema em estudo são conhecidos. Os fundamentos matemáticos e os procedimentos para se obter o adjunto de um dado programa computacional, uma tarefa fundamental para a aplicação da metodologia, são cuidadosamente examinados.
Palavras-chave: Assimilação Variacional de Dados, Equações Adjuntas, Ajuste de Trajetórias

1. INTRODUÇÃO

In the second half of the eighties, in the confluence of the spreading in West Europe of of the work made by G. I. Marchuk on Adjoint Equation Method applied to Geophysical Flux Dynamics issues, and under the strong influence of the results obtained by J. L. Lions as to systematic use of the Optimal Control Theory to systems governed by partial differential equations, in particular, in Atmospheric Flows (Talagrand, 1997), they appeared numerical methods which would become the Variational Data Assimilation operational in less than 10years.

Sensitivity Analysis, through the adjoint formalism, quickly became useful in a wide variety of research areas (Thomas et al, 2002 , 2005, Kamien and Schwartz, 1981, Reuther et al, 1996). Nevertheless, only in recent years, the Data Assimilation technique has been exploited in computational simulations not related to Numerical Weather Prediction (Elbern et al., 2007), although the method can be invaluable in computational simulations, since observations of the systems act on the numerical model as restriction which the solution must verify. One of the possible explanations for this situation might be that the high complexity of the numerical experiments of atmospheric and oceanic flows and the many peculiarities of those flows make the experiments rather accessible to the computational simulation community. Besides, with the exception of a few articles (Le Dimet and Talagrand, 1986) and books (Kalnay, 2002, Daley, 1993), the mathematical formalism is not always clearly presented. This situation is not appropriated to the development of the Variational Data Assimilation method, since important aspects for its future plans, for example, Adaptive Observations, must be placed in solid theoretical foundations.

Therefore, this work is intended to present a rigorous mathematical formulation not only for Variational Data Assimilation but also Adjoint Equation Method as well, and, as an algorithm to perform the variational assimilation of data was developed, which is also described here, both formulations can be perfectly understood.

In order to realize the numerical control of the space-time evolution of a system involving a geophysical flow it is necessary to know its initial condition, which is, the numerical counterpart of the real initial state or the initial configuration of the system under study. In the case of geophysical flows, the dimension and the geographic location of the spatial domain problems make the distribution of an observational network difficult, what causes a scarcity of data about the phenomenon studied or even turns the intial configuration unavailable for computational simulations (Talagrand, 1997). As in those flows are presented phenomena of difficult representation in a numerical model, and even in analytical models, being, in certain situations, processes highly sensitives to initial conditions variations (Lorenz, 1963), to establish an initial condition adequated for simulations, in this case, is of the paramount importance for the results.

For an adequate initial condition of the numerical experiment, one means that one which generates the numerical model states that fit, in accordance with the desired precision, the available observations. In order to decide whether or not an initial condition is appropriated to certain simulation, one should consider, in additon to the flow dynamics, the full set of observations of the system being studied. That is exactly what is performed by the methodology known as Data Assimilation (Le Dimet and Talagrand, 1986), approach in which all the knowledge on a system being analyzed, its dynamics and observations, is combined in order to produce an initial condition that minimize the error between the model generated configurations and the correspondent known configurations, in a sense to be precised later.

A problem involving geophysical flows in its discrete form produce sa very high number of variables, so that, straight away, Variational Data Assimilation is a metodology computationally intractable. A solution for this difficulty can be obtained with the use of the Adjoint Equation Method, which adequately insert the problem of the initial condition determination to simulations of geophysical flows in the domain of Optimal Control Theory (Lions, 1971), and its mathematical foundations are analyzed here with the necessary accuracy. A second consequence of the resulting metodology of combining Variational Data Assimilation with Adjoint Equation Method is the transformation of a minimization problem with restrictions, as originally proposed by the Variational Data Assimilation formalism, in a minimization problem without, what makes it possible the use of more agile routines, as L-BFGS (Liu and Nocedal, 1989), in the determination of an optimal initial condition.

The efficency of that metodology, confirmed by numerical wheather prediction operational models (Parrish and Derber, 1992), has produced, on the one hand, a tremendous increase in the scope of its applications (Elbern et al, 2007), (Kamien and Schwartz, 1981), (Reuther et al, 1996), on the other hand a continous clarification of its theoretical basis, which is necessary for the improvement of the already existing models, and for the rise of others more precise.

In the first section of this work, the mathematical foundations of two crucial methods are presented, Variational Data Assimilation and Adjoint Equation Method. In the second section, in order to fix ideas, one works with a unidimensional model which permits the exploration of several aspects of the methods used and also to develop explicitly the adjoint of a given equation. Finally, in the third section, one describes a bidimensional numerical experiment, showing with some detail (i) an algebraic formalism, fundamental in the development of a computational program to perform the experiment, formalism that has already produced, in the Computational Linear Algebra domain, a new line of investigation, known asAutomatic Differatiation (Kaminski et al, 2003) and as Automatic Adjoint Generation (Giering et al, 2005), and (ii) the structure of a FORTRAN code, created by the authors, in order to realize the Variational Data Assimilation of a flow described by the bidimensional advection-diffusion equation.

2. MATHEMATICAL FORMULATION

Definition: Let A be an open set in $I\!R^n$ e $f{:}A \to I\!R$ a differentiable function. Then, $\forall x \in A$, the differential of f in x is the linear functional tal que, $\forall v \in I\!R^{\mathrm{n}}$,

$$df(x) \cdot v = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}.$$

As $df(x)$ is a linear functional in $(I\!R^{\,n})^{*}$ *, the dual vector space of $I\!R^n$, it follows form Linear Algebra the existence and the uniqueness of the $I\!R^n$ vector, $\nabla f(x)$, such that

$$df(x) \cdot v = \langle \nabla f(x), v \rangle \tag{1}$$

where <,> is the canonical inner product in $I\!R^n$.

2.1.The Data Assimilation Method

Let S be an evolutionary system, observed during the time interval $[t_1,t_2]$, from which is available a set of observations, all of them colected in the same time interval and distributed in the spatial domain of interest on S. Besides, it is known the dynamics of S. Based on those pieces of information, one wants to determine the initial condition of the modelled S, or the numerical representation of the configuration of S at the time t_1 , the starting moment for a computational simulation of S generates configurations to approach, within an acceptable precision, to the available observations of S at correspondent instants of time, in order to obtain a configuration of the model at $t_f > t_2$ which reproduces the real configuration of S at time t_f also within an acceptable pricision. From a conceptual point of view, this process is similar to the Best Linear Unbiased Estimation (Sorensen, 1980) of the configuration of S at t_f, given the S configurations within $[t_1,t_2]$, here solved in a deterministic approach as an optimal control problem, in which the initial conditions is the control data.

