Tellus (1998), 50A, 557–572
Printed in UK – all rights reserved

TELLUS
ISSN 0280–6495

A validation of the incremental formulation of 4D variational data assimilation in a nonlinear barotropic flow

By STÉPHANE LAROCHE* and PIERRE GAUTHIER, *Meteorological Research Branch, Atmospheric Environment Service, Dorval (Québec) Canada H9P 1J3*

(Manuscript received 3 April 1998; in final form 16 July 1998)

ABSTRACT

In order to meet current operational limitations, the incremental approach is being used to reduce the computational cost of 4D variational data assimilation (4D-Var). In the incremental 4D-Var, the tangent linear (TLM) and adjoint of a simplified lower-resolution model are used to describe the time evolution of increments around a trajectory defined by a complete full-resolution model. For nonlinear problems, the trajectory needs to be updated regularly by integrating the full-resolution model during the minimization. These are referred to as outer iterations (or updates) by opposition to inner iterations done with the simpler TLM and adjoint models to minimize a local quadratic approximation to the actual cost function. In this study, the role of the inner and outer iterations is investigated in relation to the convergence properties as well as to the interactions between the large (resolved by both models) and small scale components of the flow. A 2D barotropic non-divergent model on a β-plane is used at two different resolutions to define the complete and simpler models. Our results show that it is necessary to have a minimal number of updates of the trajectory for the incremental 4D-Var to converge reasonably well. To assess the impact of restricting the gradient to its large scale components, experiments are carried out with a so-called truncated 4D-Var in which the complete model is used to compute the gradient which is truncated afterwards to retain only those components used in the incremental 4D-Var. A comparison between the truncated and incremental 4D-Var shows that the large-scale components of the gradient are well approximated by the lower resolution model. With frequent updates to the trajectory, the incremental 4D-Var converges to an analysis which is close to that obtained with the truncated 4D-Var. This conclusion is verified when perfect observations with a complete spatial and temporal coverage are used or when they are restricted to be available at a coarser resolution (in space and time) than that of the model. Finally, unbiased observational error was introduced and the results showed that at some point, the minimization is overfitting the observations and degrades the analysis. In this context, a criterion related to the level of observational noise is found to determine when to stop the minimization when the complete 4D-Var is used. This criterion does not hold however for the incremental and truncated 4D-Var, thereby indicating that it may be very difficult to establish in a more realistic context when the error is biased and the model itself is introducing a biased error. The analysis and forecasts from the incremental 4D-Var compare well to those from a full-resolution 4D-Var and are more accurate than those obtained from a low-resolution 4D-Var that uses only the simplified model.

1. Introduction

One of the reasons why the variational approach is appealing is first, that it makes it possible to assimilate new data sources that are only indirectly related to the model variables. This is the case for instance for TOVS radiances (Eyre et al., 1993; Andersson et al., 1994), Doppler radar data (Sun et al., 1991) or radio refractivity (Eyre, 1994; Zou et al., 1995). This is of course possible

* Corresponding author.

within a 3D-Var analysis (Parrish and Derber, 1992 Courtier et al., 1998; Gauthier et al., 1996) but a 4-dimensional variational assimilation (4D-Var) adds also the benefit of having an analysis that is dynamically consistent with the prediction model while being also as close as possible to the observations. Moreover, Thépaut and Courtier (1991), Rabier and Courtier (1992) and Tanguay et al. (1995) (TBG95 hereafter) showed that 4D-Var is able to infer information in the small-scales from a time sequence of large-scale observations. In TBG95, it is shown that there exists an optimal assimilation period for which a nonlinear transfer of large-scale information can fill in fine scale details below the resolution of the observational network. The optimality of the period comes from the fact that the dynamical constraint requires some time to act but, on the other hand, the assimilation interval T_a cannot be too long since at some point, it becomes impossible to control the fit to an observation by modifying the initial conditions: this is related to the predictability time limit of the model. This ability of inferring information about unobserved features of the flow will become increasingly useful because the observational network is likely to always be of a lower resolution than that of forecast models. It is indeed this ability of 4D-Var to fill in for the deficiencies of the observational network that makes it worth doing despite its high numerical cost.

The cost of 4D-Var is directly proportional to the number of iterations required to reach convergence and on the cost of one iteration. To reduce the number of iterations, preconditioning methods must be employed to speed up the convergence (Zupanski, 1993; Courtier et al., 1994; Fisher and Courtier, 1995; Yang et al., 1996). On the other hand, the cost of one iteration can also be reduced by using a simpler model. Results obtained by Thépaut and Courtier (1991) and TBG95 suggest that the corrections to the initial conditions brought in by 4D-Var are of a lower resolution than the full resolution model. This lead Courtier et al. (1994) (CTH94 hereafter) to propose a strategy, the so-called incremental approach to 4D-Var, that significantly reduces its cost so as to make it possible to envision its operational implementation. Rabier et al. (1997) have used it to develop a 4D-Var system that recently became operational at ECMWF. Their results showed a systematic positive impact on the quality of forecasts, at least with an assimilation period of 6 hrs.

In the present paper, the model introduced in TBG95, the barotropic vorticity equation on the β plane, is used to study the incremental approach of CTH94 in which the tangent linear (TLM) and adjoint of a simplified model (i.e., a low-resolution model in the present case) are used to approximate the time evolution of increments $\delta X(t)$ around a trajectory $X(t)$ obtained from an integration of the complete model. For nonlinear problems, the trajectory needs to be updated regularly by integrating the full model with the initial conditions $X_0 + \delta X_0$ and redefining the TLM so that it still provides a reasonable approximation of the time evolution of the increments. This is referred to as an outer iteration of the minimization by opposition to an inner iteration that uses the TLM and adjoint of the simpler model to minimize the cost function for the increments expressed in terms of departures from the trajectory. For more complex models, CTH94 argued that this procedure allows a progressive inclusion of physical processes. Experiments will be presented to show that it is necessary to have a minimal number of outer iterations for the incremental method to converge reasonably well. In order to assess the impact on the analysis of restricting the gradient to its large scale components, the results from a full-resolution 4D-Var experiment are compared against those obtained from a truncated 4D-Var in which the exact gradient is computed at each iteration and truncated to a lower resolution. These results are then compared against an incremental 4D-Var to assess if the large scale components of the gradient can be well approximated by a simpler lower resolution model as in the incremental approach.

The paper is organized as follows. Section 2 introduces the model used altogether with different formulations of the 4D-Var problem including the incremental approach of CTH94. Section 3 presents results of experiments when perfect observations are provided at every time step. In Section 4, experiments are conducted in which observations are only provided at a coarser resolution (spatial and temporal) and unbiased observational error is added. The results of a full-resolution 4D-Var experiment are compared against those obtained with a truncated 4D-Var and an incremental one

with different number of outer iterations. Conclusions are drawn in Section 5.

2. Description of the model and the incremental 4D-Var

As in TBG95, the nonlinear model used in this study is the barotropic vorticity equation on the β-plane

$$\frac{\partial \zeta}{\partial t} + J(\Psi, \zeta) + \beta v = f - D(\zeta). \tag{2.1}$$

The notations used are as in TBG95. Namely, the streamfunction is $\Psi = -U_0 y + \psi$, the vorticity $\zeta = \nabla^2 \psi$ and the wind components are $u = -\partial\psi/\partial y$, $v = \partial\psi/\partial x$. Moreover, $J(a, b)$ is the Jacobian operator, f is a forcing term and D, a linear dissipation operator. The model is integrated on a doubly periodic domain of length 2π using pseudo-spectral methods (Orszag, 1971). In Fourier space, (2.1) becomes

$$\left[\frac{\partial}{\partial t} + i\omega_{\boldsymbol{k}} + \nu_{\boldsymbol{k}}\right] \hat{\zeta}_{\boldsymbol{k}} = \hat{f}_{\boldsymbol{k}} + \sum_{\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{k}} A_{\boldsymbol{kp}} \hat{\zeta}_{\boldsymbol{p}} \hat{\zeta}_{\boldsymbol{q}}, \tag{2.2}$$

where the caret $\hat{}$ is used to denote a Fourier-space quantity, $\boldsymbol{k} = (k_x, k_y)$ is the wavevector, $k = |\boldsymbol{k}|$, $\omega_{\boldsymbol{k}} = k_x(U_0 - \beta/k^2)$is the linear Rossby-wave frequency, the interaction coefficient $A_{\boldsymbol{kp}} = \hat{\boldsymbol{z}} \cdot \boldsymbol{k} \times \boldsymbol{p}/p^2$, $\hat{f}$ is the forcing and $\nu_{\boldsymbol{k}} = \nu_0 + \nu k^{16}$ represents the dissipation operator D. As in TBG95, the forcing term is set to:

