Tellus (1993), 45A, 370–387
Printed in Belgium – all rights reserved

TELLUS
ISSN 0280–6495

Variational data assimilation with moist threshold processes using the NMC spectral model

By XIAOLEI ZOU, *Supercomputer Computations Research Institute, Florida State University, Tallahassee, FL 32306-4052, USA*, I. M. NAVON*, *Department of Mathematics and Supercomputer Computations Research Institute, Florida State University, Tallahassee, FL 32306, USA* and J. G. SELA, *NOAA, NMC, World Weather Building, W/NMC23, Room 204, Washington DC 20233, USA*

(Manuscript received 2 October 1992; in final form 9 February 1993)

ABSTRACT

This paper describes a detailed study of variational 4-D data assimilation including the physical processes of large-scale precipitation and deep cumulus convection. The length of the assimilation window is 6 h, and the data are NMC's operational analyses. A comparison of the minimization behavior, the computational complexity, the quality of the retrieved initial state, with and without physical processes is presented. The results demonstrate the ability to perform 4-D variational data assimilation with discontinuous physical processes. The experiments are carried out with the NMC global spectral model in a resolution of 18 layers in the vertical and a 40 wave triangular truncation.

1. Introduction

In the last several years, considerable attention has been focused on variational methods for 4-D data assimilation (LeDimet and Talagrand, 1986; Derber, 1985; Lewis and Derber, 1985; Hoffman, 1986; Talagrand and Courtier, 1987; Courtier and Talagrand, 1987; Zou et al., 1992). Only very recently has the method been applied to more complex forecast models (Thépaut and Courtier, 1991; Navon et al., 1992; Chao and Chang, 1992). Derber (1989) developed a variational continuous assimilation method which allows the forecast model error to be taken into account, thus alleviating a major deficiency of the assimilation procedure, i.e., the assumption that the forecast model is perfect. The issue of the impact of threshold processes on variational data assimilation was first addressed by Douady and Talagrand (1990). Vukićević and Errico (1993) presented some preliminary work on the adjoint of discontinuous processes relating to issues of sensitivity analysis. Zupanski (1993) derived the adjoint of moist processes in the NMC/ETA limited-area model and carried out variational data assimilation including moist processes.

The theory of variational data assimilation and its application using an adiabatic version of the NMC spectral model as constraint, is described in the papers of Navon et al., (1992) and Zou et al., (1993), where issues of the impact of horizontal diffusion and surface drag on the convergence rate as well as the control of gravity wave oscillations are addressed. However, in the adiabatic version of the NMC model, the only nonlinear terms are the quadratic terms introduced by the advection and drag terms, so the linearization of the model is therefore straightforward. When physical processes that allow moisture phase changes are included, their effect on the adjoint equations is not intuitively obvious. The mathematical modeling of the atmosphere involves physical aspects that defy convenient quantification at the level of the resolution presently used in the model. Examples of these are radiation, vertical diffusion, deep (cumulus) and shallow convection, grid-scale

* Corresponding author.

condensation and precipitation. Such physical processes are, therefore, included in the model in some parameterized form. In this paper, we consider two such types of nonlinear adjustments, namely, the grid-scale precipitation and the cumulus convection.

The specific problem addressed here is related to the presence of "threshold", or "on-off" processes, triggered when a given parameter attains a prescribed critical value. For instance, the large-scale precipitation is triggered when the moisture at a given grid point location is larger than the saturated moisture value. The deep cumulus convection is active in regions where deep layers of conditionally unstable stratification and mean low level moisture convergence are present. In addition, some physical processes require costly computations. Examples of these are the calculation of the dew point temperature from the vapor pressure, the equivalent potential temperature (θ) from temperature and pressure fields at the lifting condensation level (LCL), and the temperature and specific humidity of an ascending parcel from its potential temperature and pressure. In order to economize the computational load, the operational model employs several pre-calculated tables along with linear or bilinear interpolations. However, the interpolated variables are non-differentiable at the interpolated points. Generally speaking, minimization algorithms are slowed down, if not totally inhibited by the presence of discontinuities in the gradient and in the cost function or even by only sharp variations thereof.

The problem of whether variational data assimilation with full physics is necessary for the success of 4-D variational data assimilation entails considerable experimentation and depends on the specific forecast application. The main purpose of this paper is to describe the elements of a variational assimilation scheme that includes "on-off" physical processes related to large-scale precipitation and cumulus convection.

The basic formulation of the variational data assimilation problem is briefly described in Section 2. Section 3 reviews the large-scale precipitation and the cumulus convection parameterization schemes. Section 4 describes the method employed for developing the adjoint operations of the "on-off" physical processes. The results are presented in Section 5. Preliminary conclusions are presented in Section 6.

2. Variational data assimilation problem

The objective of 4-D variational data assimilation is to find an optimal initial state $x(t_0)$ which minimizes the following cost function J

$$J(x(t_0)) = \sum_{r=0}^{R} W(x(t_r) - X^{\text{obs}}(t_r))^2, \tag{1}$$

where $x(t_0)$ is a vector of dimension n representing the initial state of the model, $x(t_r)$ is the nonlinear model's solution at time t_r with the physical processes included, $x^{\text{obs}}(t_0)$ is the observation of the model state at time t_r, and W is an $N \times N$ diagonal weighting matrix with W_{D}, W_{ζ}, W_{T}, $W_{\ln p_s}$ and W_q as diagonal submatrices of the weighting factors for divergence, vorticity, temperature, surface pressure and moisture fields, respectively.

The gradient ∇J of the function J to be minimized yields information concerning the behavior of J around $x(t_0)$, and all large-scale unconstrained minimization algorithms require the value of this gradient.

Computation of the gradient is accomplished by the adjoint of the tangent linear perturbation model, which is obtained from the nonlinear model by linearization about its own solution. If the nonlinear model is written as

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

the tangent linear model assumes the form

$$\frac{\partial x'}{\partial t} = \frac{\partial F(x)}{\partial x} x'. \tag{3}$$

The computer algorithm (using a leap-frog time integration method) for the tangent linear model can be represented as

$$x'(t_r) = P_r(x)\, x'(t_0). \tag{4}$$

where $P_r(x)$ represents all the operations required to obtain $x'(t_r)$ from $x'(t_0)$ and x is the full trajectory in the vicinity of which the linearization is carried out.

The adjoint model is defined as

$$\begin{aligned} \hat{x}'(t_0) &= P_r^{\text{T}}(x)\, \hat{x}(t_r), \\ \hat{x}'(t_r) &= W(x(t_r) - x^{\text{obs}}(t_r)), \end{aligned} \tag{5}$$

and the gradient of the cost function to be minimized can be shown to be:

$$\nabla J(x(t_0)) = \sum_{r=0}^{R} \hat{x}^r(t_0). \tag{6}$$

The main difference in developing the adjoint model with physics as compared to that of the adiabatic problem is the treatment of the "on-off" processes appearing in the model's code as conditional statements, depending on the model variables, and table lookups with linear and bilinear interpolations.