At first, one considers the continuous version of the Variational Data Assimilation problem, which admits the appropriate mathematical treatment. In this case, the mathematical objects are:

(i) the system observations: S, $Z:[t_1,t_2]\times V\rightarrow V$, such that, $\forall t\in[t_1,t_2]$, $Z_t:V\rightarrow V$ is a differential operator defined on the vector space V provided with the inner product <,>, where $[t_1,t_2]$ is the assimilation interval or assimilation window.

(ii) the dynamics of the system S, described by

$$\frac{\partial X}{\partial t}=F(X), \tag{2}$$

Where $X:[t_1,t_2]\times V\rightarrow V$ is the system S trajectory during the assimilation interval, $[t_1,t_2]$, $E=\{Y;Y:[t_1,t_2]\times V\rightarrow V$ and $Y\in C2\}$ and $F:E\rightarrow E$ is a differential operator.

(iii) the "weight" function $W:[t_1,t_2]\rightarrow L(V)$, where $L(V)$ is the vector space of the linear operators in V, that results from the statistics information of the instruments used to collect the data on S, and that, each $t\in[t_1,t_2]$, associate the injective linear operator W(t): $V\rightarrow V$

(iv) the quadratic functional

$$\begin{aligned}&J:E\rightarrow I\!R\\&X\mapsto\frac{1}{2}\int_D<W(t)(X(t,\vec{x})-Z(t,\vec{x})),W(t)(X(t,\vec{x})-Z(t,\vec{x}))>dD\end{aligned} \tag{3}$$

where $D=\{(t,\vec{x});t\in[t_1,t_2],\vec{x}\in V\}$.

Then the Variational Data Assimilation problem is to find a trajectory of the system states S, $X:[t_1,t_2]\times V\rightarrow V$ such that X be the solution of Equation 2 which minimize the functional Equation 3, in other words, the following minimization problem with constrait:

Problem M_R: find the solution of Equation 2 minimizing Equation 3

As the solution of the Data Assimilation will be numerically obtained, it is necessary to solve its discrete version. However, this implies another difficulty: the number of resulting variables of discrete version of the equation modelling the S dynamics and its space-time domain is excessively high. Therefore, from the computational point of view, it would be very economic (and, in many problems, the only feasible way) if instead of the solution $X\in E$ one searches the solution $X_1\in E_1$, where $E_1=\{X_1; E_1=\{X_1;X_1=X|_{\{t_1\}\times V}, X\in E$ and $X\in C^2\}$ isomorph to the set of operators in V, which could be called the space of the initial configurations of S, that minimizes the restriction of the functional J to E_1. The difficulty here is the non existence of a relation between J and X_1, because J is not a function of X. However, if the problem in Equation 2 is well posed, then the knowledge of X and the knowledge X_1 are equivalent. The Adjoint Equation Method provides, from the relation between X and X_1, the differential of the restriction of J to E_1.

2.2. The Adjoint Equation Method

One rewrites the functional in Equation 3 as

$$\begin{aligned}&J:E\rightarrow I\!R\\&X\mapsto\int_D T(X(t,\vec{x}))dD\end{aligned} \tag{4}$$

where

$$\begin{aligned}&T:E\rightarrow I\!R\\&X(\cdot,\cdot)\mapsto\frac{1}{2}<W(\cdot)(X(\cdot,\cdot)-Z(\cdot,\cdot)),W(\cdot)(X(\cdot,\cdot)-Z(\cdot,\cdot))>\end{aligned}$$

As the differential of J in $X\in E$ is a linear functional in the Hilbert space E with the inner product

$$<X(\cdot,\cdot),Y(\cdot,\cdot)>_E=\int_D<W(\cdot)(X(\cdot,\cdot)),W(\cdot)(Y(\cdot,\cdot))>dD$$

by using the Riesz' Representation Theorem (Kreyszig, 1989), one obtains, $\forall H\in E$,

$$dJ(X)\cdot H=<\nabla_X J(X),H>_E\int_D<\nabla_X T(X(t,\vec{x})),H(t)>dD, \tag{5}$$

where $\nabla_X J$ e $\nabla_X T$, are the gradient of J and of T, respectively

Now, one considers the linear version of the Equation 2:

$$\frac{\partial X}{\partial t}-\frac{\partial F}{\partial X}X=0 \tag{6}$$

resulting from the substitution of the operator F for its first order approximation in an equilibrium point, ommited from the equation, and in which it was used the notation $(\frac{\partial F}{\partial X})$ for its differential, and its adjoint equation:

$$\frac{\partial X}{\partial t}+(\frac{\partial F}{\partial X})^*X+\nabla_X T=0 \tag{7}$$

where $(\frac{\partial F}{\partial X})^*$ is the adjoint of the operator $(\frac{\partial F}{\partial X})$. Then one obtains the following result:

<u>Theorem 1</u>: Given a solution $X\in E$ of M_R, $X_1\in E_1$, $X_1=X|_{\{t_1\}\times V}$, is the solution of the following unconstraint minimization problem:

Problem M_S: find $\delta X\in E_1$ minimizing $<\nabla_{X_1}J,\delta X>$.