$$\hat{f}_{\boldsymbol{k}} = \alpha \begin{cases} 1, & \text{if } \boldsymbol{k} = (0, \pm k_f) \text{ or } (\pm k_f, 0), \\ 0, & \text{otherwise,} \end{cases}$$

where $\alpha = 0.04$ and $k_f = 3$. A mean zonal wind $U_0 = 0.3$ is also imposed. We used $N = 64$ collocation points per dimension and applied circular Fourier-space truncation at $k_T = (N - 3/2)/3$ to avoid aliasing errors. All model integrations presented in this work use a timestep of $\Delta t = 0.95/k_T$ while $\nu_0 = 0.02$ and $\nu = 8.8/k_T^{16}$. More details on this model as well as for the development of its TLM and adjoint models can be found in TBG95 who discuss also the "climatology" of the model to relate their results to dimensional atmospheric variables. For example, one model time unit (≈ 22 timesteps) corresponds approximately to 0.3 days and the nonlinear turnover timescale is approximately 9 model time units (≈ 3 days).

Here, the model state $\boldsymbol{x}$ is the vorticity in Fourier space ($\hat{\zeta}$) and the observations could be wind components (u, v) defined in physical space at grid point locations or as in TBG95, spectral components of the vorticity defined in Fourier space. In the latter case, the forward model boils down to a projection of the model state onto the observed components while in the former case, H_i for each wind component is obtained from

$$u = F^{-1}\left\{-\frac{\partial}{\partial y}\Delta^{-1}\hat{\zeta}\right\}, \qquad v = F^{-1}\left\{\frac{\partial}{\partial x}\Delta^{-1}\hat{\zeta}\right\},$$

where F^{-1} is the inverse Fourier transform and Δ^{-1} is the inverse Laplacian. A projection in physical space is also part of the forward model.

The 4D-Var method attempts to find initial conditions $\boldsymbol{x}_0 \equiv \boldsymbol{x}(t_0)$ that minimize the distance to a background field $\boldsymbol{x}_b$ and to observations $\boldsymbol{y}_i$ distributed over a finite time interval (t_0, t_N) both contributions being relatively weighted by their respective error covariances. This is achieved by minimizing the functional

$$J(\boldsymbol{x}_0) = \frac{1}{2}(\boldsymbol{x}_0 - \boldsymbol{x}_b)^{\mathrm{T}} \boldsymbol{B}^{-1} (\boldsymbol{x}_0 - \boldsymbol{x}_b) + \frac{1}{2}\sum_{i=0}^{N} [H_i \boldsymbol{x}(t_i) - \boldsymbol{y}_i]^{\mathrm{T}} \boldsymbol{R}_i^{-1} [H_i \boldsymbol{x}(t_i) - \boldsymbol{y}_i], \tag{2.3}$$

where $\boldsymbol{x}_0 = \boldsymbol{x}(t_0)$, $\boldsymbol{x}_b$ is the background state and $\boldsymbol{B}$, its error covariance matrix. The vector $\boldsymbol{y}_i$ stands for the observations at time t_i with $\boldsymbol{R}_i$, the observational error covariance matrix. The observation operator H_i maps the model variables to the observation space and $(t_0, t_1, \ldots, t_N)$ are the observation times. In our case, the observation operator H_i is linear. The model state at time t_i being $\boldsymbol{x}(t_i)$, it is symbolically related to the initial conditions by $\boldsymbol{x}(t_i) = \boldsymbol{M}(t_i, t_0, \boldsymbol{x}_0)$.

Consider now the formulation of the incremental approach of CTH94. Let $\boldsymbol{x}_n(t_i)$ be a reference trajectory obtained from the initial conditions $\boldsymbol{x}_0^{(n)}$ while $\delta\boldsymbol{x}(t_i)$ is the perturbation to this trajectory caused by changes $\delta\boldsymbol{x}_0^{(n)} = \delta\boldsymbol{x}^{(n)}(t_0)$ in the initial conditions. Using the tangent linear approximation, we have that

$$\delta\boldsymbol{x}(t_i) \approx M(t_i, t_0)\delta\boldsymbol{x}_0^{(n)} \equiv M_i^{(n)}\delta\boldsymbol{x}_0^{(n)},$$

where $M(t_i, t_0) \equiv M_i^{(0)}$ is the propagator of the tangent linear model.

Introducing this form in (2.3) yields the follow-

ing quadratic functional in $\delta \boldsymbol{x}_0^{(n)}$:

$$J_n(\delta \boldsymbol{x}_0^{(n)}) = \frac{1}{2}[\delta \boldsymbol{x}_0^{(n)} - (\boldsymbol{x}_b - \boldsymbol{x}_n(t_0))]^T \times \boldsymbol{B}^{-1}[\delta \boldsymbol{x}_0^{(n)} - (\boldsymbol{x}_b - \boldsymbol{x}_n(t_0))] + \frac{1}{2}\sum_{i=0}^{N}[H_i M_i^{(n)} \delta \boldsymbol{x}_0^{(n)} - \boldsymbol{y}_i^{(n)}]^T \times \boldsymbol{R}_i^{-1}[H_i M_i^{(n)} \delta \boldsymbol{x}_0^{(n)} - \boldsymbol{y}_i^{(n)}], \quad (2.4)$$

where $\boldsymbol{y}_i^{(n)} = \boldsymbol{y}_i - H_i \boldsymbol{x}_n(t_i)$.

This procedure can then be seen as a pair of nested loops. First, inner iterations are done to minimize J_n using a descent algorithm such as a conjugate gradient or a quasi-Newton algorithm. Since this leads to a finite amplitude change to the initial conditions, (2.4) may not be an accurate approximation to the original problem (2.3). Outer iterations are then introduced to update the trajectory by integrating the nonlinear model with the initial conditions $\boldsymbol{x}_0^{(n)} + \delta \boldsymbol{x}_0^{(n)}$ to produce a new reference trajectory $\boldsymbol{x}_{n+1}(t_i)$. The functional J_{n+1} is a better local approximation to J. This is schematically depicted in Fig. 1.

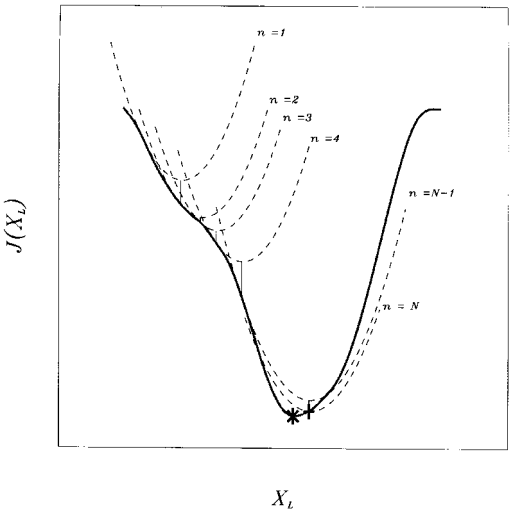

Fig. 1. Schematic representation of the cost function of the full-resolution 4D-Var (thick curve) and the incremental formulation after n updates to the background trajectory (dashed curves) in the subspace defined by the simplified model components X_L. The thick star shows the minimum of the 4D-Var formulation in the X_L direction while the thick cross is the minimum of the incremental formulation after N updates.

The incremental approach also introduces an approximation which is to use the TLM of a simpler model. In this study, the simpler model is a lower resolution version of our model, truncated at wavenumbers such that $k \leqslant k_L < k_H$ with k_H being the truncation used for the complete model. Let $\boldsymbol{x}_L = \pi_L \boldsymbol{x}$ be the projection of the complete model state onto the lower dimensional subspace of the simplified model. The outer iterations are performed by integrating the complete model with the initial conditions $\boldsymbol{x}_n(t_0) + \pi_L^T \delta \boldsymbol{x}_0^{(n)}$ to update the model trajectory to $\boldsymbol{x}_{n+1}(t_i)$. Here π_L^T is the generalized inverse of the projection operator.