For all the numerical experiments described here we used a truncation of T40 with 18 levels in the vertical. The control variables are in the grid point space. A (90×46) Gaussian grid was used and the dimension of the vector of control variables was 302220. The adjustment was performed on the 6-hour interval $[t_0, t_R]$ preceding t_R.

3. Large-scale precipitation and cumulus convection formulation

3.1. Large-scale precipitation

The large-scale precipitation algorithm is used to simulate condensation of excess water vapor into precipitation, some of which is subsequently re-evaporated into the unsaturated layers below. It is performed after each leap-frog forward time step and can be considered as an adjustment.

Starting with the highest moisture-bearing layer, the process proceeds downwards, a layer at a time, checking for supersaturation. Layers having a relative humidity in excess of 100% ($q > q_s$) are brought back by an approximate wet-bulb process to a just-saturated state, with all of the resulting liquid water assumed to condense into rain. Evaporation of falling precipitation ($q < \alpha q_s$, where $\alpha = 0.8$ for all the layers except layer 1 at which $\alpha = 0.9$) depends on a parameterized drop-size distribution and upon relative humidity.

Supersaturation: $q > q_s$

The passage from a supersaturated state at temperature T and specific humidity q, to a just-saturated state, T^* and q^*, is given by

$$T^* = T + \frac{L}{c_p}\delta q_c, \tag{7}$$

$$q^* = q - \delta q_c, \tag{8}$$

$$\mathscr{P} = \mathscr{P} + \delta q_c \frac{\Delta p}{g}, \tag{9}$$

where δq_c is the water vapor condensed and is equal to

$$\delta q_c = \frac{q - q_s}{1 + \dfrac{L}{c_p}\dfrac{L}{RT^2} q_s(T)}, \tag{10}$$

and q_s is saturation humidity calculated by

$$q_s = \frac{R_d/R_v e_s}{p + (R_d/R_v - 1)\, e_s}, \tag{11}$$

$$e_s = 0.611 e^{(L/c_p)(1/T_d + 1/T)}. \tag{12}$$

Here Δp is the pressure difference (centibars), g is the gravity (ms^{-1}) and $\mathscr{P}$ is the rainfall (m).

The required adjustments of supersaturation result in a decrease in the specific humidity by the amount of δq_c, an increase in temperature by the amount of $\delta q_c L/c_p$, and a rainfall of the condensed water by the amount of $\delta q_c \Delta p/g$.

Evaporation: $q < \alpha q_s$

If the next lower layer in the column is also supersaturated, the above process of condensation is repeated and the rain $\mathscr{P}$ is accumulated. If the layer in the column is unsaturated, the layer is moistened and cooled by the evaporation of some or all of the falling rain, also according to a wet bulb process. In the current version of the model, evaporation of falling precipitation is restricted from raising the relative humidity higher than 80% ($\alpha = 0.8$), except in the first layer, where it is halted at a relative humidity of 90% ($\alpha = 0.9$).

The total precipitation evaporated during one time step is (Kessler, 1969, NMC, 1988)

$$\text{POTEVP} = 2\,\Delta t(\alpha q_s - q) \times 0.32 \sqrt{\mathscr{P}/2\,\Delta t}\,\Delta p/g, \tag{13}$$

where POTEVP is in units of meters (depth of rainwater).

If the precipitation increment $\mathscr{P}$ entering the layer exceeds POTEVP, all of POTEVP is sub-

tracted from the layer. If not, all of the precipitation $\mathscr{P}$ must evaporate into the layer, i.e.,

$$T^* = T - \frac{L}{c_p}\delta q_c', \tag{14}$$

$$q^* = q + \delta q_c', \tag{15}$$

$$\mathscr{P} = \mathscr{P} - \delta q_c' \frac{\Delta p}{g}, \tag{16}$$

where $\delta q_c'$ is the water vapor being evaporated and is equal to

$$\delta q_c' = \frac{\min\{\text{POTEVP}, \mathscr{P}\}/(\Delta p/g)}{1 + \frac{L}{c_p}\frac{L}{RT^2}q_s(T)}. \tag{17}$$

The procedure of condensation or evaporation is repeated layer by layer and the net amount of rainfall emerging at the bottom is halved to take into account the fact that the supersaturation has accumulated over a leapfrog time step $2\,\Delta t$, while the precipitation is incremented at each time step.

3.2. *Deep cumulus convection*

Another physical process included in the NMC model is the deep convection, which is modeled by a Kuo-Anthes type of scheme (Kuo, 1965, 1974; Anthes, 1977). Moisture convergence and deep conditional instability are required in order for the deep convection to be active. The condensed water in cumulus convection may later evaporate. The condensed water usually evaporates at a different level than the level at which it was condensed.

In the following, a subscript designates a model's layer. Several checks are performed to impose realistic limits on the types of columns that are capable of supporting convection:

(1) $T_2 > 5°\text{C}$, the low temperatures must be warm;

(2) $T_2 > T_3$, no low-level inversion exists;

(3) the total moisture convergence in layer 1 to LEVDHQ is sufficient to produce a rainfall rate of at least 2 mm day^{-1}, where LEVDHQ = 5 is determined from $\sigma(\text{LEVDHQ}+1) < 0.3$;

(4) the buoyant layers $(T_c - T_e > 0)$ in the column have a combined pressure thickness greater than $0.3 \times p_s$.

The procedures for calculating the cloud base, cloud top, and in-cloud profiles of T and q are provided in the Appendix.

The vertically-integrated moisture convergence is taken to be the mass-weighted integral of the change in specific humidity during the one leapfrog time step

$$W = \int_{p_B}^{p_T} \Delta q^{(k)} \frac{dp^{(k)}}{p_s}, \tag{18}$$

where $\Delta q^{(k)} = q^{(k)}(t_{\tau+1}) - q^{(k)}(t_{\tau-1})$ is the moisture convergence at level k, p_T and p_B are the pressures at cloud top and cloud base, respectively.