<u>Proof</u>: Let Y: $[t_1, t_2]\times V\rightarrow V$ be the solution of (7) such that $Y(t_2)=0$ and $\delta X\in E_1$ a perturbation of any solution of Equation 6. Then it follows that:

$$< \frac{\partial X}{\partial t} - \frac{\partial F}{\partial X}\delta X, Y > + < \frac{\partial Y}{\partial t} + (\frac{\partial F}{\partial X})^{*} Y + \nabla_X T, \delta X > = 0$$

$$\frac{\partial}{\partial t} < Y, \delta X > + < \nabla_X T, \delta X > = 0$$

$$\therefore \frac{\partial}{\partial t} < Y, \delta X > = - < \nabla_X T, \delta X >$$

$$< Y, \delta X > \Big|_{t_1}^{t_2} = - \int_{t_1}^{t_2} < \nabla_X T, \delta X > dt = - < \nabla_X J, \delta X >$$

$$< Y_1, \ \delta X_1 > \ = \ <, \delta X > \tag{8}$$

The meaning of Equation 8 must be clear: the right side of the equality is the expression of the differential of J (relative to X), $dJ(X)$. As X_1 is the projection of X onto E_1, the left side of Equation 8 is the expression of the differential of the restriction of J (relative to X_1) to E_1. By the uniqueness of the representation of the last differential as an inner product in E_1, it follows that Y_1 is precisely the gradient of the restriction of the funtional J to E_1, that is, Y_1 is the projection of $\nabla_X J$ onto E_1. Then, one obtains

$$Y_1 = \nabla_{X_1} J. \tag{9}$$

Therefore, using the Adjoint Equation Method, one projects the solution set $C^2([t_1,t_2] \times V, V)$ onto $C^2(\{t_1\} \times V, V)$, which represents, from the computational viewpoint,a considerable reduction in the number of variables in the discrete problem. Besides, this projection transforms a constraint minimization problem in one unconstraint problem, namely, the problem of minimizing Equation 3 with the restriction Equation 2, becomes

$$\underset{\delta X \in X_1}{Min} < \nabla_{X_1} J, \delta X > \tag{10}$$

The problem M_R, involving the quadratic functional J, defined in Equation 3, has the existence of its solution guaranteed but not its uniqueness. In order to obtain the uniqueness of the solution of M_R, and of M_S, one redefines the problem using the following quadratic functional:

$$\begin{aligned} J' &: E \to I\!R \\ X &\mapsto J_B(X) + J(X) \end{aligned} \text{, (3')}$$

where, X_b is an available estimate of the initial configuration of X, and $B \in L(V)$ is the operator determined by the available statistical information on the estimate of X_b and

$$\begin{aligned} J_B &: E \to I\!R \\ X &\mapsto \frac{1}{2} \int_G < B(X_1 - X_B), B(X_1 - X_B) >_{E_1} dx \end{aligned}$$

Then, one obtains the following result:

Theorem 2: Given the Problem $M_{R'}$: find the solution of Equation 2 that minimizes Equation 3'.

There exists a unique solution of $M_{R'}$ and $X_1 = X|_{E_1} \in E_1$ is the solution of the following unconstrait minimization problem:

Problem $M_{S'}$: find $\delta X_1 \in E_1$ that minimizes

$$< \nabla_{X_1} J', \delta X_1 >_{E_1}$$

Proof: It is only necessary to redefine the inner product on E as:

$$< X(\cdot,\cdot), Y(\cdot,\cdot) >_E = \int_D \begin{matrix} < B(X(t_1,\cdot)), B(Y(t_1,\cdot)) >_{E_1} + \\ < W(\cdot)(X(\cdot,\cdot)), W(\cdot)(Y(\cdot,\cdot)) > dD \end{matrix}$$

and use all the development that preceeds Theorem 1.

Let X_b be the available estimate of the initial configuration of X. The problem of optimal solution will be iteratively solved by shearching the perturbation $\delta X_1 \in E_1$ of $\delta X \in E$ such that ($X_b + \delta X$) will be the solution of $M_{S'}$.

3. ONE UNIDIMENSIONAL EXAMPLE:

Definição: Given $D = \{(x,t) \in I\!R^2 x_1 < x < x_2 \, e \, t_1 < t < t_2\}$ and $\tilde{A}$ the border of D, one defines:

- $D = \{(x,t) \in I\!R^2; x_1 < x < x_2 \, e \, t_1 < t < t_2\}$
- $\cup\{(x,t) \in I\!R^2; x = x_2 \ e \ t_1 < t < t_2\}$
- $\cup\{(x,t) \in I\!R^2; t = t_1 \ e \ x_1 < x < x_2\}$, the inside border, that is, the locus of the inside data of D
- $\Gamma_s = \{(x,t) \in I\!R^2; x = x_1 \ e \ t_1 < t < t_2\}$
- $\cup\{(x,t) \in I\!R^2; x = x_2 \ e \ t_1 < t < t_2\}$
- $\cup\{(x,t) \in I\!R^2; t = t_2 \ e \ x_1 < x < x_2\}$, the outside border, that is, the locus of the outside data of D.
- $I\!E = C^2(D, I\!R)$, the set of functions $f: D \to IR$ twice derivable.
- $\langle u,v \rangle_{I\!E} = \int u(x,t)v(x,t)dD$, the inner product on
- $I\!E_0 = \{u|^D_{\Gamma_e}; u \in I\!E\}$, the set of functions of $I\!E$ restricted to the inside border of D.
- $\langle u_e, v_e \rangle_{I\!E_0} = \int_{\tilde{A}_e} u(x,t)v(x,t)d\tilde{A}_e$, the inner product on $I\!E_0$

One considers the following situation: a fluid (water) flowing in a river or a channel of lenght $l = x_2 - x_1$, where the window of assimilation is $[t_1,t_2]$. A certain amount of pollutant is spilled in the channel in $x<x_1$ in the instant t_1. By hypothesis, the transport of the pollutant is described by the advection-diffusion unidimensional equation.

$$\frac{\partial c}{\partial t} + a\frac{\partial c}{\partial x} - k\frac{\partial^2 c}{\partial x^2} = 0$$

with the initial and contour conditions:

$$\begin{cases} c(x,t_1) = f(x), \ x_1 < x < x_2 \\ c(x_1,t) = g(t), \ t_1 < t < t_2 \\ c(x_2,t) = h(t), \ t_1 < t < t_2 \end{cases} \tag{11}$$

where $\begin{matrix} a \ D \ \to \ I\!R \\ (x,t) \mapsto a(x,t) \end{matrix}$ is the water velocity, k is the diffusivity coefficient and $\begin{array}{l} c:D \to I\!R \\ (x,t) \to c(x,t) \end{array}$ is the pollutant concentration in D.

Let us supose the availability of a set of channel observations, $\tilde{c}D \to I\!R$, during the time interval $[t_1,t_2]$. The Data Assimilation problem consists in finding a solution of Equation 11 that minimizes the functional

$$\begin{aligned} &J : I\!E \to I\!R \\ &c \mapsto \frac{1}{2}\int_D (c-\tilde{c})^2 dD \end{aligned} \tag{12}$$

that is, one has the following problem:

$$\begin{aligned} &\underset{c \in I\!E}{Min}\, J = \frac{1}{2}\int_D (c-\tilde{c})^2 dD \\ &\frac{\partial c}{\partial t} + a\frac{\partial c}{\partial x} - k\frac{\partial^2 c}{\partial x^2} = 0 \end{aligned} \tag{13}$$

Obtaining the solution of Equation 13, one can determine the instant of time when the pollutant will reach the position $x > x_2$, and this solves the problem of the deterministic prediction of the pollutant trajectory along the channel. If one is interested just in the position x of the front of the oil stain in relation to a reference point in one of the the channel shore, the geometry of the problem is one-dimensional.