Referring to Fig. 1, the thick gray curve represents a cross section of the cost function (2.3) along a subspace direction defined by the simplified model state X_L. The dashed curves are the cost functions defined by (2.4) after each update. These curves are quadratic since both H_i and $M_i^{(n)}$ are linear operators but are not exactly tangent to the thick line because the tangent linear model in (2.4) is obtained from the simplified model while the thick curve is defined with respect to the full-resolution model. After each update of the trajectory, the solution converges to the minimum of the dashed curves. However, we can see that a global convergence is reached when the updated full-resolution trajectory converges after a given number of outer loops. It is easy to see that $\delta \boldsymbol{x}^n(t_0) \to 0$ in this case since $\delta \boldsymbol{x}_0^{(n)}$ is reset to zero after an outer iteration. The minimum reached after global convergence does not necessarily correspond to the minimum of the thick line indicated by the star since the gradient of the cost functions (2.3) and (2.4) along the subspace X_L are different. The solution corresponding to this star can be obtained by minimizing (2.3) in the subspace X_L by using the full-resolution model as a constraint. This is done by retaining only the components of gradient of the cost function in the subspace X_L. This procedure is referred to as the truncated 4D-Var in Section 3 and is used for comparison with the incremental approach.

The above description suggests that the solution obtained with the incremental formulation differs from the full-resolution 4D-Var solution for two reasons. Firstly, the minimum of the full-resolution 4D-Var is in general different from the minimum along X_L subspace (the thick star in Fig.1). Secondly, because the large-scale components in the full-resolution and simplified models are not

exactly represented and predicted in the same way, the minimum after global convergence (the thick cross in Fig. 1) does not correspond to the one along the X_L subspace.

A finite-difference variant of (2.4) can be obtained by replacing $M_i^{(n)}\delta\boldsymbol{x}_0^{(n)}$ by

$$\begin{aligned} M_i^{(n)}\delta\boldsymbol{x}_0^{(n)} &= \mathscr{L}(t_i, t_0, \pi_L\boldsymbol{x}_n(t_0) + \delta\boldsymbol{x}_0^{(n)}) \\ &\quad - \mathscr{L}(t_i, t_0, \pi_L\boldsymbol{x}_n(t_0)) \\ &= \boldsymbol{x}_L^{\delta}(t_i) - \boldsymbol{x}_L(t_i), \end{aligned} \tag{2.5}$$

where $\mathscr{L}(t_i, t_0, \pi_L\boldsymbol{x}_0)$ stands for the integration of the simplified model including the nonlinear terms. It will be implicit from now on that $\delta\boldsymbol{x}_0^{(n)} \equiv \pi_L\delta\boldsymbol{x}_0^{(n)}$ and that $M_i^{(n)}$ is the resolvent of the tangent linear model of the simplified model. Eq. (2.4) can then be rewritten as

$$\begin{aligned} J_n(\delta\boldsymbol{x}_0^{(n)}) &= \frac{1}{2}[\delta\boldsymbol{x}_0^{(n)} - (\boldsymbol{x}_b - \boldsymbol{x}_n(t_0))]^T \\ &\quad \times \boldsymbol{B}^{-1}[\delta\boldsymbol{x}_0^{(n)} - (\boldsymbol{x}_b - \boldsymbol{x}_n(t_0))] \\ &\quad + \frac{1}{2}\sum_{i=0}^{N}[H_i\boldsymbol{x}_L^{\delta}(t_i) - (\boldsymbol{y}_i^{(n)} + H_i\boldsymbol{x}_L(t_i))]^T \\ &\quad \times \boldsymbol{R}_i^{-1}[H_i\boldsymbol{x}_L^{\delta}(t_i) - (\boldsymbol{y}_i^{(n)} + H_i\boldsymbol{x}_L(t_i))]. \end{aligned} \tag{2.6}$$

As pointed out by CTH94, the practical advantage of this formulation over (2.4) is that less technical development is required once the 4D-Var problem has been implemented. The incremental formulation (2.6) is used for all the experiments presented in this work. Finally, our experiments leave aside the background term and focus on obtaining initial conditions that best fit observations spread over a finite time interval. In the next section, we investigate if it is possible to achieve convergence by only retaining the large scale components of the gradient and study the impact of the outer iterations.

3. Relative rôle of the inner and outer iterations

One of our objectives is to study the role played by the nonlinear interactions between scales in the incremental approach. Since these interactions intervene only in the second term on the right-hand-side of (2.4) and (2.6), we set $\boldsymbol{B}^{-1} = 0$ and $\boldsymbol{R} = \boldsymbol{I}$, the identity matrix so no background term is considered and the observational error is taken to be the same for all observations. Under these assumptions, the cost function (2.3) of the 4D-Var boils down to

$$J(\boldsymbol{x}_0) = \frac{1}{2}\sum_{i=0}^{N}[H_i\boldsymbol{x}(t_i) - \boldsymbol{y}_i]^T[H_i\boldsymbol{x}(t_i) - \boldsymbol{y}_i], \tag{3.1}$$

while for the incremental formulation, we get

$$\begin{aligned} J_n(\delta\boldsymbol{x}_0^{(n)}) &= \frac{1}{2}\sum_{i=0}^{N}[H_iM_i^{(n)}\delta\boldsymbol{x}_0^{(n)} - \boldsymbol{y}_i^{(n)}]^T \\ &\quad \times [H_iM_i^{(n)}\delta\boldsymbol{x}_0^{(n)} - \boldsymbol{y}_i^{(n)}]. \end{aligned} \tag{3.2}$$

The components of $\delta\boldsymbol{x}_0^{(n)}$ are defined in terms of changes to the spectral components of vorticity. In the incremental experiments, the Fourier components with a total wavenumber $k > k_L$ are never updated and the high resolution trajectory is obtained by integrating the initial conditions $\hat{\zeta}_k^{(n)}(t_0) = \hat{\zeta}_k^{(n-1)}(t_0) + \delta\hat{\zeta}_k^{(n)}(t_0)$ when $k < k_L$ and $\hat{\zeta}_k^{(n)}(t_0) = \hat{\zeta}_k^{G}(t_0)$ when $k_L < k < k_H$, ζ_k^G being the first-guess.

As in TBG95, the model was spun up from random initial conditions and integrated for 500 time units to make sure that the flow is statistically stationary and ergodic. Synthetic observations of the wind components (u, v) were generated for an assimilation period $T_a = t_N - t_0$ of 10 time units. Fig. 2 shows the vorticity fields at t_0 and t_N which represent the control fields. Unless otherwise specified, the first guess was obtained from a Cressman analysis (Cressman, 1959; Daley, 1991) of observations of u and v available at t_0 at every 8 grid points in both Cartesian directions. A radius of influence of 9 grid points was employed to perform the analysis. From this analysis, the vorticity was computed using a finite difference approximation and then mapped into the Fourier space via discrete Fourier transforms. The Cressman analysis provides a vorticity field noticeably different from the corresponding control run as shown in Fig. 3, but not completely decorrelated as the first guess used by TBG95. Consequently, a fewer number of iterations is required to reach convergence.

In all experiments presented in this paper, the minimization was performed with the variable storage quasi-Newton algorithm (M1QN3) of Gilbert and Lemaréchal (1989) that requires on average only one simulation (the evaluation of the cost function and its gradient) per iteration. All results presented here are in terms of number of

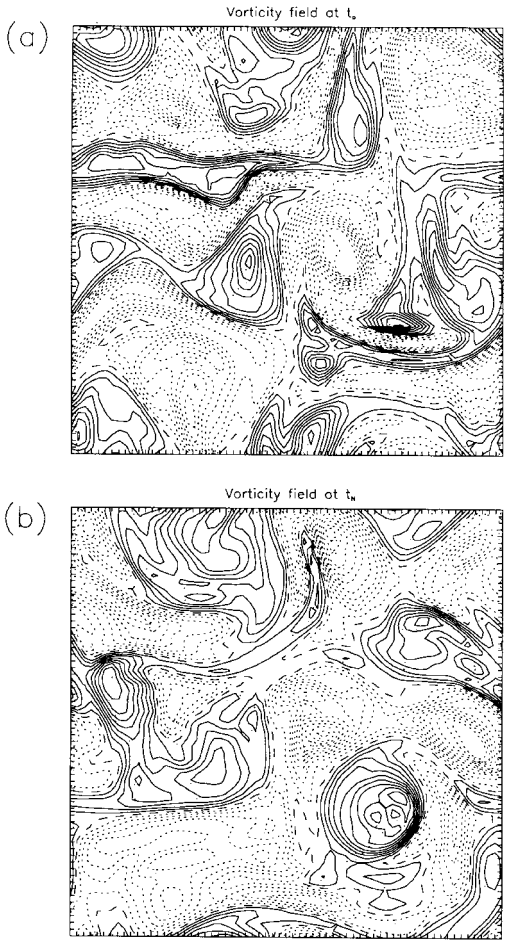

Fig. 2. Vorticity fields of the control run at the beginning (a) and the end (b) of the assimilation period.

simulations which is proportional to the computational effort. To retain the information on the Hessian matrix from one outer iteration to another, the warm restart procedure of M1QN3 is used after each update of the model trajectory as described by CTH94.