The moisture convergence between cloud base and cloud top is partitioned into a rain-producing portion $(1-b)W$ and a humidity-increasing portion bW, where b is the fraction of the area of the sky covered by the deep cumulus clouds and is calculated from the column-integrated relative humidity:

$$b = 1 - \frac{\bar{q}}{\bar{q}_s}. \tag{19}$$

The cumulus clouds are assumed to exist only momentarily, and they dissipate by mixing with the environmental air at the same level. The heat and moisture carried up by the cloud air are imparted to the environmental air on the basis of the vertical distribution of cloud-environment temperature and specific humidity differences. Thus, if the temperature and specific humidity of the air outside the cloud at level k are $T^{(k)}$ and $q^{(k)}$, while the corresponding cloud values are $T_c^{(k)}$ and $q_c^{(k)}$, then the temperature $T^{*(k)}$ and specific humidity $q^{*(k)}$ of the mean air at the level k after mixing is given by

$$T^{*(k)} = T^{(k)} + \frac{\frac{L}{c_p}(1-b)\,W}{\int_{p_B}^{p_T}(T_c^{(k)} - T^{(k)})\frac{dp^{(k)}}{p_s}} \times (T_c^{(k)} - T^{(k)}) \tag{20}$$

$$q^{*(k)} = q^{(k)} + \frac{bW}{\int_{p_B}^{p_T}(q_c^{(k)} - q^{(k)})\frac{dp^{(k)}}{p_s}} \times (q_c^{(k)} - q^{(k)}) - \Delta q^{(k)} \tag{21}$$

$$\mathscr{P}^* = \int_{p_B}^{p_T} \frac{(1-b)\,W}{\int_{p_B}^{p_T} (T_c^{(k)} - T^{(k)}) \dfrac{dp^{(k)}}{p_s}} \times (T_c^{(k)} - T^{(k)}) \frac{dp^{(k)}}{p_s}, \tag{22}$$

where $\mathscr{P}^*$ is the convective precipitation. Adjustment is limited to layers between the cloud base and cloud top.

If the air is unsaturated ($q \leqslant q_s$) in the convective column, and convective precipitation is available ($\mathscr{P}^* > 0$), the effect of evaporation of the convective falling precipitation is also included.

If the evaporation potential is larger than the amount of precipitation available to evaporate ($q_s - q \geqslant \mathscr{P}^*/\Delta\sigma$), then the temperature $T^{**(k)}$, the specific humidity $q^{**(k)}$, and the amount of precipitation $\mathscr{P}$ after evaporation will be given by

$$T^{**(k)} = T^{*(k)} - \frac{L}{c_p} \frac{\mathscr{P}^*}{\Delta\sigma}, \tag{23}$$

$$q^{**(k)} = q^{*(k)} + \frac{\mathscr{P}^*}{\Delta\sigma}, \tag{24}$$

$$\mathscr{P} = 0. \tag{25}$$

Otherwise ($q_s - q < \mathscr{P}^*/\Delta\sigma$), $T^{**(k)}$, $q^{**(k)}$, and $\mathscr{P}$ will be equal to

$$T^{**(k)} = T^{*(k)} - \frac{L}{c_p} (q_s - q^{*(k)}) \tag{26}$$

$$q^{**(k)} = q^{*(k)} + (q_s - q^{*(k)}) \tag{27}$$

$$\mathscr{P} = \mathscr{P}^* - \int_{p_B}^{p_T} (q_s - q^{*(k)}) \frac{dp}{p_s}. \tag{28}$$

The rate of convective rainfall in m/s^{-1} reaching the ground is

$$R = \mathscr{P} \times p_s/g. \tag{29}$$

4. Adjoint code development of the large-scale precipitation and cumulus convection

4.1. *"On-off" processes*

The "on-off" processes appear in the code as IF statements depending on the model variables. For example, in the large-scale precipitation code, a typical IF statement assumes the form:

```
IF      q > q_s     THEN
            instructions for precipitation
ENDIF
IF      q < 0.8q_s  THEN
            instructions for evaporation
ENDIF
```

In order to develop the adjoint code, a bit vector is used in each IF statement to record the route of the forward integration of the nonlinear model. Therefore, the modified nonlinear code assumes the form:

```
IF      q >  q_s    THEN
            IPRE = 1
            instructions for precipitation
  ELSE
            IPRE = -1
ENDIF
IF      q <  0.8q_s THEN
            IEVP = 1
            instructions for evaporation
  ELSE
            IEVP = -1
ENDIF
```

The same bit vector is used for the IF statement in the tangent linear model, i.e.,

```
IF      IPRE > 0    THEN
        linearized instructions for precipitation
ENDIF
IF      IEVP > 0    THEN
        linearized instructions for evaporation
ENDIF
```

The adjoint operation is based on the tangent linear code and follows the same route backwards

```
IF      IEVP > 0    THEN
        adjoint code of instructions for evaporation
ENDIF
IF      IPRE > 0    THEN
        adjoint code of instructions for precipitation
ENDIF
```

Thus, in order to check the correctness of the adjoint code, one more integration of the nonlinear model with bit vectors added is required in order to determine the routes for the IF statements

included in the physical processes for both the tangent linear model and the adjoint model, respectively.

4.2. *Table lookups*

Another issue involved in developing the adjoint code of the cumulus convection is related to table lookups. For the tangent linear and adjoint models, one can no longer use the original tables which were employed in the direct nonlinear model. The values in the tables are built for the nonlinear calculation. The original (analytical) formulas used to build the lookup tables are required to compute the first derivatives. This is very different from the development of the adjoint of the adiabatic model, where the only nonlinear terms are those due to the quadratic terms. We will illustrate this issue in detail by a simple example. Let us consider here the dew point temperature calculation from the vapor pressure.

Differentiating equation (A2) (see Appendix), we have

$$VP' = 0.611 * [Ax^{A-1} - (A+B) * x^{A}] \times e^{(A+B)(1-x)}x', \tag{30}$$

$$x' = -\frac{T_0}{T_d^2}T_d', \qquad x = -\frac{T_0}{T_d^2}T_d.$$

Therefore, the perturbation T_d' can be calculated if VP' is known. The value of T_d and, thus x, appearing in the coefficient is obtained by (A2) via a direct calculation (without using the table lookup).