In the sets $\{(x,t) \in I\!R^2 x = x_1 \; and \; t_1 < t < t_2\}$ and $\{(x,t) \in I\!R^2 x = x_2 \; and \; t_1 < t < t_2\}$, one can determine the border conditions of the problem, while in $\{(x,t) \in I\!R^2 t = t_1 \; and \; x_1 < x < x_2\}$ one determines the initial conditions of the problem. In the set $\{(x,t) \in I\!R^2 t = t_2 \; and \; x_1 < x < x_2\}$ one establishes the behavior of the solution of (16) at instant t_2, the final instant of time of the assimilation window. Then, for all c and $\ddot{a} \in C^2(D,I\!R)$, where $\ddot{a}c$ is a perturbation of $c \in C^2(D,I\!R)$, let J be functional defined in Equation 10. By definition, one obtains:

$$\begin{aligned} dJ(c)\cdot\delta c &= \langle \nabla J(c), \delta c \rangle_{I\!E} = \lim_{t\to 0} \frac{J(c+t\delta c)-J(c)}{t} \\ &= \lim_{t\to 0}\frac{1}{2t}\int_D (2tc\delta c + t^2(\delta c)^2 - 2t\tilde{c}\,\delta c)dD \\ &= \int_D \lim_{t\to 0}\frac{1}{2t}(2t\delta c(c-\tilde{c}) + t^2(\delta c)^2)dD \\ &= \int_D \delta c(c-\tilde{c})dD \end{aligned} \tag{14}$$

Therefore, again by using the Riesz' Representation Theorem

$$\nabla_{I\!E} J(c) = c - \tilde{c} \tag{15}$$

The gradient of the restriction of the functional J to the inside border of $I\!E$, $\tilde{A}_e$, $J_e = J\big|_{\tilde{A}_e}$, is given by one expression of the form:

$$dJ_e(u_e)\delta u_e = \langle \nabla J_e(u_e), \delta u_e \rangle_{I\!E_0} = \int_{\Gamma_e} \nabla J(u_e)\delta u_e d\Gamma_e \tag{16}$$

One shows now that the process by which it is obtained Equation 16 starting in Equation 14 is exactly the integration by parts method, with the border conditions appropriate for the current purpose.

Usiyng in (14) the fact that $\nabla J(c) = c - \tilde{c}$, one obtains one solution c of Equation 11 and a perturbation $\ddot{a}$ of c, which:

$$dJ(c)\cdot\delta c = \int_D (c-\tilde{c})(x,t)\delta c(x,t)dD$$

Defining

$$\begin{aligned} &L : C^2(D,I\!R) \to C(D,I\!R) \\ &c \mapsto \frac{\partial c}{\partial t} + a\frac{\partial c}{\partial x} - k\frac{\partial^2 c}{\partial x^2} \end{aligned}$$

for all $\ddot{a} \in I\!E$ perturbation of $c \in I\!E$ and such that $L(c) = 0$,

$$L(\delta c) = \frac{\partial \delta c}{\partial t} + a\frac{\partial \delta c}{\partial x} - k\frac{\partial^2 \delta c}{\partial x^2} = 0$$

Let $\delta' c \in I\!E$ be. Then,

$$\begin{aligned} &dJ(c)\cdot\delta c = \int_D (c-\tilde{c})(x,t)\delta c(x,t)d\mathrm{D} = \\ &\int_D [(c-\tilde{c})\delta c - (\frac{\partial \delta c}{\partial t} + a\frac{\partial \delta c}{\partial x} - k\frac{\partial^2 \delta c}{\partial x^2})\delta' c]d\mathrm{D} \therefore \\ &\therefore dJ(c)\cdot\delta c = \int_{t_1}^{t_2}\int_{x_1}^{x_2} (c-\tilde{c})\delta c dxdt - \\ &-\int_{t_1}^{t_2}\int_{x_1}^{x_2} (\frac{\partial \delta c}{\partial t} + a\frac{\partial \delta c}{\partial x} - k\frac{\partial^2 \delta c}{\partial x^2})\delta' c dxdt \end{aligned} \tag{17}$$

Integring by parts the terms of the second item of Equation 17, one obtains

$$\begin{aligned} &\int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial \delta c}{\partial t}\delta' c \mathrm{dxdt} = \int_{x_1}^{x_2} \delta c \delta' c\Big|_{t_1}^{t_2} \mathrm{dx} - \int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial \delta' c}{\partial t}\delta c dxdt \\ &a\int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial \delta c}{\partial x}\delta' c \mathrm{dxdt} = a\int_{t_1}^{t_2} \delta c \delta' c\Big|_{x_1}^{x_2} \mathrm{dt} - a\int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial \delta' c}{\partial x}\delta c dxdt \\ &-k\int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial^2 \delta c}{\partial x^2}\delta' c \mathrm{dxdt} = -k\int_{t_1}^{t_2} \frac{\partial \delta c}{\partial x}\delta' c\Big|_{x_1}^{x_2} \mathrm{dt} + \\ &+k\int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial \delta c}{\partial x}\frac{\partial \delta' c}{\partial x} dxdt = \\ &= -k\int_{t_1}^{t_2} \frac{\partial \delta c}{\partial x}\delta' c\Big|_{x_1}^{x_2} \mathrm{dt} + k\int_{t_1}^{t_2} \frac{\partial \delta' c}{\partial x}\delta c\Big|_{x_1}^{x_2} \mathrm{dt} - \\ &-k\int_{t_1}^{t_2}\int_{x_1}^{x_2} \frac{\partial^2 \delta' c}{\partial x^2}\delta c dxdt \end{aligned}$$