The high resolution model is truncated at $k_T = 20$ and ($k_H = k_T$). The simplified model is the same model but truncated at a lower resolution $k_L = 5$ which is 4 times smaller than the full-resolution model. We found that $k_L = 5$ is the best choice to enhance the interactions between the resolved and unresolved scales. This may be seen as an acid test since the resolutions used by CTH94 (i.e. T42 for the simplified model and T63 for the full-resolution model) were much closer. However, this is in line with the ECMWF operational implementation of 4D-Var (Rabier et al., 1997).

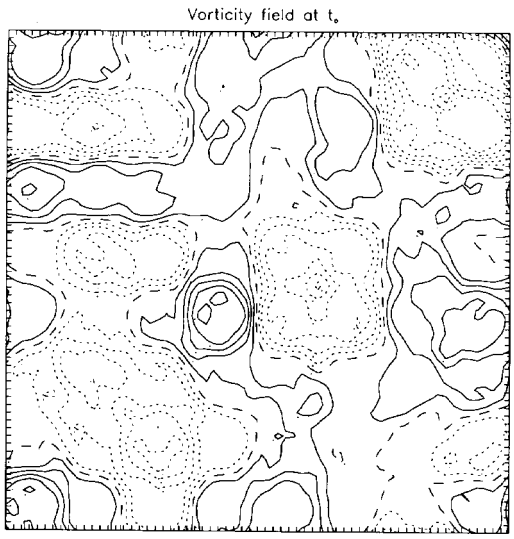

Fig. 3. Cressman analysis of the vorticity field used as a first guess.

3.1. Results obtained without updating the trajectory

First, the convergence properties of the incremental approach are studied when the reference trajectory is not updated, in other words, when only one "outer" iteration is performed. In the following experiments, observations are supplied at all grid points at all time steps (i.e. perfect observations). Fig. 4 shows the energy spectra of the error field (defined as the difference between the vorticity obtained after a given number of simulations and the control run) at the beginning t_0 and at the end t_N of the assimilation period. The spectrum at t_N is obtained by integrating the full-resolution model from initial conditions in which Fourier coefficients less and equal to k_L are given by the assimilation while coefficients greater than k_L are from the background. The spectra obtained after 80 and 160 simulations are very close which means that the minimization algorithm converges. However, the best solution at t_N for all wavenumbers is not obtained after convergence (i.e., after 160 simulations) but rather after 20 simulations. This result is somewhat disturbing

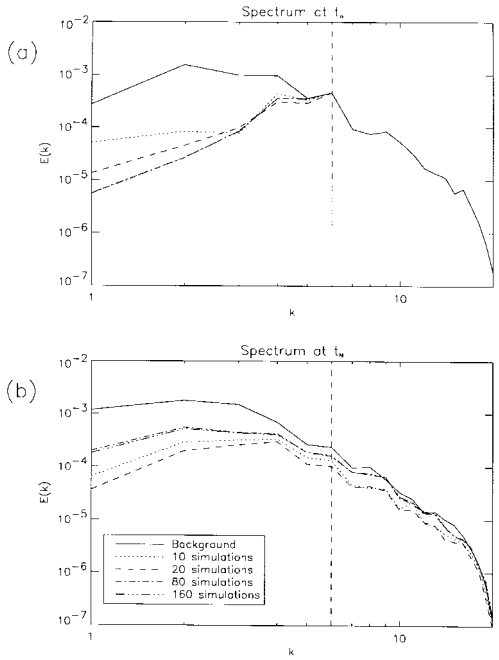

Fig. 4. Error spectra at the beginning (a)and the end (b) of the assimilation period for the incremental approach without update.

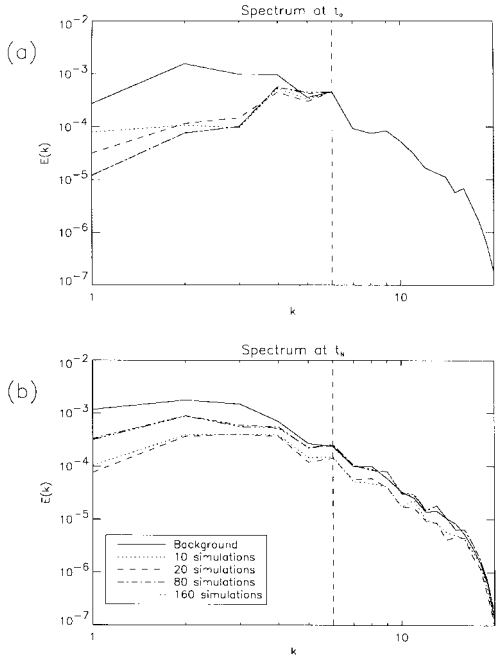

Fig. 5. Error spectra at the beginning (a) and the end (b) of the assimilation period for the low-resolution 4D-Var formulation.

because, in an operational context, the truth is not known and thus it may be difficult to find the optimal number of simulations between each update if convergence is not a suitable criterion. CTH94 used 10 simulations between each update which is probably not the optimum number of simulations. Nevertheless, we will see that it is preferable to choose a lower number of inner iterations and refresh more frequently the trajectory.

The computational efficiency of the incremental approach is comparable to that of a 4D-Var totally formulated in terms of the simpler model and its adjoint (referred to as a low-resolution 4D-Var). A supplementary computational effort is however required at each outer iteration when the complete model has to be integrated to update the reference trajectory. The final full-resolution analysis at t_0 from the low-resolution 4D-Var formulation can be obtained by adding the small-scale components of the background field to the large-scale components determined during the minimization. The full-resolution model is then integrated from this analysis to obtain the solution at the end of assimilation period. By comparison, the energy spectra of the error field obtained with the low-resolution 4D-Var are shown on Fig. 5 as a function of the number of simulations. Both at the beginning and the end of the assimilation, the best result is also obtained after 20 simulations, but it is not as good as the one obtained with the incremental approach even with a single outer iteration. By comparing the results of these two experiments, we come to the conclusion that using the large scale components of the trajectory of the model at its full resolution has a positive impact on the analysis.

3.2. *Impact of the outer iterations*

Without any outer iteration, the previous experiments have shown that the best solution is obtained after 20 simulations, a number of simulations that is likely to vary with the local structure of the cost function. There is no point in accurately minimizing a quadratic functional that does not fit well the actual cost function. As discussed earlier, the role of the outer iterations is to obtain

a better quadratic approximation of the cost function.

Next, experiments were performed with respectively 1,4, 8 and 12 outer iterations with 20 simulations used between each update for the inner iterations. Fig. 6 shows the energy spectra of the error field that indicate that the solution has converged after 12 outer loops which has also the largest number of simulations. The outer loops have improved the analysis (at $t = t_0$ and $t = t_N$) at all wavenumbers, the gain being more important for the larger scales at $t = t_N$. A comparison with the results of Fig. 4 shows that, for an equal number of simulations, the outer iterations bring a significant improvement to the analysis.

