A Penalized 4-D Var data assimilation method for reducing forecast error

M. J. Hossen

Department of Mathematics and Natural Sciences
BRAC University
66 Mohakhali, Dhaka-1212, Bangladesh

I. M. Navon*
Department of Scientific Computing
Florida State University
Tallahassee, FL 32306, USA

Dacian N. Daescu

Department of Mathematics and Statistics,
Portland State University,
P.O. Box 751, Portland, OR 97207, USA

Abstract

Four dimensional variational (4D-Var) Data Assimilation (DA) method is used to find the optimal initial conditions by minimizing a cost function in which background information and observations are provided as the input of the cost function. The corrected initial condition based on background error covariance matrix and observations improves the forecast. The targeted observations determined by using a targeting method, for instance adjoint sensitivity, observation sensitivity or singular vector may further improve the forecast . In this paper, we are proposing a new technique–consisting of a penalized 4D-Var DA method that is able to reduce the forecast error significantly. Here we are penalizing the cost function by the forecast aspect defined over the verification domain at the verification time. The result shows that the initial condition is optimally estimated, thus resulting in a better forecast by significantly reducing the forecast error .

*Corresponding author: Department of Scientific Computing, Florida State University, Tallahassee, FL 32306-4120

1 Introduction

Numerical weather prediction (NWP) is based on the the integration of a dynamic system of partial differential equations modeling the behavior of the atmosphere. Therefore discrete initial conditions describing the state of the atmosphere have to be provided prior to the integration, since they, along with the model equations and boundary conditions, control the evolution of the solution trajectory in space and time. To find the best estimate for the initial condition we use four dimensional variational(4D-Var) data assimilation (DA) techniques (LeDimet and Talagrand[1]; Derber [2]; Lewis and Derber [3]; Talagrand and Courtier [4]). In this method the best initial condition is estimated by minimizing the cost function defined as the combination of deviations of the desired analysis from a forecast and observations weighted by the inverse of the corresponding forecast and observation-error covariance matrices.

4D-Var DA method uses a flow dependent background error covariance for estimating the atmospheric state and assimilated indirect observational data such as satellite radiance without transforming them into analysis variables. The computational expense of the variational assimilation can be reduced by using the adjoint of the numerical model to calculate all of the components of the gradient of the cost function with respect to the initial conditions in one integration of the forward model followed by integration of the corresponding adjoint model. The adjoint model arises from the theory of optimization and optimal control of partial differential equations (Lions [5]; Glowinski [6]). Its theoretical aspects were presented by LeDimet and Talagrand [1]; Talagrand and Courtier [4] and LeDimet et.al [7].

Results from 4D-Var experiments with large scale numerical model were published in the early 1990s (Thepaut et al. [8]; Navon et al. [9]; M. Zupanski [10]). Thepaut et al. [11] demonstrated the ability of 4D-Var method to generate flow dependent and baroclinic structure functions in meteorological analysis.

The forecast impact of targeting of observationsis determined by the distribution and types of routine and targeted observations, the quality of the background or the first guess, and the ability of the data assimilation procedure to combine information from the both background and observations. To deploy targeted observations we need to define a target area. Typically, an objective procedure (often based on adjoint or ensemble techniques) is used a day or more in advance to identify a target region for the spatial observations identified using singulsr vector. It can also be determined on the basis of high probability for a large or a fast growing initial condition error.

The goal of the adaptive observations is to add targeted observations inside the sensitive regions in order to improve the initial conditions so that the forecast error has been reduced significantly. Adjoint based observation sensitivity techniques may be used to identify the adaptive observation space and time location that are valuable for the assimilation procedure, to conduct optimal data thinning and to design the cost-effective field experiments for collecting adaptive observations. Langland [12] shows that a small number of additional observational resources must be deployed in order to improve a specific forecast aspect. The design of cost-effective observation targeting strategies relies on the ability to a-priori identify optimal sites for collecting data of large impact on reducing forecast errors. Le Dimet et.al [7] presented a theoretical formulation of the sensitivity analysis in variational data assimilation in the context of optimal control. Daescu and Navon [13] proposed a new adjoint sensitivity approach where they considered the interaction between adaptive observations and routine observations. The Singular Vector (SV) approach provides an additional possibility of searching for directions in phase space where the errors in the initial condition will amplify rapidly. The specification of the initial and final norms plays a crucial role. In the European Center for Medium-Range Weather Forecasts (ECMWF) operational EPS, SVs are computed with the so called total energy norm at initial and final time. It can be shown that among simple norms, the total energy norm provides SVs which agree best with analysis error statistics (Palmer et al. 1998 [14]). Barkmeijer et al. [15] and [16] have shown that the Hessian of the cost function in a variational data assimilation scheme can be used to compute SVs that incorporate an estimate of the full analysis error covariance at initial time and total energy norm at final time. This type of singular vector is called Hessian singular vector. Ehrendorfer and Tribbia [17] state that such an approach to determine SVs provides an efficient way to describe the forecast error covariance matrix when only a limited number of linear integrations are possible. Though finding the Hessian matrix explicitly involves a computationally intensive effort, we can calculate Hessian vector product by using second order adjoint (see LeDimet et al. [18]). This also requires an efficient generalized eigenvalue problem solver to compute Hessian SVs. For advanced work on this topic see Hodinez and Daescu [19].