The perturbed equivalent potential temperature, θ_e', can be obtained by differentiating (A8), which can be written as

$$\theta_e = f(T, p)\, e^{g(T,p)}, \tag{31}$$

where

$$f(T, p) = \frac{T}{\left(\frac{p-e_s}{100}\right)^{\kappa}}, \qquad g(T, p) = \frac{\varepsilon L e_s}{T(p-e_s)}. \tag{32}$$

Therefore, θ_e' may be derived as

$$\theta_e' = f' e^{g} + f g' e^{g}, \tag{33}$$

where

$$f'(T', p', T, p) = \frac{T'\left(\frac{p-e_s}{100}\right)^{\kappa} - \kappa T\left(\frac{p-e_s}{100}\right)^{\kappa-1}(p'-e_s')}{\left(\frac{p-e_s}{100}\right)^{2\kappa}}, \tag{34}$$

$$g'(T', p', T, p) = \frac{\left[\begin{array}{l}\varepsilon(L'e_s + Le_s')\, T(p-e_s) \\ \quad -\varepsilon L e_s(T'(p-e_s) + T(p'-e_s'))\end{array}\right]}{T^2(p-e_s)^2}, \tag{35}$$

$$e'(T, T') = -e_{s_0}\left(\frac{T_0}{T}\right)^{C}\frac{T_0 T'}{T^2} \times \left(C\left(\frac{T_0}{T}\right)^{-1} - D\right) e^{D(1-T_0/T)}, \tag{36}$$

$$L' = (c_v - c_l)\, T', \tag{37}$$

and

$$C = \frac{c_l - c_v}{R_v}, \qquad D = C + \frac{L}{R_v T_0}. \tag{38}$$

For the perturbed cloud temperature calculation, we substitute (37) into (33) and solve for T'. After some manipulation we obtain

$$T^{c'} = \frac{\theta' + h_p e^{\varepsilon L e_s/(T(p-e_s))}p'}{h_T e^{\varepsilon L e_s/(T(p-e_s))}} \tag{39}$$

where

$$h_T = \frac{\left(\frac{p-e_s}{100}\right)^{\kappa} + T\kappa\left(\frac{p-e_s}{100}\right)^{\kappa-1} h_{e_s}}{\left(\frac{p-e_s}{100}\right)^{2\kappa}} + \frac{T}{\left(\frac{p-e_s}{100}\right)^{\kappa}} \times \frac{\left[\begin{array}{l}\varepsilon((c_v - c_l)\, e_s + L h_{e_s})\, T(p-e_s) \\ \quad -\varepsilon L e_s(p-e_s) + \varepsilon L e_s T h_{e_s}\end{array}\right]}{T^2(p-e_s)^2} \tag{40, 41}$$

$$h_p = \frac{\kappa T \left(\frac{p-e_s}{100}\right)^{\kappa-1}}{\left(\frac{p-e_s}{100}\right)^{2\kappa}} + \frac{T}{\left(\frac{p-e_s}{100}\right)^{\kappa}} \times \frac{\varepsilon L e_s T}{T^2 (p-e_s)^2}, \quad (42)$$

$$h_{e_s} = -e_{s_0}\left(\frac{T_0}{T}\right)^C \frac{T_0}{T^2}\left(C\left(\frac{T_0}{T}\right)^{-1} - D\right) \times e^{D(1-(T_0/T))}. \quad (43)$$

The perturbed cloud moisture can be derived from (A12) and written as

$$q^{c'} = 0.622 \frac{(p-e_s)\, h_{e_s} T^{c'} - e_s (p' - h_{e_s} T^{c'})}{(p-e_s)^2}. \quad (44)$$

It is pointed out that some input variables in the direct code are reused implicitly, and should be carefully accounted for.

Due to the effect of high nonlinearity and the appearance of the table lookups in the large-scale precipitation and the cumulus convection codes, the verification of the correctness of the tangent linear model prior to deriving the adjoint code becomes very important. We have verified that the tangent-linear version is the true linearization of the direct code in the vicinity of a given trajectory for each of the subroutines related to the physical processes.

5. Numerical results

5.1. *Computational complexity*

In the following experiments, the "observations" consisted of two complete states of vorticity, divergence, temperature, surface pressure field and moisture from the NMC global operational data assimilation system, 6-h apart, which are denoted by $x^{obs}(t_0)$ and $x^{obs}(t_R)$. The adjustment was performed on the 6-h interval $[t_0, t_R]$ preceding t_R. The problem is to find an optimal initial state $x^*(t_0)$ which minimizes the cost function J given by eq. (1) where the sum is performed at t_0 and t_R.

The computational cost of calculating the gradient of the cost function using the adjoint technique is independent of the dimension of the problem, meaning

$$\frac{\mathcal{T}(J, \nabla J)}{\mathcal{T}(J)} \leqslant C, \quad (45)$$

where $\mathcal{T}(J)$ and $\mathcal{T}(J, \nabla J)$ measure the CPU time to compute J and the pair of $(J, \nabla J)$, respectively, and C is a constant independent of n. This constant is small, between 2 and 4, depending on the operations performed in the direct model. Table 1 displays the CPU time used by the NMC T40, 18-layer model, its tangent linear model and its adjoint model for a six hour time integration respectively. The time step is 1800 s. All the "on-off" routes and the nonlinear model solutions at each time step were stored in memory during the direct integration to avoid the recomputation in the adjoint. We observe that the CPU time increased from 16.65 s for an adiabatic model to 29.15 s for the partially diabatic model due to the "on-off" physical processes. The tangent linear model integration is slightly more expensive than that of the nonlinear model integration, while the integration of the adjoint model is most expensive but still less than twice the CPU time required for the integration of the nonlinear model.

Table 2 displays the values of C for the gradient calculation. For the adiabatic version of the model, C is equal to 2.6. When the "on-off" physical processes are included in the direct model and the adjoint model, $C = 2.3$. In addition, we note that the cost function calculation adds 1 s of CPU time,

Table 1. *CPU time used by the NMC T40, 18-layer primitive equation model, its tangent linear model and its adjoint model for a six hour time integration with a time step of 1800s*

CPU time (second)	nonlinear model	linear model	adjoint model
adiabatic	16.65	21.79	25.14
with "on-off" processes	29.15	31.10	34.49

Table 2. *CPU time used for calculating the cost function and its gradient with respect to the initial state*

CPU time (unit: second)	$\mathcal{T}(J)$	$\mathcal{T}$ (Grad J)	$\mathcal{T}(J, \text{Grad } J)/\mathcal{T}(J)$
adiabatic	17.31	27.35	2.6
with "on-off" processes	30.01	37.73	2.3

in both cases, to the model's integration. Additional 3 seconds are required for including the forcing term in the adjoint model. Therefore, the direct access of data in the nonlinear model and adjoint model, for both the model solution and the "on-off" route, are computationally very efficient.