Keeping now the total number of simulations fixed, the number of outer iterations has been varied. It is important to emphasize that the incremental 4D-Var introduces two approximations to the problem. First, the minimization is restricted to perform the minimization within a lower dimensional subspace by considering only the large scale components of the gradient. Second, the large scale components of the gradient are in turn approximated by using a simpler model. To make a distinction between these two approximations, we have added an experiment in which the full model is used to compute the gradient but only its large scale components are kept in the minimization: this experiment is based on the truncated 4D-Var. Fig. 7 shows the normalized true cost function obtained from the truncated 4D-Var, the incremental 4D-Var when the background trajectory is not updated (referred as to incremental 1 loop) and when the trajectory is updated every 5, 10 and 20 simulations. The lowest value of the cost function and the fastest rate of convergence is obtained with the truncated 4D-Var, as expected from Fig. 1. The reduction of the cost function becomes negligible after 30 simulations for the truncated 4D-Var and after 60 simulations for the incremental formulation, so that the rate of convergence is about two times slower for the incremental approach. There is a significant reduction of the cost function when the trajectory is updated, especially in the first 40 simulations. This indicates that the trajectory should be updated more frequently at the beginning of the minimization. Fig. 8 represents the error at the end of the assimilation interval as a function of the number of inner iterations. It shows that the best results were obtained when 10 simula-

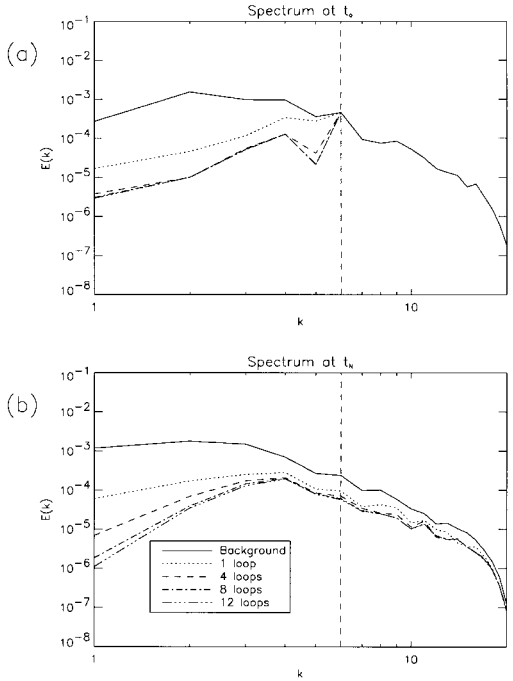

Fig. 6. Error spectra at the beginning (a) and the end (b) of the assimilation period for the incremental approach with updates. 20 simulations are used between each update.

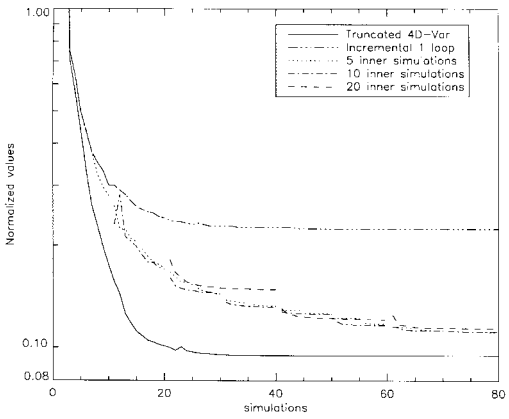

Fig. 7. Variation of the cost functions with the number of simulations obtained with the truncated 4D-Var, the incremental without update (1 loop) and with updates when 5, 10 and 20 inner simulations are used between each outer iteration.

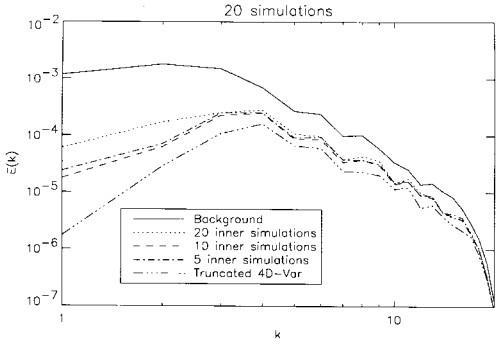

Fig. 8. Error spectra at the end of the assimilation period after a total of 20 simulations when 5, 10 and 20 inner simulations are used.

tions are used between each update. Going back to Fig. 7, we observe that when too many inner iterations are performed, we minimize more accurately a quadratic functional that differs from the true one which is a waste of effort. When an outer iteration is performed, the variational problem is redefined and this shows up on Fig. 7 as jumps in the values of the functional. This result is consistent with the schematic view of the minimization process given by Fig. 1. Finally, the fact that 10 inner iterations outperform 5 could be attributed to the better approximation of the Hessian matrix obtained through the quasi-Newton algorithm. Five iterations may not be enough to obtain a valuable estimate of the Hessian.

The fact that the truncated 4D-Var leads to better results may be attributed to the fact that, at all stages, it uses the true gradient to minimize the true cost function while only allowing changes to the control variable in its large scale components. The incremental 4D-Var on the other hand has to work with a quadratic approximation to the functional and the large scale components are approximated by integrating a lower resolution model. To what extent are the large scale components of the gradient correctly resolved by the truncated model? Considering a given point in phase space, the gradient has been computed with the full and truncated ($k_T = 5$) models. To compare these two gradients as a function of spatial scales, both were expressed as

$$\nabla J(\boldsymbol{x}) = \sum_{k=1}^{K_T} \nabla J_k, \tag{3.3}$$

where

$$\nabla J_k = \sum_{k-\frac{1}{2}<|\boldsymbol{k}'|\leqslant k+\frac{1}{2}} \frac{\partial J}{\partial x_{\boldsymbol{k}'}} \boldsymbol{e}_{\boldsymbol{k}'},$$

where $\boldsymbol{e}_{\boldsymbol{k}'}$ stands for the unit vector along the axis indexed as $\boldsymbol{k}'$ in phase space. The vector ∇J_k then regroups all components having the same spatial scale. The correlation between two components is then defined as

$$C = \cos\theta_k = (\langle \nabla J_k^{(t)}, \nabla J_k^{(T)} \rangle)(\|\nabla J_k^{(t)}\| \, \|\nabla J_k^{(T)}\|)^{-1}, \tag{3.4}$$

where the superscripts ^{t} and ^{T} stands respectively for the true gradient and the one estimated with the truncated model. Table 1 gives the values of the correlation angle θ_k and of the norms of the true and approximate gradients. These have been computed at each outer iteration at the same point in phase space obtained from the incremental experiment. At the first iteration, the truncated model is giving a good approximation of the gradient with a slight degradation as the truncation limit ($k_T = 5$) is approached. As expected, the norm of the approximate gradient is being constantly reduced but very soon (i.e., after the 3rd iteration), the two gradients start to point in different directions. Referring to Fig. 7, the incremental 4D-Var has converged after 8 outer iterations (which corresponds to 80 simulations) but Table 1 indicates that the two gradients become nearly perpendicular. At the twelfth outer iteration, the norm of the approximate gradient differs from the true one by an order of magnitude. This indicates that a truncated 4D-Var would have followed a different path and would have not considered the minimization to have converged at this particular point.

By construction, at each outer iteration, the model is integrated with the small scales of its initial conditions set to those of the first-guess. The results are then dependent on the initial point of the minimization which provides the small scales of the solution. To illustrate this dependency on the choice of the initial point, the initial small-scales components were set to those of the control run. In that case, the solution for the large-scale components converges very close to the truth (result not shown). On the other hand, setting the initial large-scale components to the true solution (i.e., perfect low components) and the small scale components to the first guess used in our experi-

Table 1. *Angle (degrees) between gradient directions of the incremental and full-resolution 4D-Var formulation and corresponding normalized norm for various wavenumbers and intervals*

	θ_k					θ	$\|\nabla J^{(T)}\|$	$\|\nabla J^{(t)}\|$
Loop	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=\{1,5\}$	$k=\{1,5\}$	$k=\{1,5\}$
1	10	12	11	25	32	14	0.909	1.000
2	12	14	18	38	43	17	0.349	0.439
3	86	17	26	41	50	46	0.092	0.130
4	47	26	47	41	45	47	0.079	0.112
5	33	31	62	56	53	39	0.060	0.147
6	37	21	43	55	56	38	0.078	0.146
7	26	29	46	53	41	34	0.080	0.148
8	53	45	44	80	69	55	0.051	0.096
9	87	66	67	91	77	84	0.029	0.066
10	76	60	43	85	69	73	0.026	0.079
11	87	66	51	90	67	83	0.023	0.074
12	69	88	90	84	100	69	0.007	0.077

ments, the solution at the end of the assimilation period is not as good as those obtained with the incremental or the truncated 4D-Var, as depicted by the thick curve on Fig. 9. This shows that it is not enough to determine perfectly the resolved components which is in agreement with the results presented in TBG95. The incremental and truncated 4D-Var both seek the best combination of resolved components (which can be varied) and unresolved components (unchanged during the minimization) such that the model trajectory fits to the best the observations during the assimilation period. To assess the quality of the analysis and the resulting forecasts, we use the correlation between the control vorticity field ($\zeta^{(t)}$) and those obtained from the various experiments (ζ). It is defined as

$$C = (\langle \zeta^{(t)}, \zeta \rangle)(\|\zeta^{(t)}\| \, \|\zeta\|)^{-1}. \tag{3.5}$$