In this paper we are proposing a new cost function which can be minimized to find the optimal estimate of the initial condition. This initial condition reduces the forecast error significantly over the verification domain at the verification time. The new cost function is obtained by penalizing the cost function with a term defined as being proportional to the square of the distance between analysis and both background and observation, with the forecast aspect being defined over the verification domain at verification time.

The structure of the paper is as follows . In section 2 we present derivation of the equations and the penalized 4-D VAR algorithm, while in Section 3 we describe the twin numerical experiments. Numerical results are then presented and discussed in Section 4 along with the pseudo-algorithm of the penalized 4-D Var approach.

Section 5 is dedicated to summary and conclusions.

2 Derivation of the Equations and Penalized 4-D VAR Algorithm

In four dimensional variational data assimilation(4D-Var DAS), an initial condition is sought such that the forecast best fits the observations within an assimilation window $[t_0, t_f]$. 4D-Var DAS provides an optimal estimate $x_0^a \in \mathbb{R}^n$ to the initial condition of a nonlinear forecast model by minimizing the cost function defined as

$$
\begin{aligned}
\mathcal{J}(x_0) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\
&+ \frac{1}{2}\sum_{i=0}^{N}(y_i - H_i x_i)^T \mathbf{R}_i^{-1}(y_i - H_i x_i) \\
& x_0^a = \arg\min \mathcal{J}
\end{aligned}
\tag{2.1}
$$

where $x_0 = x(t_0)$ denotes the initial state at the initial time t_0, x_b is a prior(background) estimate to the initial state, $y_i \in \mathbb{R}^{k_i},\ i = 0, 1, 2, .., N$ is the set of observations available at time t_i and $x_i = \mathcal{M}(t_0, t_i)(x_0)$ is the nonlinear model forecast state at time t_i and $H_i : \mathbb{R}^n \to \mathbb{R}^{k_i}$ is the observation operator that maps the state space into the observation space at time t_i. $\mathbf{B}$ is the background error covariance matrix and $\mathbf{R}_i$ is the observational error covariance matrix at time t_i. We assume that background errors and observation errors are uncorrelated with each other. In our case, we take the error covariance matrices $\mathbf{B}$ and $\mathbf{R}_i$ to be diagonal. The control variable or the variable with respect to which the cost function (2.1) is minimized is the initial state of the model x_0 . The model $\mathcal{M}$ is assumed to be perfect by imposing the model equations as the strong constraint.

In order to minimize the cost functional in (2.1) with respect to x_0, we need to calculate the gradient of the cost functional with respect to the control variable i.e. $\nabla_{x_0}\mathcal{J}$. The adjoint method provides an efficient approach to calculate the gradient of the cost function with respect to control variables. The gradient of the cost functional (2.1) is

$$
\begin{aligned}
\nabla_{x_0}\mathcal{J} &= \mathbf{B}^{-1}(x - x_0) - \sum_{i=0}^{N} \mathbf{M}_{0,i}^T \mathbf{R}_i^{-1}(y_i - H_i x_i) \\
&= \mathbf{B}^{-1}(x - x_0) - \mathbf{R}_0^{-1}(y_0 - H_0 x_0) \\
&\quad -\mathbf{M}_{0,t_a}^T \mathbf{R}_{t_a}^{-1}(y_{t_a} - H_{t_a} x_{t_a}) \\
&\quad -\mathbf{M}_{0,t_r}^T \mathbf{R}_{t_r}^{-1}(y_{t_r} - H_{t_r} x_{t_r})
\end{aligned}
\tag{2.2}
$$

where routine observations are available at $t = t_r$, t_r being the time for routine observations while adaptive observations are available at $t = t_a$, time for adaptive observations

Background error covariance is estimated by using well-known NMC-method (Parrish and Derber [20]). In this process, background errors are assumed to be well approximated by averaged forecast difference (e.g. month-long series of 24hr - 12hr forecasts valid at the same time) statistics:

$$
\begin{aligned}
& B = \overline{\epsilon_b^T \epsilon_b} = \overline{(x_b - x_t)^T (x_b - x_t)} \\
& \approx \overline{(x^{t+24} - x^{t+12})^T (x^{t+24} - x^{t+12})}
\end{aligned}
\tag{2.3}
$$

where x_t is the true atmospheric state and x_b is the background error. The bar denotes an average over time and/or space.

2.1 The adjoint sensitivity (AS) approach

The first approach to identify the adaptive observations locations is adjoint sensitivity method. In practice it is of interest to assess the observation impact on the forecast measure $\mathcal{J}^v(x_v)$ on the verification domain at the verification time t_v. The verification domain, denoted by $\mathcal{D}_v$, is the domain where forecast error is significant. The functional $\mathcal{J}^v$ is defined, see Daescu and Navon [13], as a scalar measure of the forecast error over $\mathcal{D}_v$

$$\mathcal{J}^v(x_0) = \frac{1}{2}(x_v^f - x_v^t)^T P^T E P (x_v^f - x_v^t) \tag{2.4}$$

where x_v^f is the model forecast at the verification time initialized from x_0^a and x_v^t is the verification state at t_v initialized from x_0^t that serves as a proxy to the true atmospheric state. P is a projection operator on $\mathcal{D}_v$ satisfying $P^*P = P^2 = P$ and E is a diagonal matrix of the total energy norm.
To select the adaptive observations locations, the gradient of cost functional $\mathcal{J}^v$ defined in equation (2.4) is used. The gradient of the function (2.4) at t_i is defined as

$$\nabla \mathcal{J}^v(x_i) = \mathbf{M}_{i,v}^T P^T E P (x_v^f - x_v^t) \tag{2.5}$$