Therefore,

$$\begin{aligned} dJ(c)\cdot\delta c = &\int_{t_1}^{t_2}\int_{x_1}^{x_2} [(c-\tilde{c})\delta c dxdt - \int_{x_1}^{x_2} \delta c \delta' c\Big|_{t_1}^{t_2} dx - a\int_{t_1}^{t_2} \delta c \delta' c\Big|_{x_1}^{x_2} dt \\ &+k\int_{t_1}^{t_2} \frac{\partial \delta c}{\partial x}\delta' c\Big|_{x_1}^{x_2} dt - k\int_{t_1}^{t_2} \frac{\partial \delta' c}{\partial x}\delta c\Big|_{x_1}^{x_2} dt + \\ &+\int_{t_1}^{t_2}\int_{x_1}^{x_2} \delta c\left(\frac{\partial \delta' c}{\partial t} + a\frac{\partial \delta' c}{\partial x} + k\frac{\partial^2 \delta' c}{\partial x^2}\right)dxdt \end{aligned}$$

$$\therefore dJ(c)\cdot\delta c = -\int_{x_1}^{x_2} \delta c\delta' c\Big|_{t_1}^{t_2} dx - a\int_{t_1}^{t_2} \delta c\delta' c\Big|_{x_1}^{x_2} dt \ +$$

$$+k\int_{t_1}^{t_2} \frac{\partial \delta c}{\partial x}\delta' c\Big|_{x_1}^{x_2} dt - k\int_{t_1}^{t_2} \frac{\partial \delta' c}{\partial x}\delta c\Big|_{x_1}^{x_2} dt +$$

$$+\int_{t_1}^{t_2}\int_{x_1}^{x_2} \delta c\left(\frac{\partial \delta' c}{\partial t} + a\frac{\partial \delta' c}{\partial x} + k\frac{\partial^2 \delta' c}{\partial x^2} + (c-\tilde{c})\right)dxdt$$

Then, by taking $\delta' c$ as solution of the non homogeneous adjoint equation, $\frac{\partial \delta' c}{\partial t} + a\frac{\partial \delta' c}{\partial x} + k\frac{\partial^2 \delta' c}{\partial x^2} + (c-\tilde{c}) = 0$ the initial condition $\delta' c(t_2,x) = 0$, $\forall x \in [x_1,x_2]$, and the border conditions $\delta' c(t,x_1) = \delta' c(t,x_2) = 0$, $\forall t \in [t_1,t_2]$, one obtains:

$$dJ(c)\cdot\delta c = \int_{x_1}^{x_2} \delta c(t_1,x)\delta' c(t_1,x)dx +$$

$$+k\int_{t_1}^{t_2} \frac{\partial \delta' c(t,x_1)}{\partial x}\delta c(t,x_1)dt$$

$$-k\int_{t_1}^{t_2} \frac{\partial \delta' c(t,x_2)}{\partial x}\delta c(t,x_2)dt$$

$$dJ(c)\cdot\delta c =$$

$$= \left\langle \delta' c(t_1,x) + k\left(\frac{\partial \delta' c(t,x_1)}{\partial x} - \frac{\partial \delta' c(t,x_2)}{\partial x}\right), \delta c(t,x) \right\rangle_{I\!E_0} \quad (18)$$

So, the gradient de J with respect to initial and border conditions, in other words, restrict to the space $\tilde{A}_e$, is:

$$\nabla_{I\!E_0} J = \delta' c(t_1,x) + k\left(\frac{\partial \delta' c(t,x_1)}{\partial x} - \frac{\partial \delta' c(t,x_2)}{\partial x}\right) \quad (19)$$

Therefore, in the iterative resolution, by means of a minimization routine, using the gradient of the functionl J, of problem in Equation 13,, one considers the problem:

$$\underset{\delta' c \in I\!E_0}{Min} < \nabla_{I\!E_0} J, \delta' c >$$

what represents a considerable gain in the necessary computations.

4.Results

4.1.Implementing the Adjoint Equation Method

Another interesting and also important feature of the implementation of Variational Data Assimilation is the development of automatic routines to perform the data assimilation using the Adjoint Equation Method. Nowadays, due to the several academic and industrial application of the method and considering the extension of the work needed to produce such computational routine, the automatic generation of the adjoint to a given code can be realized by computational codes developed for this task (Faure and Papegay, 1998). As a good knowledge of the rules used in the preparation of the linear and the adjoint programs to a given computational code precede the correct use of such automatic routines, it is presented here some illustrative examples of the rules to be followed when manually generating these codes. Of course, the rules are the same used by automatic routines.

A computational code in, *exempli gratia*, FORTRAN language is nothing but an ordered sequence of command lines and sometimes one or more subroutines or functions, which are also made of an ordered sequence of command lines. One begins defining a rule to obtain the adjoint instruction to a given command line, and, in the sequel, it is derived the adjoint instruction to the instruction resulting of two consecutive command lines. Let one consider the following FORTRAN command line,

*Z = 2*X + Y**3*,

which can be identified with the fuctio

$$f : I\!R^2 \to I\!R,\quad (x,y) \mapsto 2x + y^3 \text{, where } z = f(x,y), \quad (20)$$

or even with the function

$$g : I\!R^3 \to I\!R^3,\quad (x,y,z) \mapsto (x,y,2x+y^3) \text{, where } z = 2x + y^3, \quad (21)$$

a very useful process in order to clarify the many stages in the work out of the adjoint instruction and its implementation, decreasing the possibility of errors when writing the computational code.

Although Equation 20 and Equation 21 are mathematically equivalents, notation in Equation 21 is much more convenient to write the computational codes.

As g is a nonlinear operator and nonlinear operators do not posses adjoints, when a command line is equivalent to one nonlinear operation, the first action to do is to take the linear version of the nonlinear. This can be achieved simply taking the differential of the operator, or in matricial notation, much more useful for the adjoint process, considering the jacobian matrix of the operator. In the case of function g, its jacobian is given by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 3y^2 & 0 \end{pmatrix} \quad (22)$$

Then the linear of the instruction represented by g, dg, is:

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_o = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 3y_i^2 & 0 \end{pmatrix}\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_i \quad (23)$$

where the subscripts o and i denote, respectivelly, outside and inside amounts stored in the variables, just after and before the execution of the command line, that is,

$$\begin{cases} dx_o = dx_i \\ dy_o = dy_i \\ dz_o = 2dx_i + 3y_i^2 dy_i \end{cases} \tag{24}$$

in this way, the adjoint instruction is obtained by the matrix in Equation 21:

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_s = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3y_e^2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}_e \tag{25}$$

From operation in Equation 25 one obtens the adjoint instruction to the instruction in expression (19):

$$\begin{cases} dx_o = dx_i + 2dz_i \\ dy_o = dy_e + 3y_e^2 dz_e \\ dz_o = 0 \end{cases} \tag{26}$$

By associating each instruction line of one code to a funciton, therefore, this means that a program is a composition of functions. So, the linear version of the given program is obtained by using the Chain Rule of the Calculus, while the adjoint of a given code, by multipliyng the reverse order the adjoint matrices that correspond to each command line. The processes, although simple, demands careful, since one outside variável in a command line can or not be defined using the same variable (in this case, seem as na inside variable), as it will be considered later.