For the different experiments conducted in this section, Fig. 10 shows the values of C as a function of time during 40 time units, the first 10 time units being the assimilation period and the subsequent 30 time units representing the resulting forecasts. The incremental approach with 12 outer loops and the truncated 4D-Var are quite comparable and correspond to the best result. The solution is better correlated with the control at the end of the assimilation period. The error at the beginning of the assimilation period is mainly due to the specified small-scale components from the background field. However, the large-scale components obtained through the minimization with the incremental formulation lead to a continuous improvement during which nonlinear information is transferred in time from the large to smaller scales. This is also true when we extend the assimilation

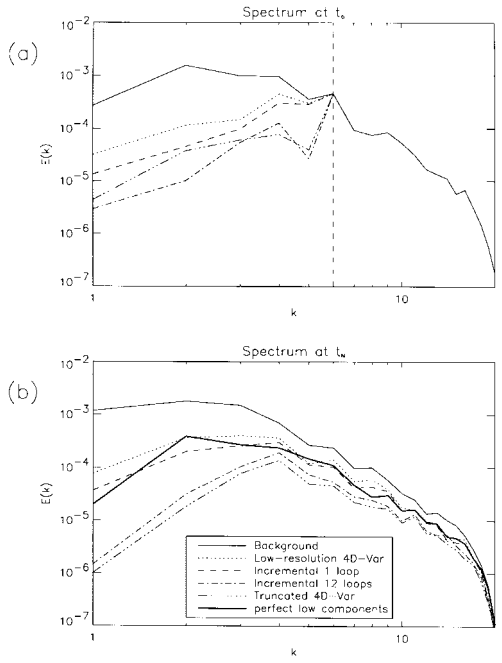

Fig. 9. Error spectra at the beginning (a) and the end (b) of the assimilation period for the various variational formulations when perfect observations are used.

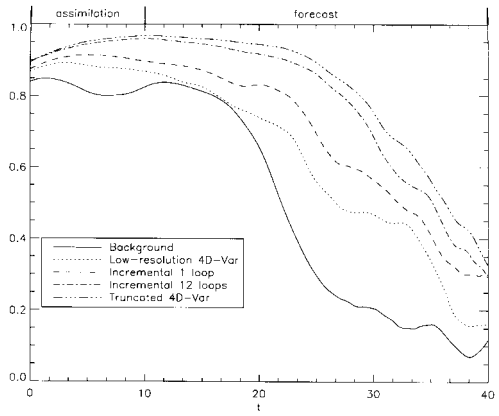

Fig. 10. Correlation between the control vorticity field and those obtained for the various variational formulations when perfect observations are used.

period to 20 time units (results not shown). These results show that in this case, the outer iterations are bringing a significant improvement to the results.

The RMS errors on the vorticity fields obtained at $t = t_N$ with the truncated 4D-Var and incremental formulation are 0.141 and 0.157 (in non dimensional units) respectively. Thus, the error due to the minimization along the subspace defined by the large-scale components account for 90% of the total error while only 10% of the total error is due to the error introduced by the simplified model (second source of error).

Given that a full-resolution 4D-Var converges to zero for this experiment, the next section will repeat these experiments in a more realistic context in which only the large scale part of the flow is observed and when observational error is present.

4. Low-resolution observational network with observational error

So far, the incremental 4D-Var has been looked at in a context where perfect observations are available at every grid point and time step. However, the impact of the resolution of the observational network on the approach is important since the operational observational network has a much lower resolution in space and time than operational NWP models. In the following experiments, observations are supplied at every 8 grid points in both directions and at every 1.66 time units. According to the scaling arguments of TBG95, the distance between observations corresponds to 875 km and the time interval to 12 hours in a real atmosphere. The resolution of this observational network is somewhat lower than the global observing system but remains representative.

With perfect observations (no observational error), the full 4D-Var yields a solution that is very close to the truth at $t = t_N$ while being quite different at the beginning of the assimilation period (results not shown). These results are similar to those presented in TBG95. As discussed in Pires et al. (1996), this can be explained by the fact that initially the difference lies along the stable manifold, a component which is quickly damped during the integration. Fig. 11 shows the energy spectra of the error field at the beginning and end of the assimilation period for the various formulations studied in this paper. A comparison of the error spectra at $t = t_N$ shows that the incremental

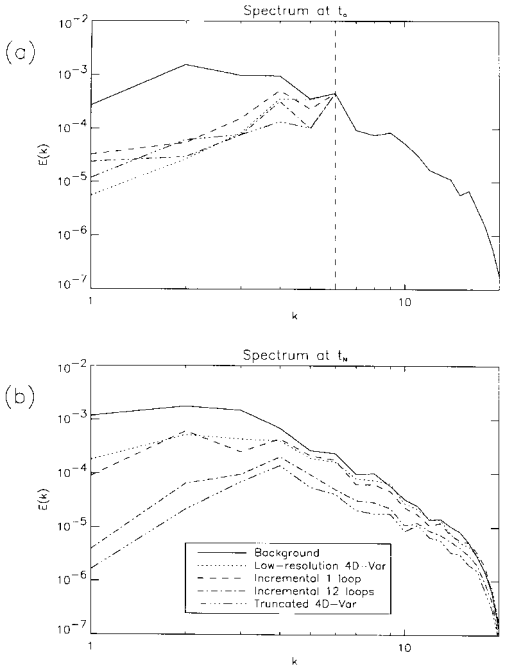

Fig. 11. Error spectra at the beginning (a) and the end (b) of the assimilation period for the various variational formulations when observations are distributed in a low-resolution observation network.

4D-Var with 12 outer iterations (or loops) recovers a solution that is fairly close to the result obtained with the truncated 4D-Var which still gives the best result. These results confirm those presented in the previous section.

Next, the impact of observational error was investigated by adding a random unbiased error with a Gaussian distribution to the observations. Its error variance was chosen as representative of what is found operationally (i.e., 10% of the average amplitude). The matrix $\boldsymbol{R}$ is taken to be diagonal so that observational error correlations are not considered here. The assimilation was first performed by using the full-resolution 4D-Var and Fig. 12 shows the error spectra as a function of the number of simulations. The best solution is obtained after 20 simulations and degrades afterwards, particularly in the small-scales at the beginning of the assimilation period. The explanation of such a behavior lies with the fact that degrees of freedom are adjusted to fit the observational noise, a phenomenon reported by Courtier and Talagrand (1987) and Thépaut et al. (1993). On the other hand, Gauthier (1992) showed that observational error creates an error on the gradient which makes the minimization to converge to a different point than the true minimum, $\boldsymbol{x}_0^{\mathrm{t}}$. If the observation is expressed as $\boldsymbol{y}_i = \boldsymbol{y}_i^{\mathrm{t}} + \varepsilon_i$, with ε_i an unbiased random error, it can be viewed as one realization of the observation vector. If this is introduced into the functional and an ensemble average is performed on the result, one obtains that, for any point $\boldsymbol{x}_0$ in phase space,

$$\overline{J(\boldsymbol{x}_0)} = J_{\mathrm{t}}(\boldsymbol{x}_0) + \frac{1}{2}\sum_{i=0}^{N} \overline{[\varepsilon_i^{\mathrm{T}}\,\varepsilon_i]^2} \equiv J_{\mathrm{t}}(\boldsymbol{x}_0) + \mathscr{B}^2, \tag{4.1}$$

where the overbar stands for the ensemble average and $J_t(\boldsymbol{x}_0)$ represents the "true functional" defined with respect to the true observations. When, in the course of the minimization, the functional J goes below the noise level $\mathscr{B}^2$, the changes brought afterwards are not significative in that they may or may not bring the analysis closer to the true minimum. Therefore, the results will vary from one realization to another. The solid line on Fig. 13 shows the value of the functional as a function of simulations, the line with short dashes indicating the noise level. The crossing point occurs around 12 simulations and this agrees well

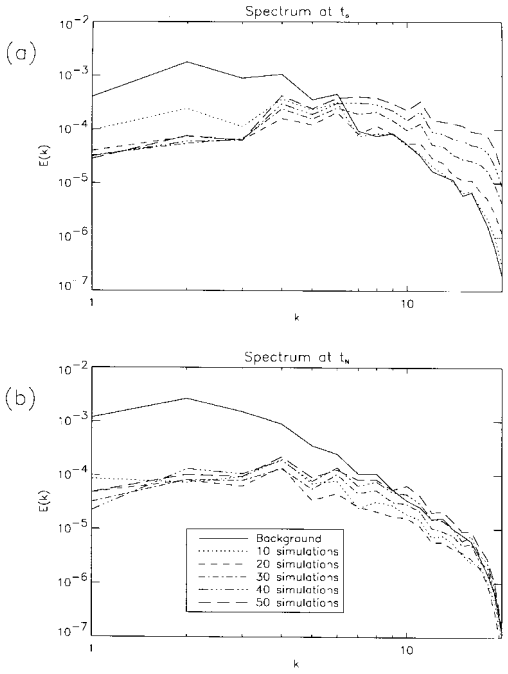

Fig. 12. Error spectra at the beginning (a) and the end (b) of the assimilation period for the full-resolution 4D-Var formulation when random noise is added to the observations in the low-resolution observation network.

