International Journal of Pure and Applied Mathematics
Volume 69 No. 1 2011, 15-32

THE NEWTON METHOD IN PROBLEMS OF VARIATIONAL DATA ASSIMILATION: APPLICATION TO AN INFILTRATION MODEL

Pierre Ngnepieba[1 §], Desmond Stephens[2], Francois-Xavier Le Dimet[3]

[1,2]Department of Mathematics
Florida A&M University
Tallahassee, Florida, 32307, USA
[1]e-mail: Pierre.Ngnepieba@famu.edu
[2]e-mail: Desmond.Stephens@famu.edu

[3]Laboratoire Jean-Kuntzmann
University of Grenoble and INRIA
Grenoble Cedex, 38400, FRANCE
e-mail: Francois-Xavier.Ledimet@imag.fr

Abstract: The problem of four-dimensional variational data assimilation (4D-Var) seeks to find an optimal initial field minimizing a cost function defined as the squared distance between model solutions and observations within an assimilation window. It requires minimization algorithms along with adjoint models to compute gradient information needed for the minimization. In this paper, an alternative method is suggested based on the implementation of the Newton algorithm stemming from the optimality system. This new method solves the 4D-Var-minimization problem efficiently, and it is equivalent to the LBFGS algorithm using the exact Hessian of the cost function. An application to the data assimilation problem in hydrology is presented. Numerical results are discussed.

AMS Subject Classification: 65N30
Key Words: Newton's method, data assimilation, infiltration, optimality system, adjoint, Hessian, Richards equation

Received: October 18, 2010

§Correspondence author

1. Introduction

In data assimilation tasks, the adjoint model is used to compute (in reverse mode) the gradient of the cost function with respect to the control variable [4, 6]. The adjoint method eliminates the need for repeatedly computing the gradient of the cost function at each iteration of the 4D-Var minimization routine during a forward integration. Additionally, for large systems, the adjoint model reduces the number of iterations and the computed gradient is exact. Once the gradient is computed an iterative gradient descent algorithm (quasi-Newton type Gilbert and Lemaréchal [5], Lui and Nocedal [7]) is used to approximate the optimality condition. Optimization of the Numerical Weather Prediction (NWP) initial conditions can be regarded as a class of *inverse problem*. Indeed, we assume that the forward problem is solvable, i.e. given a set of initial and boundary conditions, a set of forward models (governing equations) can be run to predict the observations. This paper investigates an alternative for solving the data assimilation problem based on Newton's method applied to the optimality system. In this study, we focus the attention on the optimality condition only rather than the full optimality system as described by Ngnepieba et al in their seminar paper [10].

The first part of the paper is devoted to the theoretical results. Section 2 provides a formulation of the problem. In Sections 3-5 we present Newton's method and results on its solvability in a specific class of functional spaces. The second part of the paper will focus on an application modeling underground water. Using this model we consider four cases corresponding to different temporal distributions of data in a Grenoble sand [8, 9]. These cases correspond to observations distributed for all time steps, for all at 60 seconds, for all at 300 seconds, and finally for all at 720 seconds.

2. Formulation of the Problem

Consider the evolution problem

$$\begin{cases} \dfrac{dX}{dt} = F\left(X\right), \quad t \in (0,T), \\ X\left(0\right) = u, \end{cases} \tag{1}$$

where $X = X\left(t,u\right)$ is the unknown function belonging for all time $t \in [0,T]$ in a Hilbert space $\mathcal{X}$, $u \in \mathcal{X}$ is the control parameter and F is a non-linear operator such that $F:\ \mathcal{X} \longrightarrow \mathcal{X}$. We define $Y = L^2\left(0,T;\mathcal{X}\right),\ (x,y)_{L^2(0,T;\mathcal{X})} = (x,y)$

and $\|x\| = (x,x)^{\frac{1}{2}}$ as the scalar product and norm in the space Y. Let us introduce the cost function

$$J(u) = \frac{\alpha}{2}\|u - u^0\|_{\mathcal{X}}^2 + \frac{1}{2}\int_0^T \|C.X - X_{obs}\|_{Y_{obs}}^2 dt, \tag{2}$$

where $\alpha \geq 0$ is a constant, u^0 be an estimate of the initial condition u. Y_{obs} is a subspace of Y, C is a linear operator such that $C: Y \longrightarrow Y_{obs}$ and $X_{obs} \in Y_{obs}$ represents the vector of observation.

We consider the following data assimilation problem with the aim of identifying the initial condition, namely find X and u such that:

$$\begin{cases} \dfrac{dX}{dt} = F(X), \quad t \in (0,T), \\ X(0) = u, \\ J(u) = \operatorname*{Inf}_{\tau \in \mathcal{X}} J(\tau). \end{cases} \tag{3}$$

The necessary optimality condition (see [4, 6]) reduces (3) to the following system

$$\begin{cases} \dfrac{dX}{dt} = F(X), \quad t \in (0,T), \\ X(0) = u, \end{cases} \tag{4}$$

$$\begin{cases} \dfrac{dP}{dt} + \left[F'(X)\right]^T.P = C^T(C.X - X_{obs}), \quad t \in (0,T), \\ P(T) = 0, \end{cases} \tag{5}$$

$$\nabla J(u) = \alpha\left(u - u^0\right) - P(0) = 0, \tag{6}$$

where X, P, and u are unknowns.