5.2. Numerical results

The first test consisted of two parallel assimilations with and without the large-scale precipitation and cumulus convection, both including the horizontal diffusion and a simplified surface drag (see Navon et al., 1992). The minimization was found to be similar in both cases with the value of the cost function decreasing to about 15% of its original value; see Fig. 1a. The norm of gradient decreased almost two orders of magnitude; see Fig. 1b. The total CPU times required for 60 iterations of the minimization with and without the physical processes were 76 and 44 minutes, respectively. The maximum memory used was 7.4MW and 6.4MW respectively. From this experiment, we conclude that "on-off" threshold processes when included in the model do not hinder the minimization process. Theoretically, if the gradient of the cost function calculated by integrating the adjoint model is always a one-sided gradient, i.e., the model integration follows the same "on-off" route at each iteration, the large-scale unconstrained minimization will still work. We note however that this is not the case in our minimization experiment, and despite this fact, the minimization process performed successfully. We now consider the horizontal distribution of the model's variables displayed over an area encompassing North America. Figs. 2a, 2b, 2c show three examples of the large-scale saturation, evaporation, and the cumulus convection during the first 3 iterations of minimization, respectively. In Fig. 2(a) we plotted the amount of δq_c (see eq. (3.4)) when large-scale saturation occurs ($q > q_s$) at a model grid point located at (90° E, 35° N, 225 mb). We observe that the saturation process was turned on from the fourth leap-frog time step during the first and second iterations. During the third iteration, this process was also turned on at the fourth time step but it was turned off at the last time step of the assimilation window. We can draw a similar conclusion from Figs. 2b, c for the cumulus convection process.

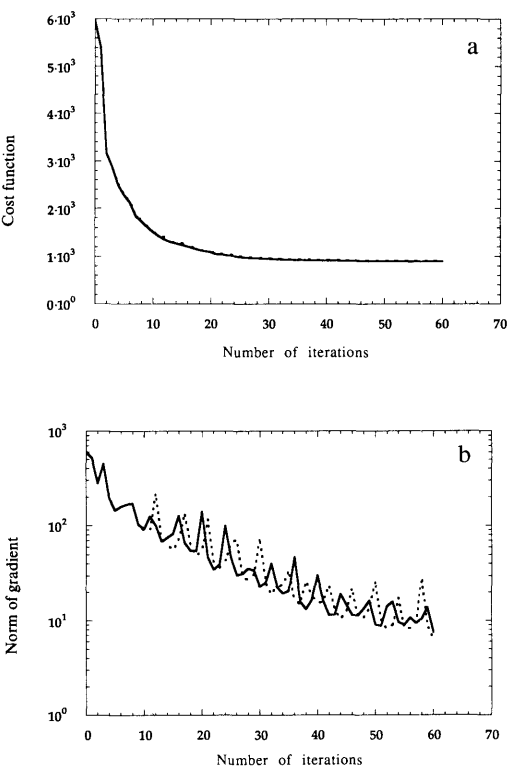

Fig. 1. Variations of (a) the cost function J and (b) the gradient ∇J with the number of iterations of the minimization with (full line) and without (dotted line) the inclusion of "on-off" physical processes for operational analysis.

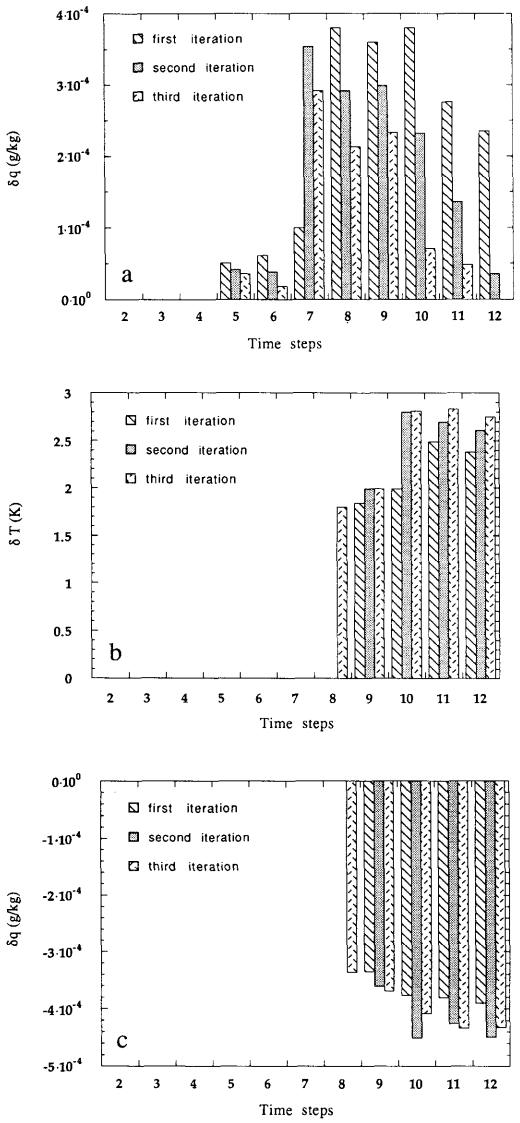

Fig. 2. (a) Large scale supersaturation amounts at the point (90°E, 35°N, 225 mb) as a function of time. (b) Cumulus convective heating of the environment at the point (25°W, 38°N, 920 mb) as a function of time. (c) Cumulus convective moistening of the environment at the point (25°W, 38°N, 920 mb) as a function of time.

In the following, we compare the differences between the retrieved initial states from the adiabatic and partially diabatic runs. First, the RMS of the retrieved divergence field was found to increase slightly when the large-scale precipitation and the cumulus convection processes were included in the variational data assimilation (Fig. 3).

In order to observe the differences in the retrieved initial conditions with or without the precipitation processes, we plotted in Figs. 4a–h the difference of the divergence, vorticity, temperature and the moisture fields at high and low levels respectively. In the area where the divergence in the higher layers and the convergence in lower layers increases due to the inclusion of the "on-off" processes in the minimization, (we shall call it the active region), the temperature and the vorticity were also found to have increased. At the lower levels, as the convergence increased, the temperature decreased and the moisture increased. In addition, a northwest–southeast wave train is observed in the vorticity field which propagates toward the equator.

Fig. 5 displays the retrieved initial stream functions near 500 mb (level 9). We observe that there is a trough upstream of the active region, and that the differences in the minimization solution with or without the physical processes are strongly correlated with the weather system.

In order to separate the effects of large-scale precipitation and cumulus convection, the latter was turned off in the second experiment, and only the large-scale precipitation was active. We find that most of the differences in Fig. 4 are due to the large-scale precipitation and the deep cumulus

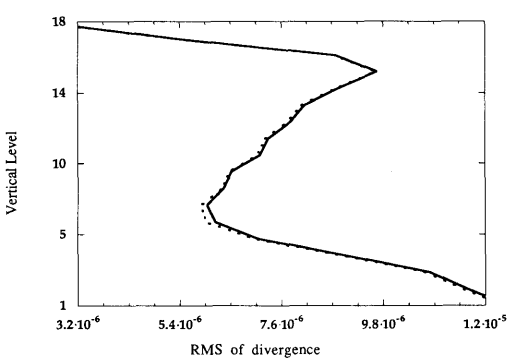

Fig. 3. RMS of the retrieved initial divergence at all vertical levels after the minimization with (full line) or without (dotted line) the "on-off" processes.