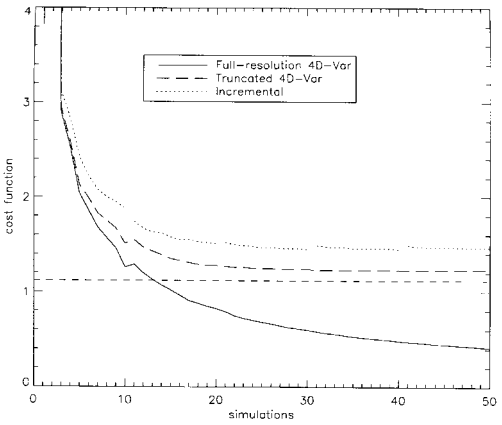

Fig. 13. Variation of the cost function with the number of simulations obtained for the full-resolution 4D-Var (solid line), the truncated 4D-Var (long dashes) and incremental 4D-Var (dotted) when random noise are added to the observations in the low-resolution observation network. The short dashed line indicates the total variance of the random noise.

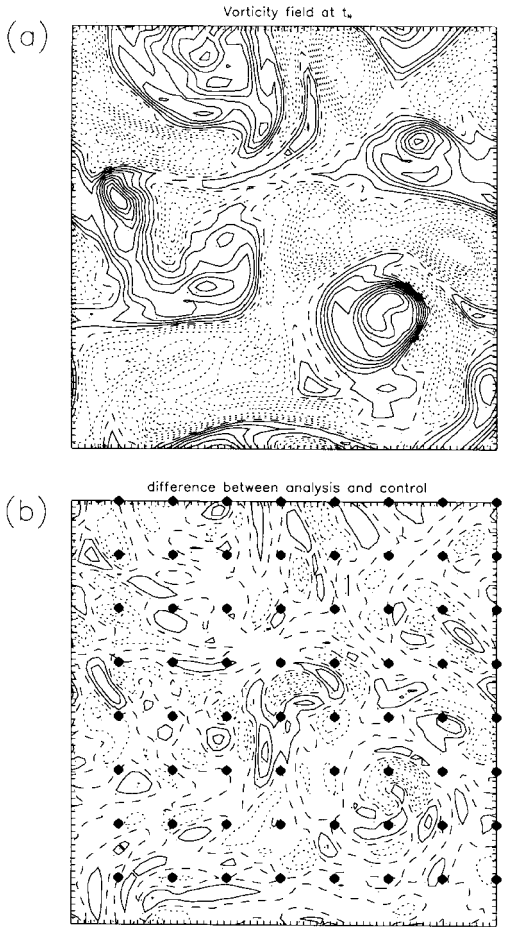

Fig. 14. (a) Vorticity field at the end of the assimilation period obtained with the full-resolution 4D-Var formulation after 20 simulations when random noise is added to the observations in the low-resolution observation network. (b) Difference between (a) and the control run vorticity field. The large dots indicate the position of the observations.

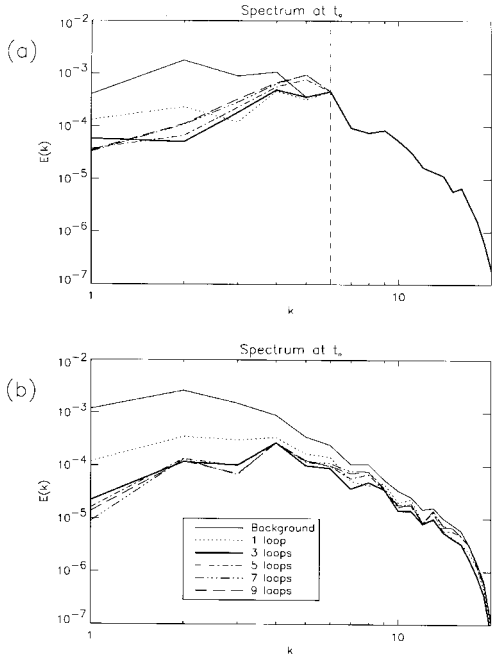

Fig. 15. Error spectra at the beginning (a) and the end (b) of the assimilation period as a function of the number of outer iterations used in the incremental approach when random noise is added to the observations in the low-resolution observation network.

with the result that 20 simulations are better than 10 or 30 simulations. At least within the present context, the noise level provides a criterion to stop the minimization based on the value of the functional. As shown on Fig. 14, the largest errors are mostly found in data void regions. These differences continue to grow as the number of simulations increases (result not shown).

Fig. 15 shows the error spectra obtained with the incremental formulation. The best result at t_N (giving the lowest RMS error of the vorticity field) is obtained after the third loop (the thick full line). Beyond that, the small-scale components degrade as in the experiment with the full-resolution 4D-Var while the large-scale components remain virtually unchanged. The dotted and long-dashed lines on Fig. 13 represent the value of the functional for the incremental and truncated 4D-Var respectively. It is interesting to notice that by not being able to change the small scales, the minimization cannot fit the observational noise. For both formulations, the functional reaches an asymptotic level that is close to the noise level but always above it. Therefore, the criterion of convergence that works well for the full-resolution 4D-Var fails here. It is to be expected that it may be difficult to derive a similar one for more complex models because this criterion neglects many factors. However, the idea of randomizing the observations and the state variables has been applied in Rabier and Courtier (1992) to estimate the error of their 4D-Var analyses based on a primitive-equation model. In Courtier et al. (1994),

it was used to estimate the Hessian of the cost function to precondition the minimization.

Finally, Fig. 16 shows the time evolution (over 40 time units) of the correlations (as defined by (3.5)) between the truth and the results of the different experiments. These results are compared against those obtained with the full-resolution 4D-Var stopped at 20 simulations which gave the best results. This comparison shows that the truncated 4D-Var is very close to the full-resolution 4D-Var during the assimilation period and, for this particular realization, is better for a large portion of the forecast period. This shows the negative impact of fitting the observational error. Moreover, the incremental formulation after 3 loops is close to both the truncated and the full-resolution 4D-Var. This indicates that, at a much lower computational cost, the incremental approach provides an analysis as good as the full-resolution 4D-Var and is much better than the low-resolution 4D-Var as shown on Fig. 16.

5. Summary and conclusion

Several aspects of the incremental formulation of the 4D-Var proposed by CTH94 were examined in the context of the simple model of quasi-geostrophic turbulence used in TBG95. By using a lower resolution model, two simplifications are made in the incremental approach: the minimization is constrained to a lower dimensional subspace and the gradient components along that subspace are computed approximately by using a truncated model. To make a distinction between these two aspects, the results obtained with the incremental 4D-Var were compared against those of a truncated 4D-Var that computes the full gradient but retains only those components used in the incremental minimization. In the incremental 4D-Var, the background (or reference) trajectory is obtained by integrating the full-resolution model while the analysis increments are propagated in time with a simplified (lower resolution) model. The procedure can then be seen as minimizing successive quadratic approximations of the full-resolution 4D-Var problem in a subspace defined by the simplified model components. Outer iterations are used to update the quadratic approximation and keep it to be a good local approximation of the cost function. Our results have shown that it is necessary to have a minimal number of outer iterations for the incremental minimization to converge to a good solution. This conclusion holds if the observations are perfect or not and if the spatial and temporal coverage of the observations is full or partial.