where $x_i = x(t_i)$

2.1.1 Location of adaptive observations by AS

We use the gradient of the function defined in (2.5) to evaluate the sensitivity of the forecast error with respect to the model state at each targeting instant t_i. A large sensitivity value indicates that small variations in the model state x_i will have a significant impact on the forecast at the verification time. The adjoint sensitivity field with respect to total energy metric is defined as

$$F_v(\lambda, \theta) = ||\nabla_{x_i} \mathcal{J}^v||_E \tag{2.6}$$

where E is the total energy metric and weighted norm is defined as

$$||x||_E = \frac{1}{2}(u^2 + v^2) + \frac{h^2}{h_0} \tag{2.7}$$

where h_0 is the mean geopotential height of the reference data at the initial time. The adaptive observation at target instant t_i are deployed at the first n_i locations (λ, θ) where $F_v(\lambda, \theta)$ attains largest values.

2.2 Penalized four dimensional variational method

The fundamental idea of the penalty method is to replace a constrained optimization problem by a series of unconstrained problems whose solutions ideally converge to the solution of the original constrained problem. The unconstrained problems are formed by adding a term to the objective

function that consists of a penalty parameter and a measure of violation of the constraints. The severity of the penalty is determined by a parameter, r, known as penalty parameter.
The general form of constrained minimization is

$$\min_x \mathcal{J}(x) \tag{2.8}$$

$$\text{subject to } \ c(x) = 0; \tag{2.9}$$

where x is an n-dimensional vector and $c(x)$ is an m-dimensional vector. Instead of solving the constrained optimization problem we can solve an unconstrained minimization problem by defining the quadratic penalty function,

$$Q(x;r) = \mathcal{J}(x) + \frac{1}{2} r |c(x)|^2$$

for any scalar $r > 0$. We seek the approximate minimizer x_k of the function $Q(x;r)$ as $r_k \to \infty$ as $k \to \infty$.
The algorithm of exterior penalty method (since it uses a sequence of infeasible points and feasibility is obtained only at the optimum) can be summarized as:

1. Start with an initial point x_0 and an initial value of parameter $r_0 > 0$. set $k = 0$

2. Minimize $Q(x_k; r_k)$ with x_k by using an unconstrained minimization method and obtain x_k^*.

3. Test whether x_k^* is a solution of the problem i.e. satisfying the constraints $c(x) = 0$ within some prescribed accuracy criteria. If this is true, terminate the process, otherwise, set $r_{k+1} = \mu r_k$ where $\mu \in [4, 10]$ suggested by Bertsekas [21]

4. Set $k = k + 1$, use as a new starting point $x_k = x_k^*$ and go to step 2.

The method depends for its success on sequentially increasing the penalty parameter r to high values. The approximate minimizer becomes increasingly accurate as r gets higher.

2.2.1 Penalized cost function

In this work, we penalize the cost functional $\mathcal{J}(x)$ defined in (2.1) by adding a penalty term

$$r\mathcal{J}_v \tag{2.10}$$

to the cost function in order to reduce the forecast error over the verification domain at the verification time. $\mathcal{J}^v$ is the forecast aspect defined in the equation (2.4). In this work we employ the penalty method in a weak sense that we try to find the minimizer by reducing the forecast aspect $\mathcal{J}_v$ until it reaches a prescribed small value ϵ instead of attaining a perfect steady state where forecast error is absolutely zero i.e. $\mathcal{J}_v = 0$. That is, we are looking for the optimal initial condition so that the cost function is minimized subject to the constraint that the forecast error is very small. In this case, the penalty parameter is sufficiently large but does not tend to infinity which is equivalent to imposing an inequality constraint of the form

$$\mathcal{J}_v \leq \epsilon \tag{2.11}$$

The modified cost function is

$$\mathcal{J}' = \mathcal{J} + \frac{1}{2} r \mathcal{J}_v \tag{2.12}$$

That is,

$$\begin{aligned} \mathcal{J}'(x_0) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\ &+ \frac{1}{2}\sum_{i=0}^{N}(y_i - H_i x_i)^T \mathbf{R}_i^{-1}(y_i - H_i x_i) \\ &+ r\frac{1}{2}(x_v^f - x_v^t)^T P^T E P (x_v^f - x_v^t) \end{aligned} \tag{2.13}$$

where $x_v^f = M(x_0)$. The minimizer of the cost function (2.13) is obtained by using an unconstrained minimization routine which requires the gradient of the cost function (2.13). The gradient of the penalized cost function is obtained by using the following formula

$$\begin{aligned} \nabla_{x_0} \mathcal{J}' &= \mathbf{B}^{-1}(x - x_0) - \sum_{i=0}^{N} \mathbf{M}_{0,i}^T \mathbf{R}_i^{-1}(y_i - H_i x_i) \\ &+ r \mathbf{M}_{i,v}^T P^T E P (x_v^f - x_v^t) \end{aligned} \tag{2.14}$$