Newton's algorithm will now be employed to find the solution to the system (4)-(6) under the assumption that such a solution exists.

3. Newton's Method

The system (4)-(6), with the three unknowns X, P and u, can be rewritten in terms of an operator $\mathcal{F}$ such that:

$$\mathcal{F}(U) = 0, \tag{7}$$

where $U = (X, P, u)$.

The implementation of Newton's method requires the calculation of the derivative $\mathcal{F}'(U)$ of the operator $\mathcal{F}$ at the point U.

Newton's algorithm to solve the problem is,

$$U_{n+1} = U_n - \left[\mathcal{F}'(U_n)\right]^{-1}.\mathcal{F}(U_n), \quad \text{where } U_n = (X_n, P_n, u_n)\,. \tag{8}$$

The algorithm can be decomposed into two steps:

1. Find $V_n = [\mathcal{F}'(U_n)]^{-1}.\mathcal{F}(U_n)$ as the solution of $[\mathcal{F}'(U_n)].V_n = \mathcal{F}(U_n)$. Then $V_n = \left(\bar{X}_n, \bar{P}_n, \bar{u}_n\right)$ is the solution of:

$$\begin{cases} \dfrac{d\bar{X}_n}{dt} - \left[F'(X_n)\right].\bar{X}_n = \dfrac{dX_n}{dt} - F(X_n)\,, \\ \bar{X}_n(0) - \bar{u}_n = X_n - u_n, \end{cases} \tag{9}$$

$$\begin{cases} \dfrac{d\bar{P}_n}{dt} + \left[F'(X_n)\right]^T.\bar{P}_n \quad + \left[F''(X_n).\bar{X}_n\right]^T P_n - C^T.C.\bar{X}_n, \\ \qquad\qquad\qquad = \dfrac{dP_n}{dt} + \left[F'(X_n)\right]^T.P_n - C^T\left(C.X_n - X_{obs}\right) \\ \bar{P}_n(T) = P_n(T)\,, \end{cases} \tag{10}$$

$$\alpha\bar{u}_n - \bar{P}_n(0) = \alpha\left(u_n - u^0\right) - P_n(0) \tag{11}$$

2. Next, we set $U_{n+1} = U_n - V_n$ where,

$$X_{n+1} = X_n - \bar{X}_n, \quad P_{n+1} = P_n - \bar{P}_n, \quad u_{n+1} = u_n - \bar{u}_n. \tag{12}$$

With $U_{n+1} = U_n - V_n$, the equations (9), (10) and (11) can be reformulated as: for X_n, P_n, u_n, find X_{n+1}, P_{n+1}, u_{n+1} such that

$$\begin{cases} \dfrac{dX_{n+1}}{dt} - \left[F'(X_n)\right].X_{n+1} = F(X_n) - \left[F'(X_n)\right].X_n, \\ X_{n+1}(0) = u_{n+1}, \end{cases} \tag{13}$$

$$\begin{cases} \dfrac{dP_{n+1}}{dt} + \left[F'(X_n)\right]^T.P_{n+1} = & -\left[F''(X_n).(X_{n+1} - X_n)\right]^T P_n \\ & + C^T\left(C.X_{n+1} - X_{obs}\right), \\ P_{n+1}(T) = 0, \end{cases} \tag{14}$$

$$\alpha\left(u_{n+1} - u^0\right) - P_{n+1}(0) = 0. \tag{15}$$

4. Linear Data Assimilation at Each Step

For a fixed n, the problems (13)-(15) can be seen as a linear data assimilation problem. Indeed, it is the optimality system associated with the minimization of the following problem: find $\widetilde{X}, \widetilde{u}$ such that:

$$\begin{cases} \dfrac{d\widetilde{X}}{dt} - \left[F'(X_n)\right].\widetilde{X} = \widetilde{f}, \quad t \in (0,T), \\ \widetilde{X}(0) = \widetilde{u}, \\ J_n(\widetilde{u}) = \inf\limits_{v \in \mathcal{X}} J_n(v), \end{cases} \tag{16}$$

where $\widetilde{f} = F(X_n) - \left[F'(X_n)\right].X_n$ and the cost function

$$\begin{aligned} J_n(\widetilde{u}) = \frac{\alpha}{2}\left\|\widetilde{u} - u^0\right\|_{\mathcal{X}}^2 \ &+ \ \frac{1}{2}\int_0^T \left\|C\widetilde{X} - X_{obs}\right\|_{\mathcal{X}}^2 dt \\ &- \ \frac{1}{2}\int_0^T \left(P_n, \left[F''(X_n).\left(\widetilde{X} - X_n\right)\right]\left(\widetilde{X} - X_n\right)\right)_{\mathcal{X}} dt. \end{aligned} \tag{17}$$

The Hessian H_n of the linear problem (16) is defined as successive solutions to:

$$\begin{cases} \dfrac{dR}{dt} - \left[F'(X_n)\right].R = 0, \quad t \in [0,T], \\ R(0) = u, \end{cases} \tag{18}$$

$$\begin{cases} \dfrac{dQ}{dt} + \left[F'(X_n)\right]^T.Q = -\left[F''(X_n).R\right]^T P_n - C^T C.R, \\ Q(T) = 0, \end{cases} \tag{19}$$

$$H_n v = \alpha v - Q(0). \tag{20}$$

Thus, if Newton's method converges, and $X_n \to X$, $P_n \to P$, $u_n \to u$ as $n \to +\infty$, then $H_n \to H(u)$ when $n \to +\infty$, where $H(u)$ is the Hessian of the function of cost $J(u)$ at the point u.