Let the instructions be:

$Z = X + 2*Y$

$W = Y + Z**2,$

Which will be refered as i_1 and i_2 respectively.

The following functions are associated with the previous instructions:

$$f_1 : \mathbb{R}^4 \to \mathbb{R}^4 \quad (x,y,z,w) \mapsto (x,y,x+2y,w) \quad \text{and} \quad f_2 : \mathbb{R}^4 \to \mathbb{R}^4 \quad (x,y,z,w) \mapsto (x,y,z,y+z^2)$$

To make clearer the processes of taking the linear and the adjoint of a given instruction, as well as the writting of the correspondent instructions in FORTRAN, one adopts the following convention: $\tilde{f}_1 E_e \to E_i$ e $\tilde{f}_2 E_i \to E_s$, where E_e , $E_{in} = \tilde{f}_1(E_i)$ and $E_s = \tilde{f}_2(E_{in})$ are, respectivelly, the inside space, the intermediate space and the outside space, all isomorphs to $\mathbb{R}^4$. Then

$$\tilde{f}_1 : E_e \to E_i \quad (x,y,z,w) \mapsto (x,y,x+2y,w)$$

and

$$\tilde{f}_2 : E_i \to E_s \quad (x,y,x+2y,w) \mapsto (x,y,x+2y,y+(x+2y)^2)$$

where , $\tilde{f}_2 : E_i \to E_s$, $(x,y,x+2y,w) \mapsto (x,y,x+2y,y+(x+2y)^2)$

$v_I = (x,y,z,w)$ is a vector of inside data and $v_O = \tilde{f}(v_e)$, is a vector of outside data, resulting of the action of the instructions in the vector of inside data v_i. Given a perturbation d_i of the vector of inside data, one obtains the correspondent perturbation d_o in the vector of outside data, in order words, $d\tilde{f}(v)\, d_i = d_o$, which, by the Chain Rule, it is expressed by $d_o = d\tilde{f}_2(v_{in}) \cdot d\tilde{f}_1(v_i) d_i$, where $v_{in} = \tilde{f}_1(v_i)$ is the vector of intermediate data.

Computing $d\tilde{f}_2(v_{in})$ and $d\tilde{f}_1(v_i)$, in the case $v_i = (x,y,z,w) \in E_e$ e $v_{in} = (x,y,x+2y,w) \in E_{in}$, tem-se:

$$d\tilde{f}_2(v_{in}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 2(x+2y) & 0 \end{pmatrix} \quad \text{e} \quad d\tilde{f}_1(v_i) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Therefore, given any direction, $d_i = (dx_i, dy_i, dz_i, dw_i) \in E_e$, one obtains:

$$(dx_s, dy_s, dz_s, dw_s) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 2(x+2y) & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dx_e \\ dy_e \\ dz_e \\ dw_e \end{pmatrix}$$

that is:

$$\begin{cases} dx_s = dx_e \\ dy_s = dy_e \\ dz_s = dx_e + 2dy_e \\ dw_s = (2x_e + 4y_e)dx_e + (4x_e + 8y_e + 1)dy_e \end{cases}$$

Leaving the inside and outside índexes (which are not written in a real code) and discarding the unchanged variables, as $dx = dx$ and $dy = dy$, the linear instruction of instructions i_1 and i_2 above are the following FORTRAN instructions:

$$dZ = dX + 2*dY$$
$$dW = (2*X + 4*Y)*dX + (4*X + 8*Y + 1)*dY$$

as it was expected.

Considering that $\tilde{f}_1 E_e \to E_i$ and $\tilde{f}_2 E_i \to E_s$, the adjoint to the instructions i_1 and i_2 are $\tilde{f}_1^* E_i \to E_e$ and ,

$$\tilde{f}_1^* E_i \to E_e$$

from it follows that $[d\tilde{f}(v_s)]^*(d_s) = [d\tilde{f}_1(v_i)]^* \circ [d\tilde{f}_2(v_s)]^*(d_s) = d_e$, where:

$$[d\tilde{f}_1(v_i)]^* = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad [d\tilde{f}_2(v_s)]^* = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2z_s \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Then, given $d_s = (dx_s, dy_s, dz_s, dw_s) \in E_s$, one obtains:

$$(dx_e, dy_e, dz_e, dw_e) = [d\tilde{f}(v_s)]^*(d_s) =$$

$$\begin{pmatrix} 1 & 0 & 1 & 2z_s \\ 0 & 1 & 2 & 4z_s \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} dx_s \\ dy_s \\ dz_s \\ dw_s \end{pmatrix} = \begin{pmatrix} dx_s + dz_s + 2zdw_s \\ dy_s + 2dz_s + 4zdw_s \\ 0 \\ 0 \end{pmatrix}$$

that is:

$$\begin{cases} dx_e = dx_s + dz_s + 2z_s dw_s \\ dy_e = dy_s + 2dz_s + 4z_s dw_s \end{cases}$$

Considering that $v_s = (x_e+2y_e)$ and leaving the índices, one obtains the adjoint instruction in FORTRAN as:

$$dX = dX + dZ + 2* \ (X + 2Y)dW$$
$$dY = dY + 2*dZ + 4* \ (X + 2Y)* dW$$

And now consider the case in which one of the variables appears in one command line as an outside variable and as an inside variable. So, let i_3 and i_4 be the instructions:

*Y = X*Y**2*

*Y = Y**3*X**2*

To attain the necessary clarity in the development of the adjoint code, one defines the functions associated with the above instructions using the already introduced notation of inside, intermediate and outside spaces, E_e, E_i and E_s, respectively, in this case, all isomorphs to $\mathbb{R}^2$:

$$\tilde{f}_1 : E_e \to E_i \quad (x, y) \mapsto (x, xy^2) \qquad \text{and} \qquad \tilde{f}_2 : E_i \to E_s \quad (x, xy^2) \mapsto (x, x^2(xy^2)^3)$$

such that, to the inside vector $v_e = (x,y)$, the two instructions produce the outside vector v_s given by:

$$\tilde{f} : E_e \to E_s$$
$$v_e \mapsto v_s = (\tilde{f}_2 \circ \tilde{f}_1)(v_e)$$

For the Chain Rule, the linear of the composition of the instructions is given by $d_s = d\tilde{f}_2(v_i) \cdot d\tilde{f}_1(v_e) d_e$, where d_e is a perturbation in the inside vector, d_s, and the correspondent perturbation in the outside vector is $v_i = \tilde{f}_1 \ (v_e) = (x,xy^2)$.