3 Description of Twin Numerical Experiments

3.1 Experimental setup

The numerical experiments were performed in the twin experiment framework using a finite volume global two dimensional shallow water equations model that has been widely used as an essential tool for testing promising numerical methods for solving geophysical science problems. The shallow water(SW) equations, a first prototype of the partial differential equations, describes the horizontal dynamics of the atmosphere. The shallow water equations in spherical coordinates are written as

$$\frac{\partial h}{\partial t} + \frac{1}{a\cos\theta}\left[\frac{\partial}{\partial\lambda}(hu) + \frac{\partial}{\partial\theta}(hv\cos\theta)\right] = 0 \tag{3.1}$$

$$\frac{\partial u}{\partial t} + \frac{1}{a\cos\theta}\left[u\frac{\partial u}{\partial\lambda} + v\cos\theta\frac{\partial u}{\partial\theta}\right] - \left(f + \frac{u}{a}\tan\theta\right)v + \frac{g}{a\cos\theta}\frac{\partial h}{\partial\lambda} = 0 \tag{3.2}$$

$$\frac{\partial v}{\partial t} + \frac{1}{a\cos\theta}\left[u\frac{\partial v}{\partial\lambda} + v\cos\theta\frac{\partial v}{\partial\theta}\right] + \left(f + \frac{u}{a}\tan\theta\right)u + \frac{g}{a}\frac{\partial h}{\partial\theta} = 0 \tag{3.3}$$

where $f = 2\Omega\sin\theta$ is the Coriolis parameter, Ω is the angular speed of the rotation of the earth, h is the height of the homogeneous atmosphere, u and v are the zonal and meridional wind components respectively, θ and λ are the latitudinal and longitudinal directions, respectively, a is radius of the earth and g is the gravitational constant.

We consider a spatial discretization on a 72×37 grid ($5^\circ \times 5^\circ$ resolution). As a result of this, the dimension of the discrete state vector $x = (h, u, v)$ is 7776. For numerical stability we choose the integration time step, $\Delta t = 900s$. For our numerical experiment we consider the 500mb European Center for Medium-Range Weather Forecasts(ECMWF) ERA-40 data valid for March 15, 2002 00h as a true (reference) atmospheric state x_0^t. The model states at the initial time and after 30h integration are displayed in Figure 1 and 2. The background field x_b is obtained from a 6h integration of SW model initialized at $t_0 - 6$h with x_0^t. Observational data for the data assimilation procedure is generated from the SW model trajectory initialized with x_0^t and corrupted with random errors from a normal distribution $N(0, \sigma^2)$. We choose the standard deviation $\sigma_h = 5$ for the height and $\sigma_u = \sigma_v = 0.5$ for the velocities. The background error covariance matrix is calculated by using NMC method as described before. We assumed that background and observation errors are uncorrelated. Therefore, the error covariance matrices are diagonal.

4D-Var DAS is carried out in the assimilation window $[t0, t_0 + 6h]$. The routine observation for our experiment is available at t_0 and $t_0 + 6$h only on a coarse $10^\circ \times 10^\circ$ mesh grid and the total number of observation locations are 648. So the observation operator is thus a 648×2664 matrix with entries of 0 and 1 only. At the verification time $t_v = t_0 + 30h$ the forecast error is calculated by using reference state, $x_v^t = \mathcal{M}_{t_0 \to t_v}(x_o^t)$ and the forecast from background $x_v^f = \mathcal{M}_{t_0 \to t_v}(x_b)$. The forecast error is displayed in Figure 3 calculated by using $||\mathcal{M}_{0,t_v}(x_b) - \mathcal{M}_{0,t_v}(x_0^t)||_E$ at t_v obtained by using background estimate as the initial conditions. $||\mathcal{M}_{0,t_v}(x_0^a) - \mathcal{M}_{0,t_v}(x_0^t)||_E$ at t_v obtained by using optimal analysis x_0^a as the initial condition. We have found that the forecast error is large over the domain $D_v = [65^\circ S, 35^\circ S] \times [100^\circ W, 65^\circ W]$ which will be considered as the verification domain for our experiment.

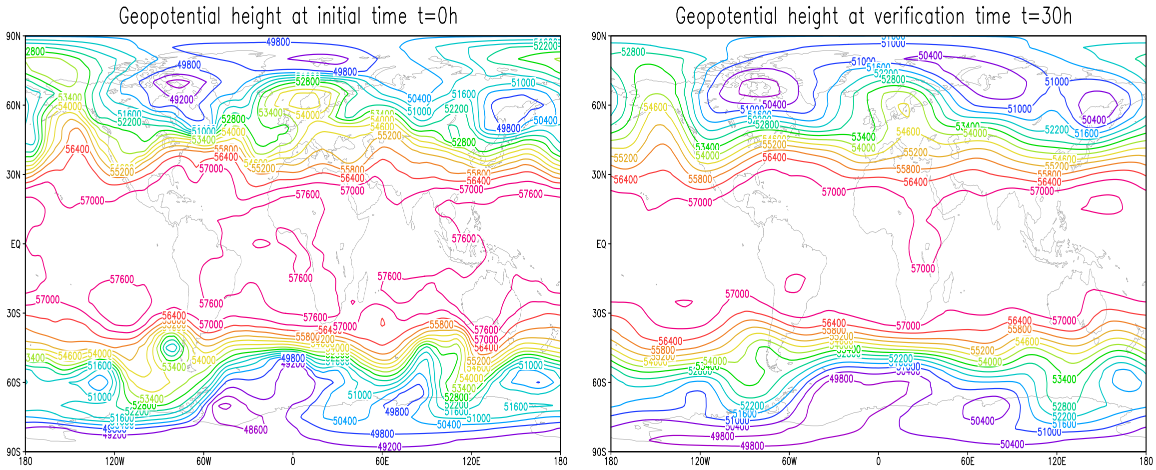


Figure 1: Graph of the geopotential height at the initial time $t = 0$h (left) and at verification time, $t = 30$h (right)