The Hessian H_n is symmetric. Moreover,

$$\begin{aligned} (H_n v, v)_{\mathcal{X}} = \ & (\alpha v - Q(0), v)_{\mathcal{X}} = \alpha(v,v)_{\mathcal{X}} - (Q(0), v)_{\mathcal{X}} \\ &= \alpha(v,v)_{\mathcal{X}} - \left(Q, \left[F''(X_n).R\right]^T P_n - C^T C.R\right) \\ &= \alpha(v,v)_{\mathcal{X}} + (CR, CR) - \left(\left[F''(X_n).R\right] R, P\right). \end{aligned}$$

When $C = I_{\mathcal{X}}$ (identity operator of $\mathcal{X}$) and under the condition

$$\left\|\left[F''(X_n).Q\right]\right\| \left\|P_n\right\| \leq 1, \tag{21}$$

H_n is positive definite for $\alpha > 0$:

$$(H_n v, v)_{\mathcal{X}} \geq \alpha (v, v)_{\mathcal{X}}. \tag{22}$$

Therefore the linear problem (16) has a unique solution $\widetilde{u}$, which provides a minimum to the function $J_n(\widetilde{u})$ defined by (18), whenever condition (21) is satisfied for each iteration.

5. Convergence of Newton's Method

Consider the ball

$$S_R(X_0) = \{X \in Y \ : \|X - X_0\| \leq R) .$$

Let $X_0 \in Y$ and $R \in I\!R_+^*$. Assume the initial problem satisfies the following conditions for all $X \in S_R(X_0)$:

1) The solution of

$$\begin{cases} \dfrac{d\varphi}{dt} - \left[F'(X)\right].\varphi = f, \quad t \in [0, T], \\ \varphi(0) = v, \end{cases}$$

satisfies the inequality

$$\|\varphi\| \leq c_1 \left(\|f\| + \|v\|_{\mathcal{X}}\right). \tag{23}$$

2) The solution of the adjoint problem

$$\begin{cases} -\dfrac{d\phi}{dt} - \left[F'(X)\right]^T.\phi = d, \\ \phi(T) = g, \end{cases}$$

satisfies

$$\|\phi\| + \|\phi(0)\|_{\mathcal{X}} \leq c_1^* \left(\|d\| + \|g\|_{\mathcal{X}}\right), \quad c_1^* = c_1^*(R, X_0) > 0. \tag{24}$$

3) The operator F is three times continuously Fréchet-differentiable and satisfies:

$$\left\|F''(X)\right\| \leq c_2, \quad c_2 = c_2(R, X_0) > 0 \ \forall X, \tag{25}$$

$$\left\|F'''(X)\right\| \leq c_3, \quad c_3 = c_3(R, X_0) > 0 \ \forall X. \tag{26}$$

We present Newton's method to solve the equation $\nabla J(u) = 0$ which is equivalent to the optimality system (4)-(6). Under the hypothesis of complete observation, $C = \mathrm{I}_{\mathcal{X}}$ ($\mathrm{I}_{\mathcal{X}}$ being the identity operator), the convergence of the Newton's algorithm is presented in Theorem 1 below:

We have

$$u_{n+1} = u_n - \left[J''(u_n)\right]^{-1} \nabla J(u_n). \tag{27}$$

The implementation of the Newton's algorithm requires the calculation of the Hessian ($H = J''$) of the cost function J and resolution of the linear system $J''(u_n)(u_{n+1} - u_n) = \nabla J(u_n)$. By using the definition of the Hessian, we note that Newton's method consists on the following steps:

1. Specify X_n, P_n by solving:

$$\begin{cases} \dfrac{dX_n}{dt} = F(X_n), \quad t \in [0, T], \\ X_n(0) = u_n, \end{cases} \tag{28}$$

$$\begin{cases} \dfrac{dP_n}{dt} + \left[F'(X_n)\right]^T . P_n = C^T (C.X_n - X_{obs}), \quad t \in [0, T], \\ P_n(T) = 0, \end{cases} \tag{29}$$

next evaluate $\nabla J(u_n) = \alpha (u_n - u_{obs}) - P_n(0)$

2. Determine $v_n = [J''(u_n)]^{-1} . \nabla J(u_n)$ as the solution of the following system:

$$\begin{cases} \dfrac{dR_n}{dt} = \left[F'(X_n)\right] . R_n, \ \ t \in [0, T], \\ R_n(0) = v_n, \end{cases} \tag{30}$$

$$\begin{cases} \dfrac{dQ_n}{dt} + \left[F'(X_n)\right]^T . Q_n = \left[F''(X_n) R_n\right]^T . P_n + C^T C . R_n, \\ Q_n(T) = 0, \end{cases} \tag{31}$$

$$\alpha v_n - Q_n(0) = H(u_n) v_n = \alpha (u_n - u_{obs}) - P_n(0). \tag{32}$$

3. Update u_{n+1} by setting:

$$u_{n+1} = u_n - v_n. \tag{33}$$

Next, after we have the value of the Hessian, and the solution to the corresponding equations, it is not a problem to solve the linear optimal control problem of the form (30)-(32) by the appropriate methods.