Therefore, given $d_e = (dx_e, dy_e)$, one obtains

$$d\tilde{f}_2(v_i) = \begin{pmatrix} 1 & 0 \\ 2x(xy^2)^3 & 3x^2(xy^2)^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2x^4y^6 & 3x^4y^4 \end{pmatrix}$$

and

$$d\tilde{f}_1(v_e) = \begin{pmatrix} 1 & 0 \\ y^2 & 2xy \end{pmatrix}$$

Then, $d_s = (dx_s, dy_s)$ is given by

$$(dx_s, dy_s) = \begin{pmatrix} 1 & 0 \\ 2x^4y^6 & 3x^4y^4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y^2 & 2xy \end{pmatrix} \begin{pmatrix} dx_e \\ dy_e \end{pmatrix}$$

that is

$$\begin{cases} dx_s = 1dx_e \\ dy_s = 5x^4y^6 dx_e + 6x^5y^5 dy_e \end{cases}$$

or

$$dy = 5x^4y^6dx + 6x^5y^5dy$$

So, in FORTRAN, the linear of the commands i_3 and i_4 is given by:

dY = Y**2*dX + 2*X*Y*dY
Y = X*Y**2
dY = 2*X*Y**3*dX + 3*Y**2*X**2*dY
Y = Y**3*X**2

and the adjoint of the instructions is obtained from ,

$$d_e = [d\tilde{f}(v_e)]^*(d_s) = [d\tilde{f}_1(v_e)]^* \cdot [d\tilde{f}_2(v_i)]^*(d_s)$$

where d_e and d_s, respectively, are perturbations in the inside and outside vectors, tha is:

$$(dx_e, dy_e) = \begin{pmatrix} 1 & y^2 \\ 0 & 2xy \end{pmatrix} \begin{pmatrix} 1 & 2x(xy^2)^3 \\ 0 & 3(xy^2)^2x^2 \end{pmatrix} \begin{pmatrix} dx_s \\ dy_s \end{pmatrix} \therefore$$

$$\therefore (dx_e, dy_e) = \begin{pmatrix} 1 & 5x^4y^6 \\ 0 & 3x^4y^6 \end{pmatrix} \begin{pmatrix} dx_s \\ dy_s \end{pmatrix} \therefore$$

$$\therefore \begin{cases} dx_e = dx_s + 5x^4y^6 dy_s \\ dy_e = 6x^5y^5 dy_s \end{cases}$$

Therefore, in FORTRAN, the adjoint instructions to i_3 and i_4 are given by:

I_0 = Y
*Y = X*Y**2*
I_1 = Y
*Y = Y**3*X**2*

Y = I_1
*dX= dX + 2*X*Y**3*dY*
*dY = 3*Y**2*X**2*dY*

Y = I_0
*dX = dX + Y**2*dY*
*dY = 2**Y*X*dY*

4.2 Numerical experiment

One studied the oil stain trajectory determination in a region whose hydrodynamics were known and a set of oil spot observation within the time interval $[t_1,t_2]$ was available. It was also supposed that all observations were incomplete, so that, at any time $t_1 < t < t_2$, the avalilable observation $Z(t)$ can not reproduce integrally the oil spot, and, therefore, could not be used as an "initial condition" in the integration of the numerical model.

The first step of the experiment was to simulate an oil spillage in spatial domain of 6000m × 4400m, discretized by means of bidimensional grid of 21 × 21 nós, where Δ x = 300m, Δ y = 220m and the time interval of analysis was [0,30000] in seconds, with Δ t = 300s. It was taken the velocity of the medium as u = 0.1 × 10^{-1} m/s^{-1} and v = 0.0, C was the concentration of oil, and D was the variable coefficient of diffusivite. The transport of the oil spot was described by the bidimensional advection-diffusion

$$\frac{\partial C}{\partial t}+u\frac{\partial C}{\partial x}+v\frac{\partial C}{\partial y}-D(\frac{\partial^2 C}{\partial x^2}+\frac{\partial^2 C}{\partial y^2})=0, \qquad (27)$$

which is the universal transport model of oil in the sea surface (Lehr and Cekirge, 1980). In Figure 1, one sees the flow chart to the Data Assimilation using the Adjoint Equation Method.

One has integrated then the Equation 27 using as initial condition the artificial oil spill (Figure 2) and, with the integration of Equation 27 (in G(U)), were storage, as system observations, the configurations obtained at the times t = 10 × i, i = 1,...9. In Figures 3 and 4, are shown the partial configurations of the oil spot for t = Δ t × 20 and t = Δ t × 90. One aletory perturbed the initial configuration, shown in Figure 2, by using the RAND function from the FORTRAN 90 Library, and it was obtained the perturbed initial configuration (Figure 5) , which was the begining of the Variational Data Assimilation process , and after the first integration of Equation 27, in order to generate the (toy) observations, the Equation 27 is then integrated with an initial condition obtained from the minimization routine, and the error between the model generated configurations and the observations are measured in J(G(U)). If the value measured is above the tolerance, the main program calls the subroutine D(G(U))*, in order to obtain the integration backward in time of the adjoint Equation 7, which gives the gradient of the functional J in relation to the initial conditions of the problem, which is then sent to the minimization subroutine (L-BFGS), giving a new initial configuration U which, after integration in G(U), will decrease the value of J(G(U)). When the error generated by the two configurations, model generated and observations, measured in J(G(U)), is bellow the tolerance, the process of Data Assimilation stopped, producing the optimal initial condition (Figura 6).

In the considered example, the optimal initial condition (Figura 3) was obtained, recovering the initial condition of the problem with an error less than 2%.