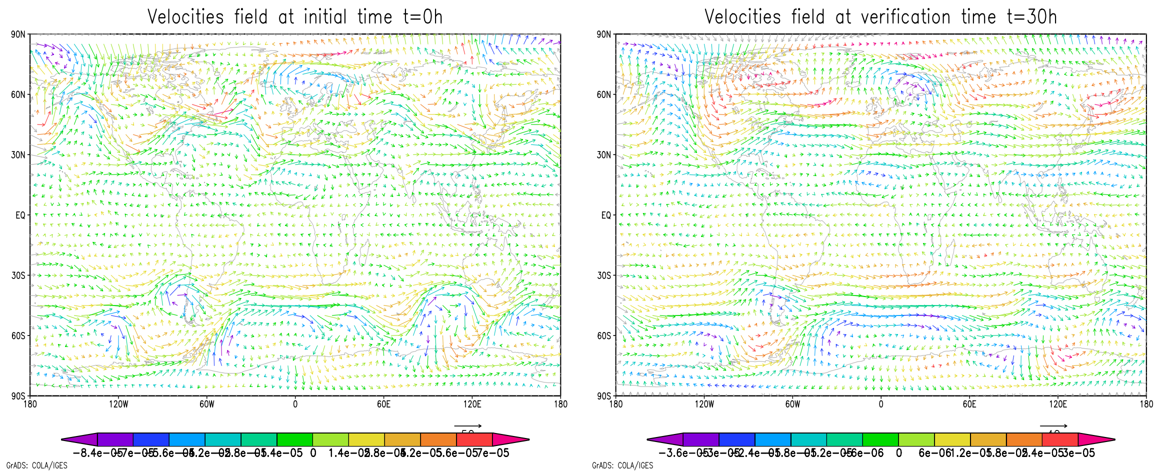


Figure 2: Graph of the velocities at the initial time $t = 0$h (left) and at verification time, $t = 30$h (right)

4 Results

In our twin experiment, we first minimize the cost function without adding the penalty term to the cost function. The minimization process terminates successfully after 16 iterations and 21 function evaluations. The cost function is decreased by 10 orders of magnitude while the norm of the gradient is decreased by 4 orders of magnitude. We used the resulting optimal initial condition to compute the forecast error. The result is shown in Figure 4.

We then estimate the optimal initial condition by taking some adaptive observation with the routine observations by using the adjoint sensitivity method. The result shows that only few adaptive observation added to the routine observations improve the forecast slightly. To compute the adjoint sensitivity, we used the algorithm mentioned below. The adjoint sensitivity and adaptive observation locations are displayed in Figures 5(a)- 5(f). The forecast error over D_v at t_v is obtained by

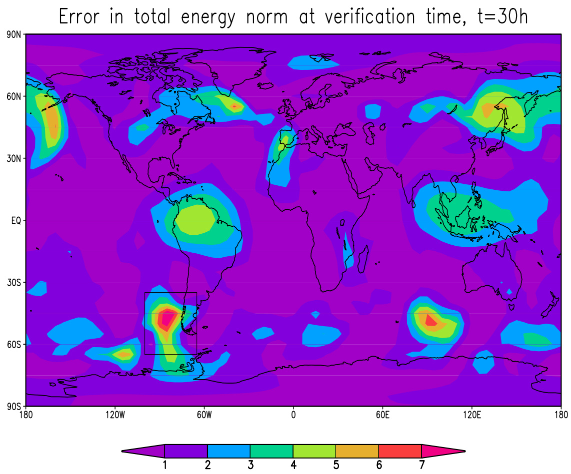


Figure 3: Graph of the forecast error $t = 30$h from the background

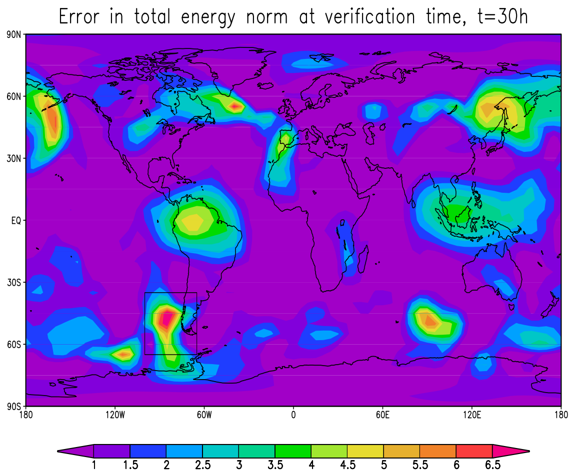


Figure 4: Graph of the forecast error $t = 30$h from the optimal analysis obtained by using only routine observations

using optimal initial condition x_0^a with routine plus adaptive observations are displayed in Figure 6. The algorithm to compute the adjoint sensitivity are given below:–

Algorithm 1

- Calculate model solution x_v^t at t_v with initial condition (true atmospheric state) x_0^t by

$$x_v^t = \mathcal{M}_{0,v}(x_0^t) \tag{4.1}$$

- Obtain optimal initial condition x_0^a by minimizing the cost functional $\mathcal{J}$ defined in (2.1) with only routine observations. Calculate model forecast

$$x_v^f = \mathcal{M}_{0,v}(x_0^a) \tag{4.2}$$

- Compute $\nabla_{x_v}\mathcal{J}^v = P^T EP(x_v^t - x_v^f)$ and use it as initial condition for adjoint model.
- Integrate Adjoint model backward from t_v to t_i : $\nabla_{x_i}\mathcal{J}^v = \mathbf{M}_{i,v}^T \nabla_{x_v}\mathcal{J}^v$