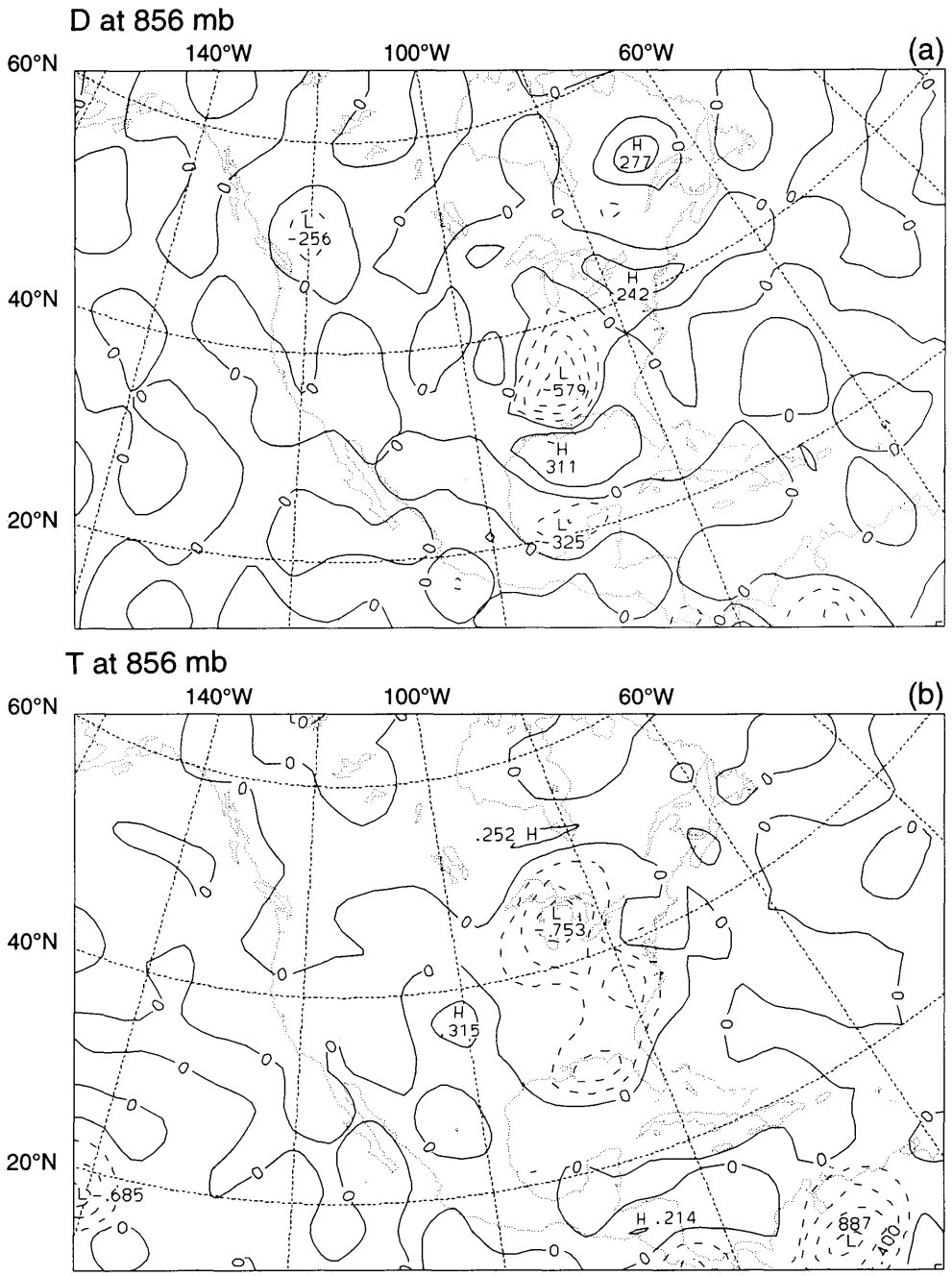

Fig. 4. The differences of divergence, temperature, vorticity and the moisture fields at (a–d) the 856 mb level and (e–h) the 225 mb level between the two retrieved initial states. Contour intervals for divergence, vorticity, temperature and moisture are $8.0 \times 10^{-7}\,s^{-1}$, $2.0 \times 10^{-6}\,s^{-1}$, 0.07 K and $8.0 \times 10^{-5}\,g\,kg^{-1}$ in the upper layer and $1.5^{-6}\,s^{-1}$, $5.0 \times 10^{-6}\,s^{-1}$, 0.2 K and $8.0 \times 10^{-5}\,g\,kg^{-1}$ in the lower layer respectively. Values on the isolines of divergence and moisture fields are scaled by 10^8 and 10^6 respectively. Values on the isolines of vorticity in the upper and lower layers are scaled by 10^8 and 10^7, respectively.