5.CONCLUSIONS

Data Assimilation, process in which observations of the studied systems are considered as conditions to the computational simulation must verify, together with Adjoint Equation Method, produce a methodology to obtain reliable and highly accurate computational configurations of a system

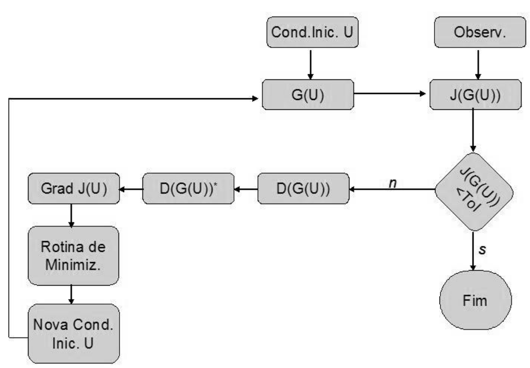

Figure 1 - General flux

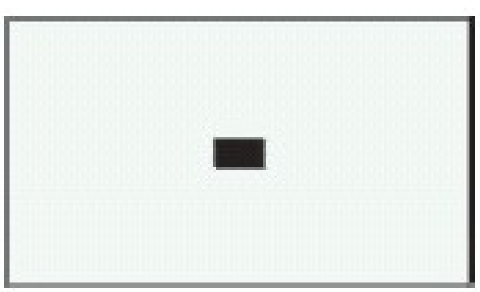

Figure 2 - Inital configuration

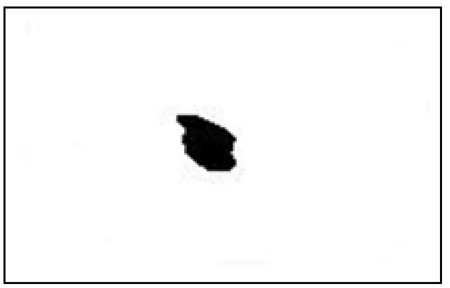

Figure 3 - Configuratio for t = Δt x 20

Figure 4 - Configuration for t = Δt x 90

Figure 5 - Random perturbation

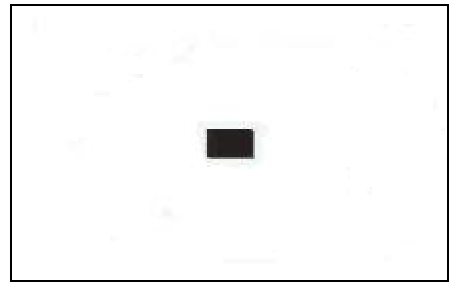

Figure 6 - Initial configuration Incial

at a feasible computational cost and within a time interval that allows na effective decision in an oil spill.

A further question, also in the realm of the formalism presented here, is the optimal location of measurement instruments, enhancing an existing network.

6. REFERENCES

DALEY, R. **Atmospheric Data Analysis**, Cambridge University Press, 472, 1993.

ELBERN, H., STRUNK, A., SCHIMDT, H., TALAGRAND, O. Emission rate and chemical state estimation by 4-dimensional variational inversion, **Atmos. Chem. Phys**.7, 3749-69, 2007.

FAURE, C., PAPEGAY, Y. Odyssée User's Guide, Version 1.7, INRIA, 1998.

GIERING, R., KAMINSKI, T., TODLING, R., ERRICO, R., GELARO, R., WINSLOW, N. Generating tangent linear and adjoint versions of NASA/GMAO's Fortran-90 global weather forecast model. In H.M. Bücker, G. Corliss, P. Hovland, U. Naumann, and B.Norris, editors, Automatic Differentiation: Applications, Theory, and Implementations, **Lecture Notes in Computational Science and Engineering**, v. 50, pages 275-284. Springer, New York, NY, 2005.

KALNAY, E. **Atmospheric Modeling, Data Assimilation and Predictability**, Cambridge University Press, 364, 2002.

KAMIEN, M.I., SCHWARTZ, N.L. **Dynamic optimization: the calculus of variation and optimal control in economics and management**, North-Holland, 1981.

KAMINSKI,T., GIERING, R., SCHOLZE, M., RAYNER, P., KNORR, W. An example of an automatic differentiation-based on modelling system, Conference proceeding, Computational Science and Its Applications ICCSA 2003, **Proceedings of the International Conference on Computational Science and its Applications**, Montreal, Canada, May 18--21, 2003. Part II, Springer, 2003.

KREYSZIG, E.; **Introductory Functional Analysis with Applications**, John Wiley and Sons, 1989.

LE DIMET, F. X., TALAGRAND, O. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, **Tellus**, 38A, 1986.

LEHR, W.J., CEKIRGE, H. M. Oil slick movements in the Arabian Golf, **Proceeds of Petroleum and Marine Environment**, Monaco, 1980.

LIONS, J.L. **Optimal Control of Systems Governed by Partial Differential Equations**,
Springer - Verlag, 1971.

LIU, D., NOCEDAL, D. On the limited memory BFGS method for large scale minimization, **Mathematical Programming**, 45, pp.: 503-528, 1989.

LORENZ, E.N.; Deterministic non-periodic flows, Journal of Atmospheric Sciences, 20, pp.: 130-141, 1963.

PARRISH, D.F., DERBER., J.D. The national meteorological center spectral statistical interpolation analysis system**. Mon. Wea. Rev**., 120:1747–1763, 1992.

REUTHER, J., JAMESON, A., FARMER, J., MARTINELLI, L., SAUNDERS, D. Aerodynamic shape optimization of complex aircraft configurations via an adjoint formulation, **AIAA Journal** , 96-0094, 1996.

SORENSEN, H.W.; **Parameter estimation: principles and problems**, Marcel Dekker, 1980.

TALAGRAND, O.; Assimilation of observations, an introduction, **Journal of Meteorological Society of Japan**, v. 75, n. 1B, pp.: 191-209, 1997.

TALAGRAND; .A study of the dynamics of four-dimensional data assimilation, **Tellus**, 33, 43-60, 1986.

THOMAS, J. P., HALL, K. C. Sensitivity analysis of coupled aerodynamic/structural dynamic behavior of blade rows. **Extended Abstract for the 7th National Turbine Engine High Cycle Fatigue (HCF) Conference**, Palm Beach Gardens, Florida, 14-17 May 2002.

THOMAS, J. P., HALL, K. C., DOWELL, E.H. A discrete adjoint approach for modeling unsteady aerodynamic design sensitivities. **AIAA Journal**., 43(9):1931-1936, 2005.