Newton's Method for $\nabla J(u) = 0$: Convergence

Assuming that the hypothesis (23)-(26) are satisfied, we find the solution of $\nabla J(u_n) = 0$ in the ball

$$S_r = \{u : \ \|u - u_0\|_{\mathcal{X}} \le r\} \quad u_0 \in X, \ \ r > 0. \tag{34}$$

Let $u_0 \in \mathcal{X}, \ \ u \in S_r$ and X_0, X as solutions of the initial non-linear problem (1) with the same initial condition $X_0(0) = u_0$ and $X(0) = u$. We also suppose that for $r > 0$, there exists $R > 0$ such that

$$\|X - X_0\|_{\mathcal{X}} \le R. \tag{35}$$

Let us introduce the following notation:

$$\begin{aligned} \eta = \|\nabla J(u_0)\|_{\mathcal{X}}, \quad K = c_1^2 c_1^* \{2c_1^* c_2 (c_2\beta + \|C^*C\|) + c_3\beta + c_1 c_2 \|C^*C\|\}, \\ \beta = c_1^* \|C^*C\| R + c_1^* \|C^* \left(CX_0 - \widehat{X}\right)\|, \end{aligned} \tag{36}$$

where the constants c_i, c_1^* are defined in Section 5. For the hypothesis (23)-(26), we have the following result:

Theorem 1. *Let $u_0 \in \mathcal{X}$ such that the Hessian $H(u_0) = J''(u_0)$ is positive definite with the appropriate minimum values $\lambda_{min} > 0$. Then, if*

$$h = \frac{K\eta}{\lambda_{min}^2} \le \frac{1}{2}, \tag{37}$$

and

$$r \ge r_0 = \frac{1 - \sqrt{1 - 2h}}{h} \frac{\eta}{\lambda_{min}}, \tag{38}$$

the equation $\nabla J(u) = 0$ has a solution $u \in S_r$ for which Newton's Method will converge.

The rate of convergence is defined by the following formula:

$$\|u - u_n\|_{\mathcal{X}} \le \frac{2h^{2^n}}{2^n} \frac{\eta}{h\lambda_{min}}. \tag{39}$$

In order to show the convergence of Newton's method, it is necessary to study the third derivative $J'''(u)$ of the cost function J. The complete proof of this theorem can be found in [8, 2, 1] and many references therein.

6. Application to the Data Assimilation in Hydrology

As an application we consider a one-dimensional model of infiltration in an unsaturated ground. The state variable is h, the water pressure, in a domain between the surface at $z = 0$ and the bottom at $z = Z$.

$$\begin{cases} C(h)\dfrac{\partial h}{\partial t} = \dfrac{\partial}{\partial z}\left[K(h)\left(\dfrac{\partial h}{\partial z} - 1\right)\right] & (t,z) \in [0,1] \times [0,1], \\ h(0,z) = h_{ini}(z), \\ h(t,0) = h_{surf}(t), \\ h(t,Z) = h_{bot}(t), \end{cases} \tag{40}$$

$C(h)$ and $K(h)$ are diffusivity hydraulic and conductivity hydraulic, respectively, and are given by:

$$C(h) = \begin{cases} \dfrac{\theta_s(2-n)}{h_g}\left(\dfrac{h}{h_g}\right)^{n-1}\left[1 + \left(\dfrac{h}{h_g}\right)^n\right]^{\frac{2}{n}-2}, \text{ if } h < 0, \\ C(h) = 0, \text{ if } h \geq 0, \end{cases} \tag{41}$$

$$K(h) = \begin{cases} K_s\left[1 + \left(\dfrac{h}{h_g}\right)^n\right]^{\eta(\frac{2}{n}-1)}, \text{ if } h < 0, \\ K(h) = K_s, \text{ if } h \geq 0 \text{ (cas saturé)}, \end{cases} \tag{42}$$

We have five hydrodynamic parameters

$$K_s, h_g, \theta_s, \eta, \text{et } n, \tag{43}$$

where h_g is the inflection point of the curve

$$h = f(\theta), \quad f(x) =: \theta_s\left[1 + \left(\frac{x}{h_g}\right)^n\right]^{\frac{2}{n}-1}.$$

For the numerical results the five parameters are $K_s = 4.28 \times 10^{-3}$, $h_g = -16.40$, $\theta_s = 0.312$, $\eta = 6.73$, and $n = 2.79$, $h_{surf}(t) = 0.3024$, $h_{surf}(t) = 0.2873$. The aim is to identify the initial condition. If all the parameters of the model are known, then it is possible to compute a cumulated infiltration given by:

$$I_{\text{cum}}(t) = \int_0^Z (\theta(t,z) - \theta_{\text{ini}})\, dz, \tag{44}$$

where $\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K(\theta)\left(\frac{\partial h}{\partial z} - 1\right)\right]$, θ, is the water content of the ground. The observation bear on the cumulated infiltration $I_{\rm obs}$ (see [8] and many references therein).