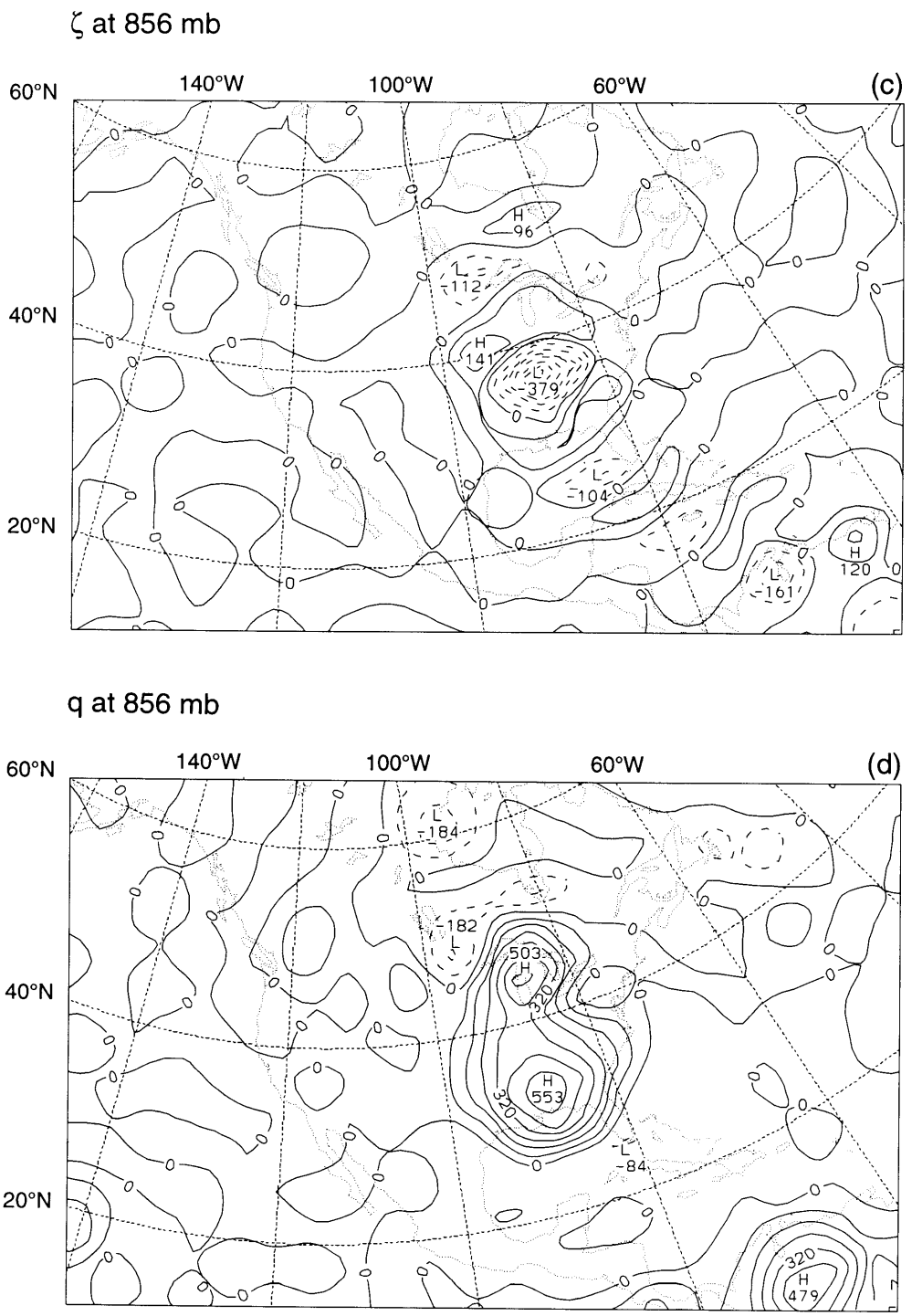

Fig. 4–Continued

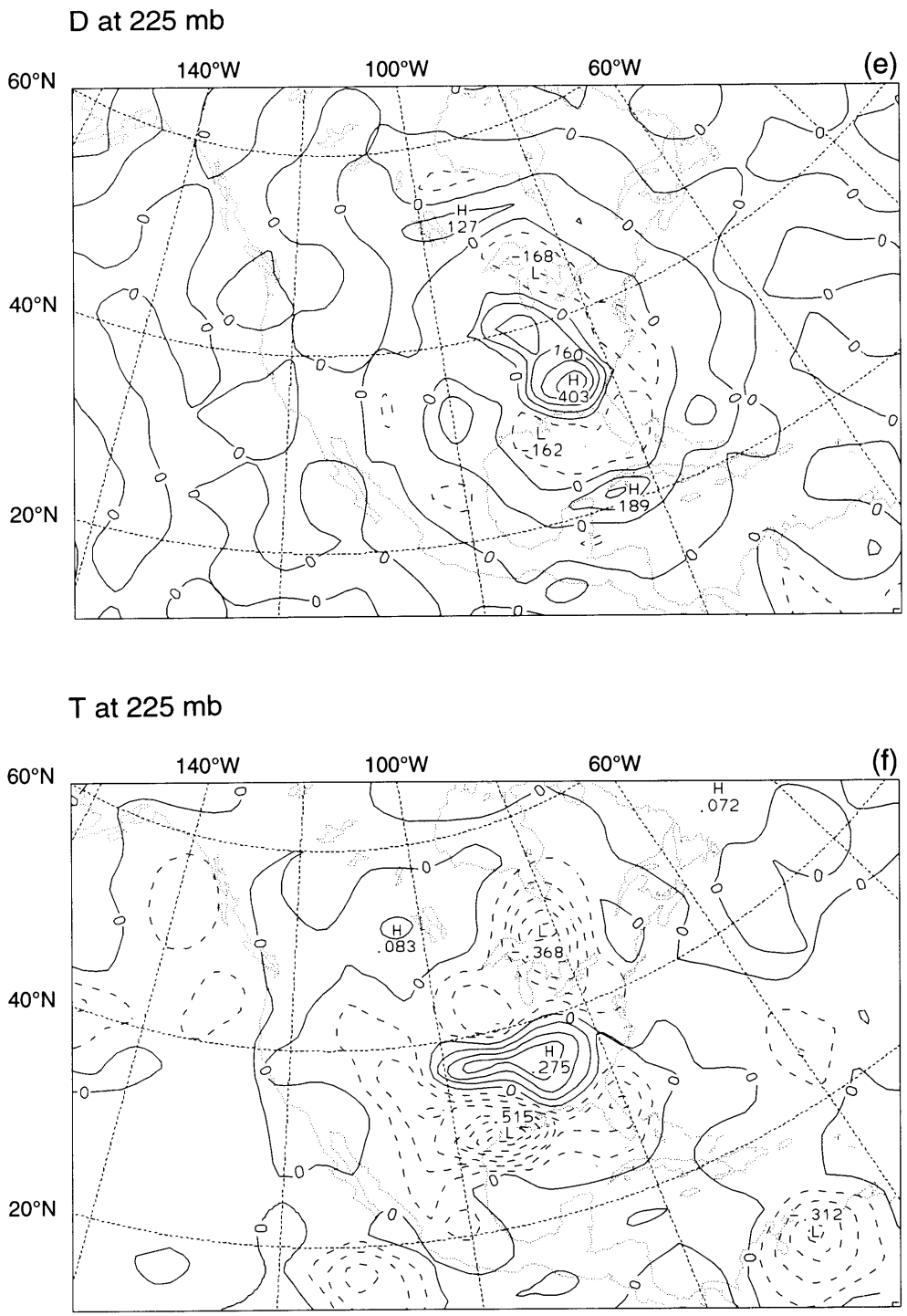

Fig. 4–Continued

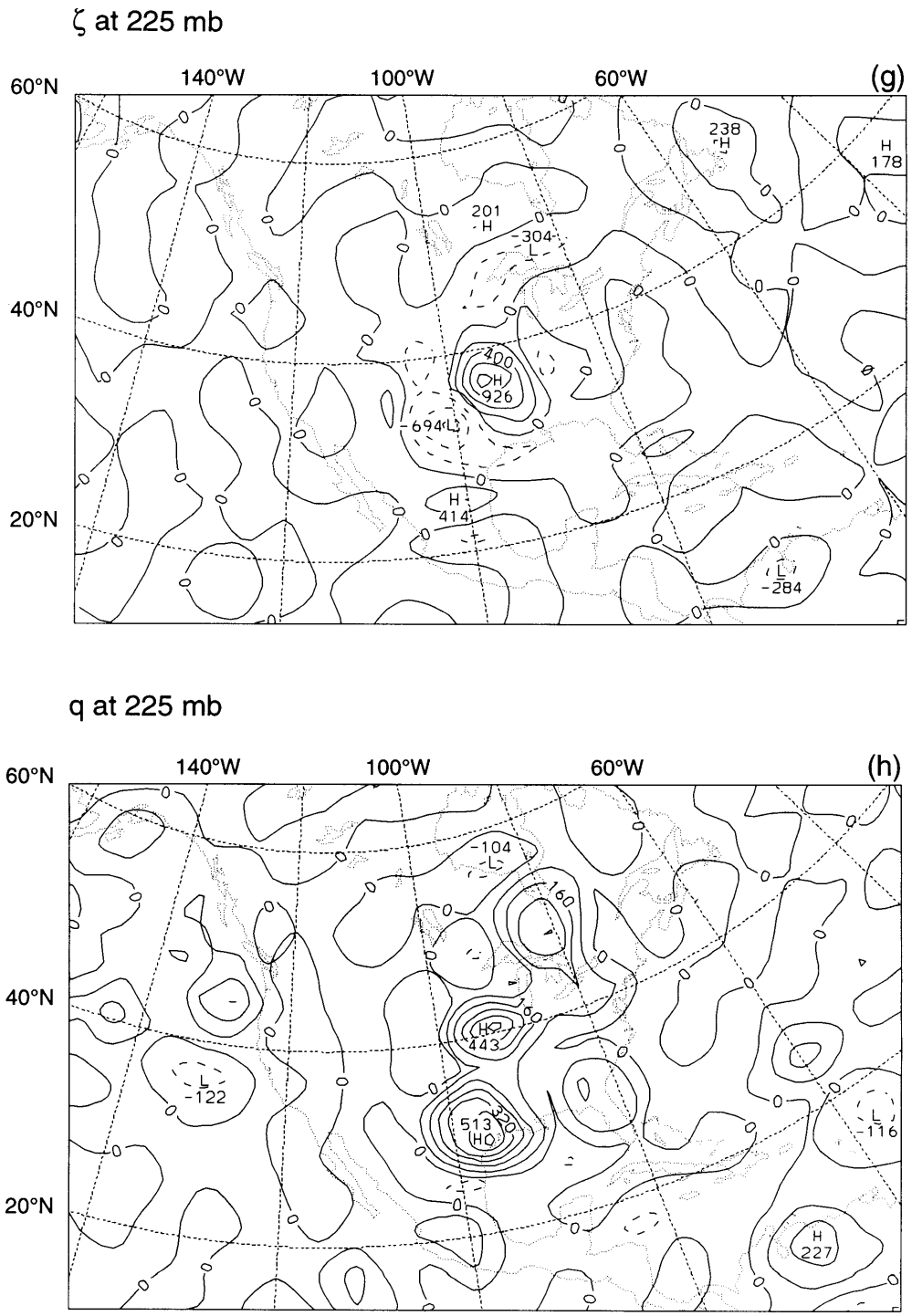

Fig. 4–Continued

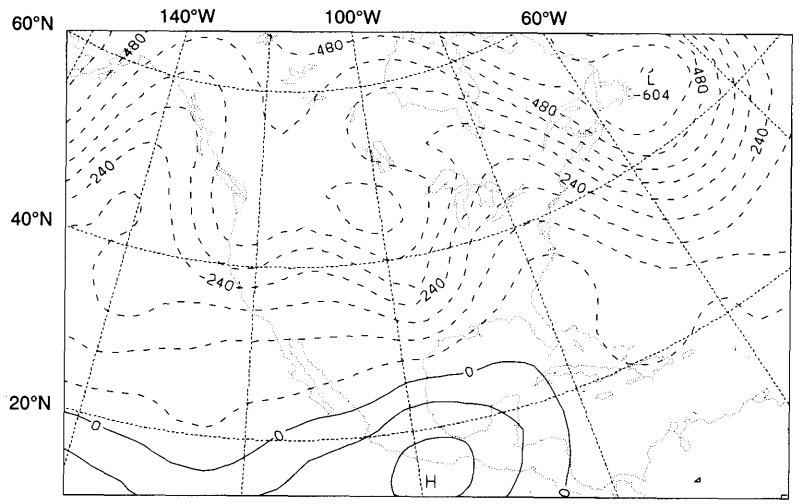

Fig. 5. The distribution of the retrieved initial stream function at 500 mb obtained by the minimization including the "on-off" processes. Contour interval is 6.0×10^6 m^2 s^{-1} and values on the isolines are scaled by 10^{-5}.

convection impacts positively on the dynamic features in front of the trough south of Lake Erie. Fig. 6 shows the divergence difference field between the retrieval with both the large-scale precipitation and the cumulus convection turned on and