We then carried out several sequential minimizations of the penalized cost function with the penalty term defined in equation (2.13) aiming to reduce the forecast error to certain minimum level. We attained the minimum forecast error when a very large value of the penalty parameter r was employed. The value of the penalty parameter was adaptively increased based on the value of the cost function $\mathcal{J}_v$ over verification domain where $r_{k+1} = \frac{\mathcal{J}_v(k)}{\mathcal{J}_v(0)} r_k$. In our experiment, the initial value for the penalty parameter is $r = 1$ and this value is sequentially increased based on the value of the $\mathcal{J}_v(k)$ for each call of the unconstrained minimization routine. The algorithm for finding the optimal minimizer of the penalized cost function is provided below. We have found that the minimization routine performed well for a smaller value of the parameter ($r < 10^5$). For the large value of r i.e $r \geq 10^5$, the minimization failed to converge. However, in our experiment we have found that the forecast error was reduced significantly which means that the initial condition is estimated optimally by adding the penalty term to the cost function. The forecast error computed with the optimal initial condition is displayed in Figure 7.

Algorithm 2

1. Initialization: $r_0 = 1$,$\beta_0 = 6$
2. Calculation of $\mathcal{J}_v(0)$ with the starting point x_0.
3. Do loop $k = 1, 2, ...$
4. Forward integration of the forecast model
5. Calculation of $\mathcal{J}_v(k)$
6. Calculation of $\beta_k = \frac{\mathcal{J}_v(0)}{\mathcal{J}_v(k)}$
7. If $\beta_k > 1$ then $r_{k+1} = \beta_k r_k$ else $r_{k+1} = 6r_k$
8. Unconstrained minimization
9. End do

It is well-known that the performance of minimization routine is very sensitive to the different values of penalty parameter. The reason is that the condition number of the Hessian matrix of the cost with respect to the control variables evaluated at the minimum increases as r is getting larger. Moreover, if the initial value of r is too large, it is very difficult to find the minima for any robust unconstrained minimization routine due to the slow convergence induced by the increasingly larger condition number of the Hessian of the penalized cost function. For this reason, we solved the problem of the penalty function sequentially by using the unconstrained minimization routine M1QN3 [22] equivalent to L-BFGS routine, with moderately increasing values of the penalty parameter. We have found that the method performs very well if the penalty parameter is chosen by using a cost function $\mathcal{J}_v$ that is decreasing slowly and consequently the value of parameter is increasing slowly. In our experiment, each successive $x^*(r_k)$ is used as a new starting point for solving an unconstrained minimization problem with the next increased value of the penalty parameter until an acceptable convergence criterion is attained.

5 Summary and Conclusions

In numerical weather prediction, we can reduce the forecast error by optimizing the initial condition. To obtain the optimal initial condition we need to minimize the cost function defined by equation (2.1) which depends on background information and observations. Studies show that only a few adaptive observations included with the existing routine observations can improve weather forecast. Several targeting methods have already been developed. In this paper we use the adjoint sensitivity method to compare our proposed method, the penalized 4D Var DA method. The approach proposed in this paper is able to estimate the initial condition optimally by minimizing the penalized cost function $\mathcal{J}'(x_0)$ defined by (2.13). The evolution of the forecast error obtained by using different approaches are displayed in Figure 8. We have found that the forecast error is reduced significantly by the new approach of employing the penalized 4D Var DA method. From the results obtained we conclude that penalized 4D Var approach performs better than the adjoint sensitivity method. Therefore, the penalized 4D-Var method enables us to obtain the optimal initial condition that provides better forecast than the other method without adding any observations to the existing routine network observations.

Acknowledgments

This research was funded by the National Aeronautics and Space Administration(NASA) (grant no. NNG06GC67G). Logistic support was provided by the Department of Mathematics and Natural Sciences, BRAC University, Dhaka, Bangladesh. The first author would like to thank Prof A.A.Z. Ahmad for his critical reading and making suggestions to correct the manuscript.

References

[1] F.X. LeDimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, 38 A:97–110, 1986.

[2] J. Derber. Variational four-dimensional analysis using quasi-geostrophic constraints. *Mon. Wea. Rev.*, 115:998–1008, 1987.

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

[4] O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. *Q. J. Roy. Meteor. Soc.*, 113:1311–1328, 1987.

[5] J.L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations.* Springer-Verlag, Berlin Heidelberg New York, 1971.

[6] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems.* Springer-Verlag, New York, 1984.

[7] F.X. LeDimet, H.E. Ngodock, B. Luong, and J. Verron. Sensitivity analysis in data assimilation. *J. Meteor. Soc. Japan*, 75:245–255, 1997.

[8] J.N. Thepaut and P. Courtier. Four-dimensional variational assimilation using the adjoint of a multilevel primitive-equation model. *Q. J. Roy. Meteor. Soc.*, 117:1225–1254, 1991.

[9] I. M. Navon, X. Zou, J. Derber, and J. Sela. Variational Data Assimilation with an Adiabatic Version of the NMC Spectral Model. *Mon. Wea. Rev.*, 120:1433–1446, 1992.

[10] M. Zupanski. Regional 4-dimensional variational data assimilation in a quasi-operational forecasting environment. *Mon. Wea. Rev.*, 121:2396–2408, 1993.