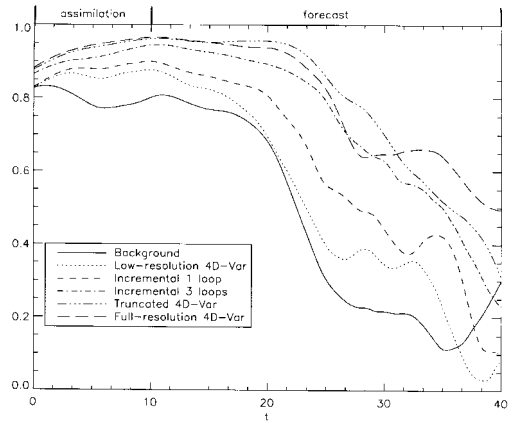

Fig. 16. Correlation between the control vorticity field and those obtained for the various variational formulations when random noise is added to the observations in the low-resolution observation network.

The solution was shown to converge globally when the gradient of the full-resolution 4D-Var and that of the incremental 4D-Var become nearly orthogonal. A balance must be reached between inner and outer iterations. On the one hand, too many inner iterations only succeed to minimize a quadratic functional that may not be representative of the actual cost function. On the other hand, a minimum number of inner iterations is needed for the quasi-Newton minimization algorithm to construct a useful approximation of the Hessian matrix. In this study, 10 inner simulations was found to be optimal, a number of simulations that agrees with that used in CTH94. The convergence rate of the incremental approach is about twice slower than the truncated 4D-Var formulation but its computational cost is approximately 64 times cheaper which makes it very efficient. Moreover, the error due to the incremental formulation account for only 10% of the total error, the remaining 90% of the total error being attributed to the fact that the cost function is minimized along a lower dimension subspace.

The impact of unbiased observational error was examined. For the full-resolution 4D-Var, our

results have shown that if the minimization is pursued for values of the cost function below some threshold value, no improvement is brought to the analysis. This threshold value was related to the characteristics of the unbiased random observational error and agrees with the result given in Gauthier (1992). This criterion does not hold however for the incremental and truncated 4D-Var for which the small scales cannot be used to fit the random noise. The functional then asymptotes to a level that is close but above the noise level. No clear criterion is then available to tell when the minimization should be stopped. This shows that such a criterion may be very difficult to establish in practice because the error on the observations is usually biased and the assimilating model is also introducing an error that is biased too.

Results with simple models cannot be readily extended to more complex models. The forcing term used here for example, is a stationary source term of barotropic energy which does not appear explicitly in the TLM (see eq. (4) in TBG95). Information about the forcing can then entirely be passed to the TLM through the reference trajectory derived from the full model. In more complex models, forcing associated with orography and parameterized sub-grid scale processes would also be approximately modelled in the TLM. Studies by Mahfouf et al. (1996) and Rabier et al. (1997), have shown that it is not sufficient to pass the information about these forcings through the reference trajectory: they need to be taken into account explicitly in the TLM. Currently, this is a very active field of research that reflects the importance of properly addressing such forcings in 4D-Var.

Acknowledgements

The authors would like to thank their colleagues, Monique Tanguay and Peter Bartello, who kindly made their model available to us. We also thank them for many stimulating discussions. We would also like to thank Philippe Courtier, Saroja Polavarapu, Florence Rabier and two anonymous reviewers for their comments on the manuscript that helped to improve the final version.

REFERENCES

Andersson, E., Pailleux, J., Thépaut, J. R. Eyre, A. P. McNally, G. A. Kelly and P. Courtier, 1994. Use of cloud-cleared radiances in 3D and 4D variational data assimilation. *Q. J. R. Meteorol. Soc.* ***120***, 627–653.

Courtier, P. and O. Talagrand, 1987. Variational assimilation of meteorological observations with the adjoint vorticity equations. Part II: Numerical results. *Q. J. R. Meteorol. Soc.* ***113***, 1329–1347.

Courtier, P., J.-N. Thépaut and A. Hollingsworth, 1994. A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Meteorol. Soc.* ***120***, 1367–1387.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljević, M. Hamrud, A. Hollingsworth, F. Rabier and M. Fisher, 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part I: Formulation. *Q. J. R. Meteorol. Soc.*, in press.

Cressman, G. P., 1959. An operative objective analysis scheme. *Mon. Wea. Rev.* ***88***, 367–374.

Daley, R., 1991. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, 457 pages.

Eyre, J. R., G. A. Kelly, A. P. McNally, E. Andersson and A. Persson, 1993. Assimilation of TOVS radiances information through one-dimensional variational analysis. *Q. J. R. Meteorol. Soc.* ***119***, 1427–1463.

Eyre, J. R., 1994. Assimilation of radio occultation measurements into a numerical weather prediction system. *ECMWF Technical Memorandum* ***199***, Reading, U.K, 34 pages.

Fisher, M. and P. Courtier, 1995. Estimating the covariance matrices of analysis and forecast error in variational data assimilation. *ECMWF Technical Memorandum* ***220***, Reading, U.K, 26 pages.

Gauthier, P. 1992. Chaos and quadri-dimensional data assimilation: a study based on the Lorenz model. *Tellus* ***44A***, 2–17.

Gauthier, P., L. Fillion, P. Koclas and C. Charette, 1996. Implementation of a 3D variational analysis at the Canadian Meteorological Centre. *Proceedings of the 11th AMS Conference on Numerical weather prediction.* Norfolk, Virginia, 19–23 August 1996.

Gilbert, J. C. and C. Lemaréchal, 1989. Some numerical experiments with variable-storage quasi-Newton algorithms. *Mathematical Programming* ***45***, 407–435.

Mahfouf, J.-F., R. Buizza and R. M. Errico, 1996. Strategy for including physical processes in the ECMWF variational data assimilation system. In: *Proceedings of the ECMWF Workshop on Nonlinear aspects of data assimilation.* Shinfield Park, Reading, U.K. RG2 9AX, 9–11 September 1996.

Orszag, S. A., 1971. Numerical simulation of incompress-

ible flows within simple boundaries (I). Galerkin (spectral) representations. *Stud. in Appl. Math.* ***50***, 293–327.

Parrish, D. F. and J. C. Derber, 1992. The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.* ***120***, 1747–1763.

Pires, C., R. Vautard and O. Talagrand, 1996. On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus* ***48A***, 96–121.

Rabier, F. and P. Courtier, 1992. Four-dimensional assimilation in the presence of baroclinic instability. *Q. J. R. Meteorol. Soc.* ***118***, 649–672.

Rabier, F., J. F. Mahfouf, M. Fisher, H. Järvinen, A. Simmons, E. Andersson, F. Bouttier, P. Courtier, M. Hamrud, J. Haseler, A. Hollingsworth, L. Isaksen, E. Klinker, S. Saarinen, C. Temperton, J. N. Thépaut, P. Undén and D. Vasiljević, 1997. Recent experimentation on 4D-Var and first results from a simplified Kalman filter. *ECMWF Technical Memorandum* ***240***. Reading, U.K, 42 pages.

Sun, J., D. W. Flicker and D. K. Lilly, 1991. Recovery of three-dimensional wind and temperature fields from single-Doppler radar data. *J. Atmos. Sci.* ***48***, 876–890.

Tanguay, M., P. Bartello and P. Gauthier, 1995. Four-dimensional data assimilation with a wide range of scales. *Tellus* ***47A***, 974–997.

Thépaut, J.-N., D. Vasiljević, P. Courtier and J. Pailleux, 1993. Variational assimilation of conventional meteorological observations with a multilevel primitive-equation model. *Q. J. R. Meteorol. Soc.* ***119***, 153–186.

Thépaut, J.-N. and P. Courtier, 1991. Four-dimensional data assimilation using the adjoint of a multilevel primitive equation model. *Q. J. R. Meteorol. Soc.* ***117***, 1225–1254.

Yang, W., M. Navon and P. Courtier, 1996. A new Hessian preconditioning method applied to variational data assimilation experiments using NASA general circulation models. *Mon. Wea. Rev.* ***124***, 1000–1017.

Zou, X., Y.-H. Kuo and Y.-R. Guo, 1995. Assimilation of atmospheric radio refractivity using a nonhydrostatic adjoint model. *Mon. Wea. Rev.* ***123***, 2229–2249.

Zupanski, M. 1993. A preconditioning algorithm for large-scale minimization problems. *Tellus* ***45A***, 478–492.