The data assimilation problem is to determine $U = (h_{\rm ini})$ minimizing the cost function, J, defined by:

$$J(U) = \frac{\Delta t}{2} \sum_{j=0}^{M} \left(I_{\rm cum}(t_j) - I_{\rm obs}\right)^2 + \frac{\lambda}{2}\|U - U_e\|^2, \tag{45}$$

where $U_e = h^0_{\rm ini}$ is an estimation of the initial condition.

If P is the adjoint variable, then adjoint model is defined by (see [8, 9]):

$$\begin{cases} -\dfrac{\partial}{\partial t}(C.P) + \left[\dfrac{\partial C}{\partial h}\right] \cdot \left(P. \left[\dfrac{\partial h}{\partial t}\right]\right) - \dfrac{\partial}{\partial z}\left[K \dfrac{\partial P}{\partial z}\right] + \left[\dfrac{\partial K}{\partial h}\right] \cdot \left(\dfrac{\partial P}{\partial z}. \left[\dfrac{\partial h}{\partial z} - 1\right]\right) \\ = \left(I_{cal} - I_{obs}\right) \dfrac{\partial I_{cal}}{\partial h} \delta(t - t_i), \quad i = 0, \ldots.(N-1), \\ P(t = T, z) = 0, \\ P(t, z = 0) = 0, \\ P(t, z = Z) = 0. \end{cases} \tag{46}$$

From the backward integration of the adjoint system we deduce the gradient of the cost function with respect to the initial condition:

$$\begin{aligned} \nabla J\left(h_{\rm ini}\right) &= \int_0^Z \left((C(h)P)_{|t=0} - \left(I_{cal}(t) - I_{obs}(t)\right) \delta(t - t_j) \frac{\partial f(h)}{\partial h} \right) dz \\ &+ \lambda \left(h_{\rm ini} - h^0_{\rm ini}\right). \end{aligned} \tag{47}$$

It is well known that there is no commutativity between discretization and derivative of the adjoint, therefore, all the former calculations should also be carried out on the discretized model.

The model is discretized with a finite difference scheme in space with $Z =$ 1m, the grid size is $\Delta Z = 1$cm. The temporal scheme is an implicity Euler scheme with $T = 1$ hour and a time step of $\Delta T = 10$ seconds.

The numerical experiment has been realized on a material known as Grenoble sand, see [8]. These observation (cumulative infiltration) are distributed using four different configurations: discrete data distributed every 10 seconds (case 1), every 60 seconds (case 2), every 300 seconds (case 3), and every 720 seconds (case 4).

We used BFGS (see Gilbert and Lemaréchal [5]) to determine the optimal initial and boundary conditions that will be used to initialize the Newton's method. The decay of the functional (45) during the iteration process of the minimization procedure is shown in Figure 1 for the all the cases. Figure 1 shows the convergence of the of the assimilation process.

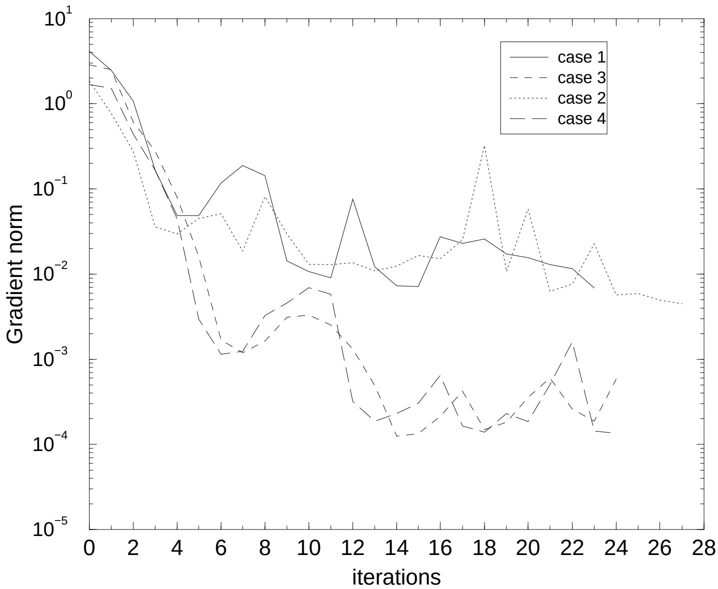

Figure 1: Gradient norm of the cost function during the iterations (using BFGS algorithm [5])

Numerical Results for $\nabla J(u) = 0$

The first step consists of finding the initial and boundary conditions u_0 such that the Hessian of the cost function at u_0, $H(u_0)$, is positive definite. This value of u_0 found will serve as the starting point of the algorithm (28)-(33) to solve for $\nabla J(u) = 0$. This starting point, u_0, is easily found (see Figure 1). Indeed, at this point u_0, $H(u_0)$ is necessarily positive define. This does express the necessary and sufficient condition for the existence of the optimal control problem (at least locally). To check it out, we have carried out a spectral decomposition of the Hessian following the four different configurations of the observation. In each case, we have carried out 100 iterations of the Lanczos algorithm. This step is used to check if the Hessian matrix at the point u_0 is positive define. Table 1 displays the result and clearly shows that all the eigenvalues are positive irrespective of the temporal distribution of the observation. Therefore, the Hessian is positive define at u_0.