[11] J.N. Thepaut, R.N. Hoffman, and P. Courtier. Interactions of dynamics and observations in a 4-dimensional variational assimilation. *Mon. Wea. Rev.*, 121:3393–3414, 1993.

[12] R.H. Langland. Issues in targeted observing. *Q. J. Roy. Meteor. Soc.*, 131:3409–3425, 2005.

[13] D.N. Daescu and I.M. Navon. Adaptive observations in the context of 4D-Var data assimilation. *Meteorol. Atmos. Phys.*, 85:205–226, 2004.

[14] T.N. Palmer, R. Gelaro, J. Barkmeijer, and R. Buizza. Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, 55:633–653, 1998.

[15] J. Barkmeijer, M. Van Gijzen, and F. Bouttier. Singular vectors and estimates of the analysis-error covariance metric. *Q. J. Roy. Meteor. Soc.*, 124:1695–1713, 1998.

[16] J. Barkmeijer, R. Buizza, and T.N. Palmer. 3D-Var Hessian singular vectors and their potential use in the ECMWF Ensemble Prediction System. *Q. J. Roy. Meteor. Soc.*, 125:2333–2351, 1999.

[17] M. Ehrendorfer and J. J. Tribbia. Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.*, 54:286–313, 1997.

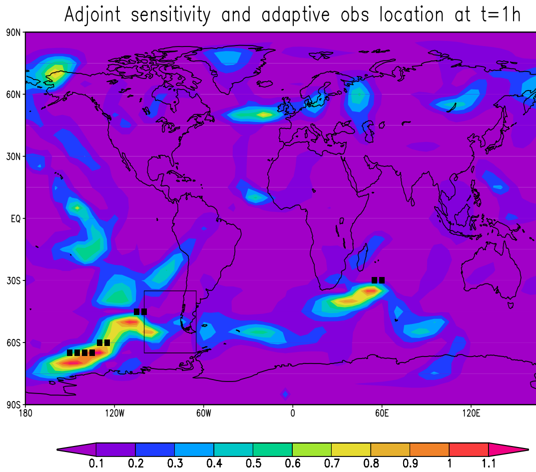
[18] F.X. LeDimet, I.M. Navon, and D.N. Daescu. Second order information in data assimilation. *Mon. Wea. Rev.*, 130(3):629–648, 2002.

[19] Humberto C. Godinez and Dacian N. Daescu. Observation tarteting with a second-order adjoint method for increased predictability. *Computational Geosciences(accepted)*, 2010.

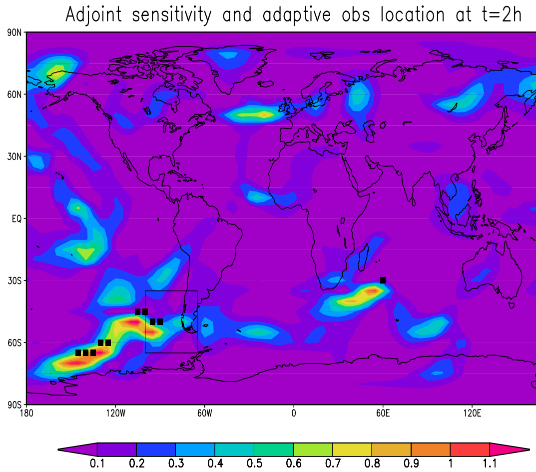
[20] D.F. Parrish and J.D. Derber. The national meteorological center spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, 120:1747–1763, 1992.

[21] D.P. Bertsekas. *Constrained optimization and Lagrange multiplier methods, in Computer Science and Applied Mathematics, Series(ed. Werner Rheinboldt).* Academic Press, Inc., 111 Fifth Avenue, New York, New York 10003, 1982.

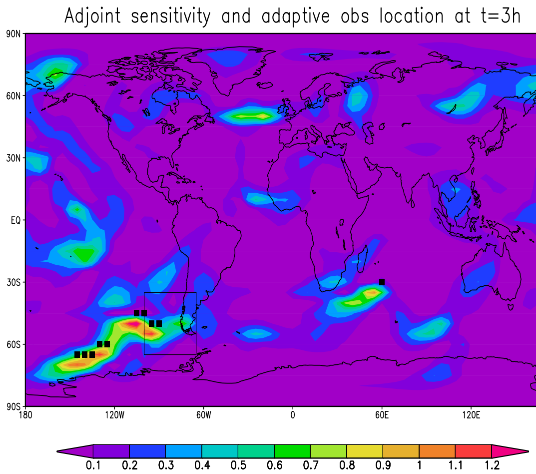
[22] J.Ch. Gilbert and C. Lemarećhal. Some numerical experiments with variable-storage quasinewton algorithms. *Mathematical Programming*, 45:407–435, 1989.