Problems	λ_{min}	λ_{max}	Cond.	Criteria of Convergence
Case 1	0.643031	0.982707	1.528241	
Case 2	0.878232	0.878984	1.000855	$\|\lambda_n - \lambda_{n-1}\| \leq 10^{-10}$
Case 3	0.8205396	1.053039	1.283350	n values of the n-th iteration
Case 4	0.817760	0.9305687	1.1379480	

Table 1: Smallest eigenvalue (λ_{min}), largest eigenvalue (λ_{max}) and condition number of the Hessian. The four cases correspond to the different temporal distribution of the observation.

Let us present the results obtained by deploying the Newton's algorithm (28)-(33). We fix the same stopping criterion:

$$\|v_n\|_X = \|u_{n+1} - u_n\|_X \leq \varepsilon = 10^{-10}. \tag{48}$$

The inversion of (27) is made by solving the minimization problem:

$$v^* = \inf_{v \in \mathcal{X}} g(v),$$

with

$$g(v) = \frac{1}{2}(H(u_n)v, v)_{\mathcal{X}} - (\nabla J(u_n), v)_{\mathcal{X}}.$$

Thus, a method of gradient descending type is used to minimize g.

Figures 2, 4, 6 and 8 display the norm of the gradient during Newton's iterations for all four configurations. Figures 3, 5, 7, 9, display in a continuous line, the residual norm during the process of inverting the Hessian, in the dotted line, the norm of the residual at the last inversion.

7. Concluding Remarks

The Newton's method to find optimality condition, $\nabla J(u) = 0$, gives good results. We are able to obtain a decay of the cost function's gradient in the order of magnitude of 10^{-8} (Cases 1 and 4) and 10^{-11} (Case 3). This results are much better comparatively to the classical decent algorithm (quasi-Newton [5] see Figure 1) in which we do not have any information about the Hessian of the cost function. As a matter of fact, in this type of algorithm, the Hessian

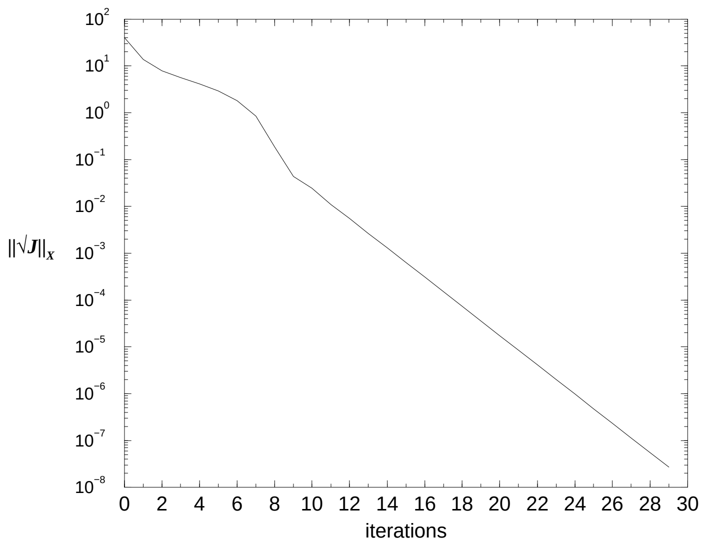

Figure 2: Evolution of the gradient of the cost function during the iterations (case 1)

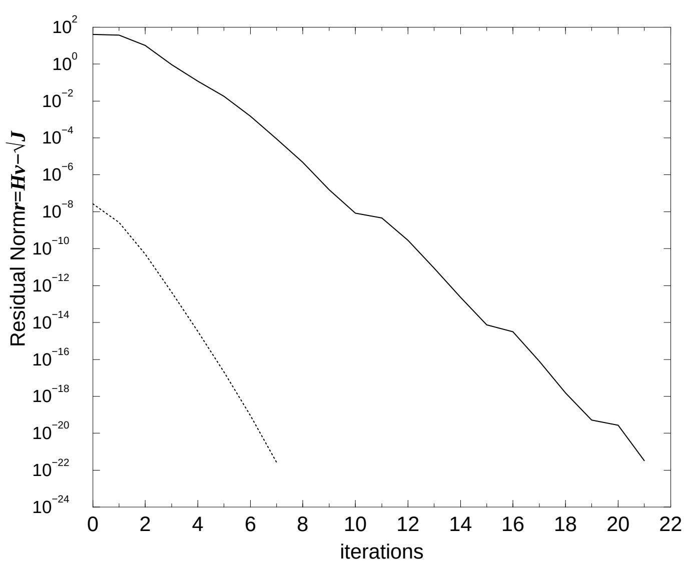

Figure 3: Residual norm during the inversion process: Continuous line 1-st inversion, dotted line, 2-nd inversion (case 1)

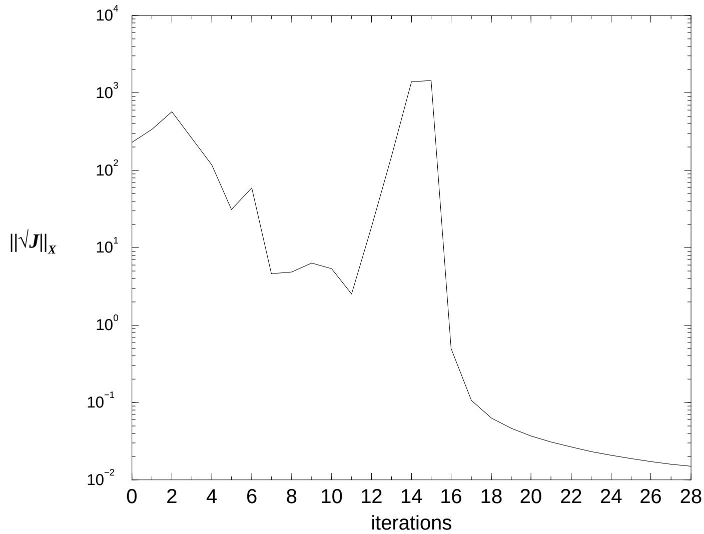

Figure 4: Evolution of the gradient of the cost function during the iterations (case 2)