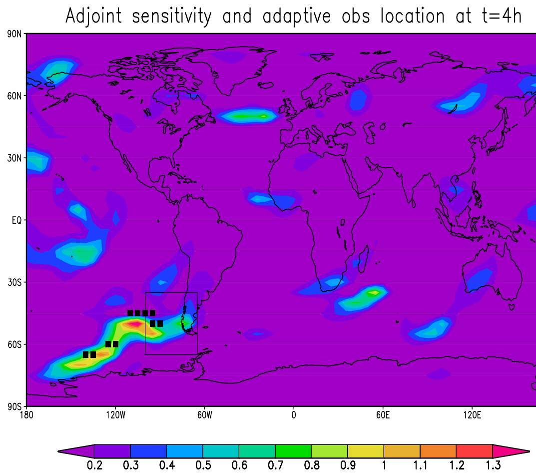
(a) t=1h



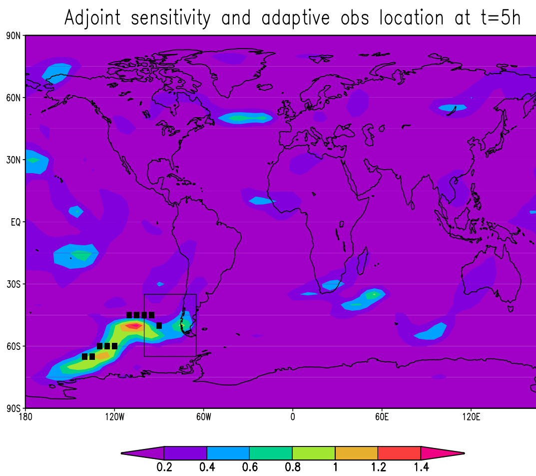
(b) t=2h



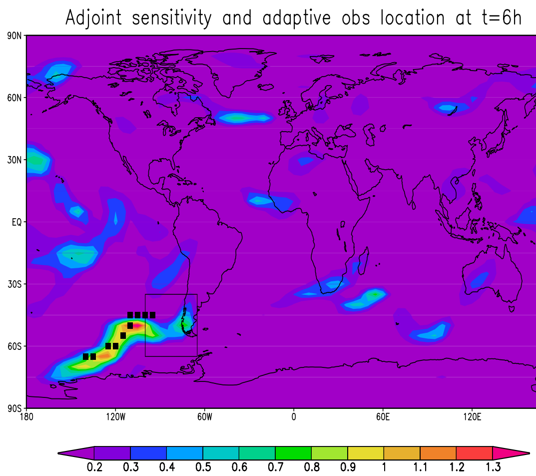
(c) t=3h



(d) t=4h



(e) t=5h



(f) t=6h

Figure 5: Time evolution of the sensitivity field corresponding to Adjoint sensitivity method. The location of the adaptive observation at target instant t_i over assimilation window is marked with "□".

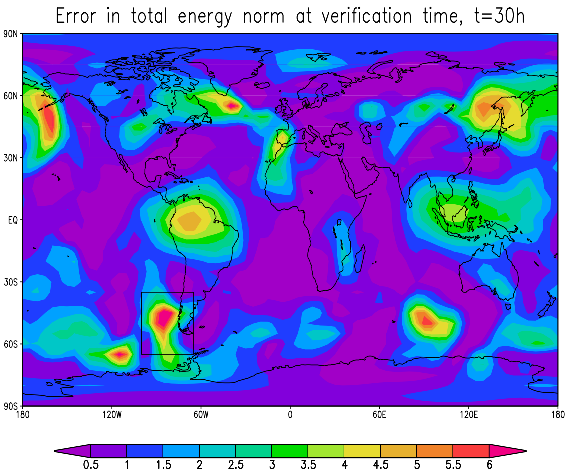


Figure 6: Graph of the forecast error $t = 30$h from the optimal analysis obtained by using routine plus adaptive observations

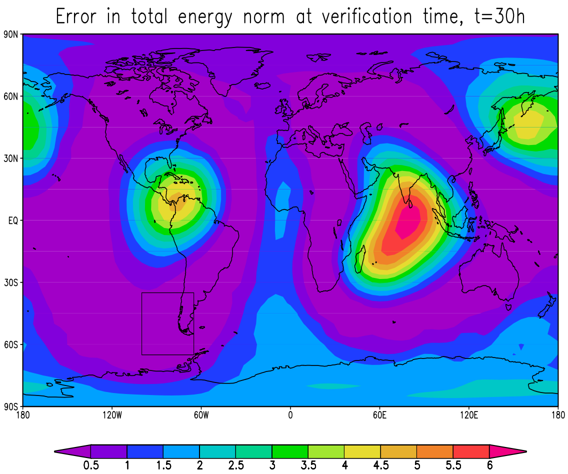


Figure 7: Graph of the forecast error $t = 30$h with the optimal analysis x_0^p obtained by minimizing the penalized cost function. In the figure we see that there is no forecast error over the verification domain if the optimal analysis x_0^p used as initial condition for the forecast model.

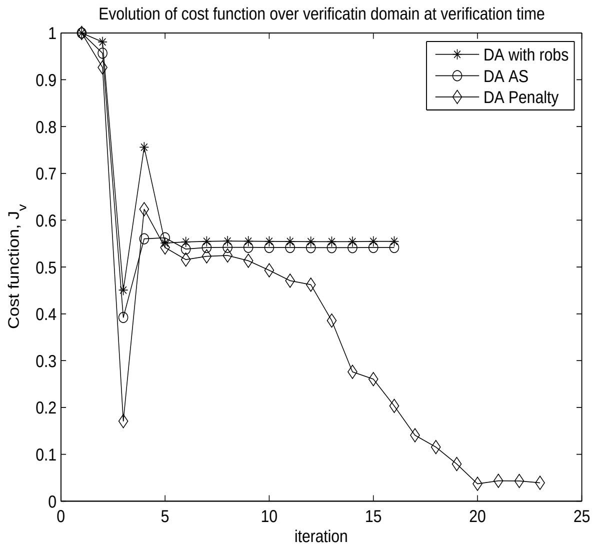


Figure 8: The normalized forecast error over verification domain at the verification time t_v