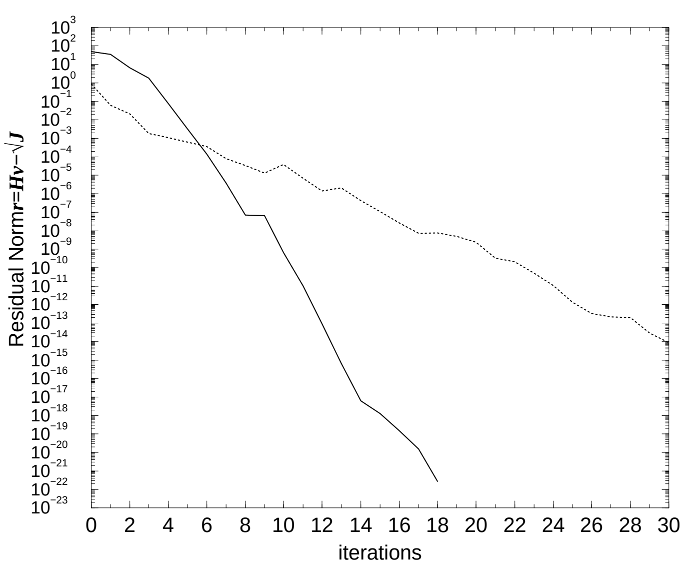

Figure 5: Residual norm during the inversion process: Continuous line 1-st inversion, dotted line, 2-nd inversion (case 2)

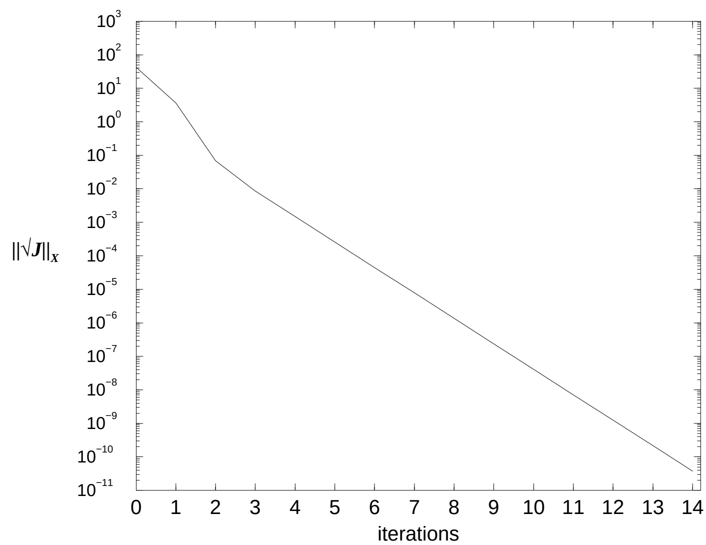

Figure 6: Evolution of the gradient of the cost function during the iterations (case 3)

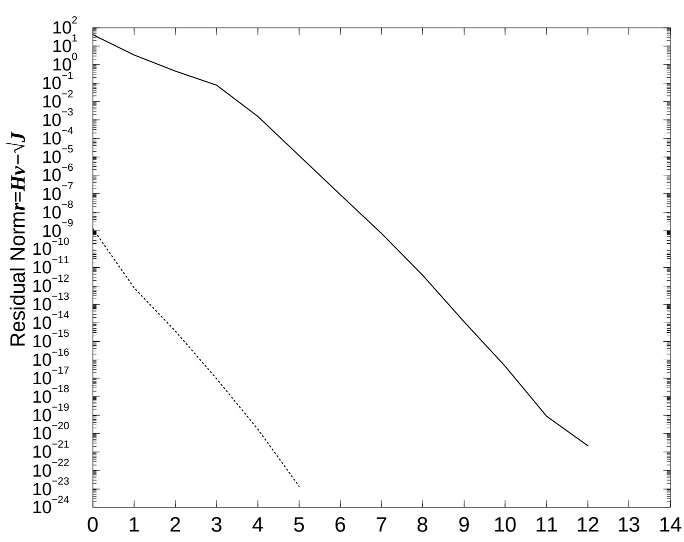

Figure 7: Residual norm during the inversion process: Continuous line 1-st inversion, dotted line, 2-nd inversion (case 3)

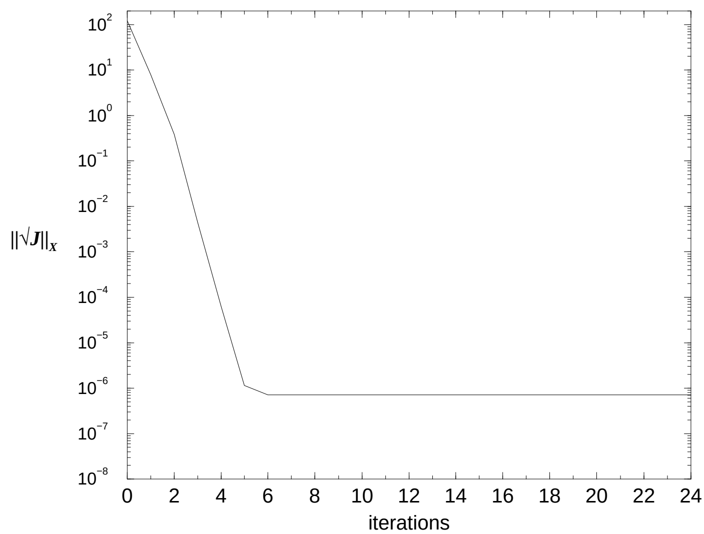

Figure 8: Evolution of the gradient of the cost function (left) and the normal of u (right) during the iterations (case 4)

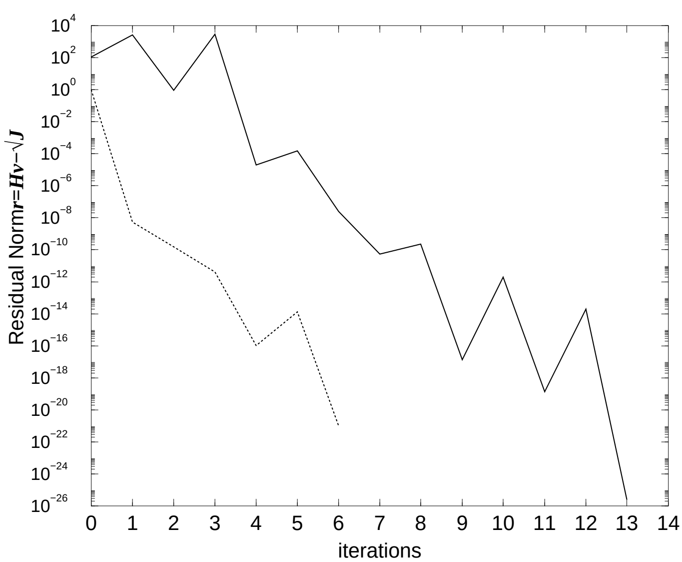

Figure 9: Residual norm during the inversion process: Continuous line 1-st inversion, dotted line, 2-nd inversion (case 4)

of the cost function is approximated by the successive variations of the cost function the gradient between two consecutive iterations. The price paid here by implementing Newton's method is the use of the second order adjoint obtained by constructing the exact Hessian. The quality of results obtained here is at that price.

The Newton's methods is an extremely powerful tool that can be extended in other applications especially in data assimilation in geophysical physical flows. Its implementation requires the use of the second order adjoint that can be used for the sensitivity analysis. For geophysical models, generally non-linear, it is probably one of the tools which can better compute characteristics of non-linear models.

Acknowledgments

This work was partially supported by the Department of Defense, Army Research Laboratory, DAD19-01-02-0013.

References

[1] V.I. Agoshkov, G.I. Marchuck, On solvability and numerical solution of data assimilation problems, *Russ. J. Numer. Anal. Math. Modelling*, **8**, No. 1 (1993), 1-16.

[2] F.-X. Le Dimet, V. Shutyaev, On Newton methods in data assimilation, *Russ J. Numer. Anal. Math. Modelling*, **15**, No. 5 (2000), 419-434.

[3] F.-X. Le Dimet, V.P. Shutyaev, On data assimilation for quasilinear parabolic problems, *Russ. J. Numer. Analys. Math. Modelling*, **16**, No. 3 (2001), 247-259.

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

[5] J.C. Gilbert, C. Lemaréchal, Some numerical experiments with variable storage quasi-Newton algorithms, *Mathematicals Programming*, **45** (1989), 407-435.

[6] J.L. Lions, *Optimal Control of Systems Gouvered by Patial Differantial Equations*, Dunod (1968).

[7] D.C. Lui, J. Nocedal, On the limited memory BFGS method for large scale optimization, *Mathematicals Programming*, **45** (1989), 503-528.

[8] Pierre Ngnepieba, *Data Assimilation and Parameters Identification: An Application in Hydrology*, Ph.D. Dissertation, University Joseph Fourier of Grenoble 1, Grenoble, France (2001).

[9] P. Ngnepieba, F.-X Le Dimet, A. Boukong, G. Nguetseng, Parameter identification: An application to the Richards equation, *ARIMA's Review*, Volume 1 (2002), 127-157.

[10] P. Ngnepieba, F.-X. Le Dimet, A. Jones, Jr., V. Shutyaev, On the Newton's algorithm for the data assimilation problem: An application in hydrology, *International Journal of Pure and Applied Mathematics*, **52**, No. 1 (2009), 67-89.

[11] Pierre Ngnepieba, D. Stepehens, F.-X. Le Dimet, The Newton method in problems of variational data assimilation: Application to an infiltration model, In: *Proceedings of the First Buea International Conference on Mathematical Sciences*, University of Buea, May 12-17, 2009. Buea, Cameroon (2009).

[12] M.M. Vainberg, *Variational Methods for the Study of Nonlinear Operators*, London Holden-Day (1964).