retrieval with only the large-scale precipitation being included in the minimizations. We conclude that the divergence in the active region is intensified due to the inclusion of the cumulus convection process.

We also made 6 h forecast of the precipitation from the two initial conditions (Fig. 7). We find

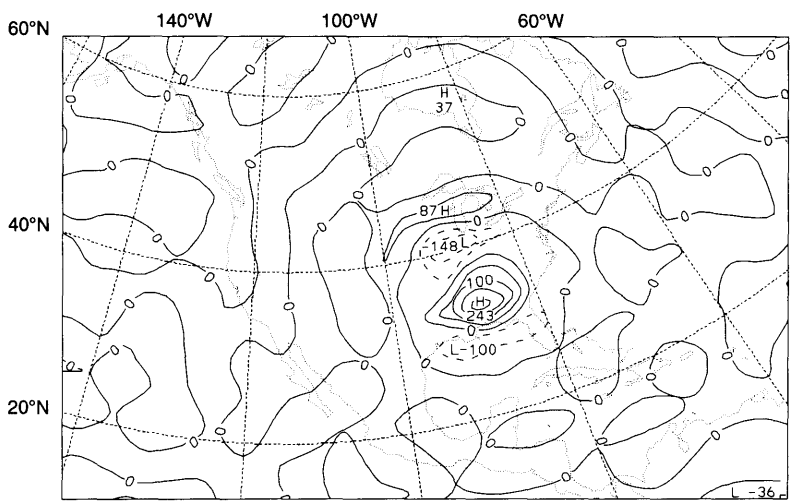

Fig. 6. The differences of the divergence field at 225 mb between the retrieved initial condition when both the large-scale precipitation and the cumulus convection are included in the minimization and with only the large-scale precipitation process included. Contour interval is 6.0×10^{-7} s^{-1} and values on the isolines are scaled by 10^8.

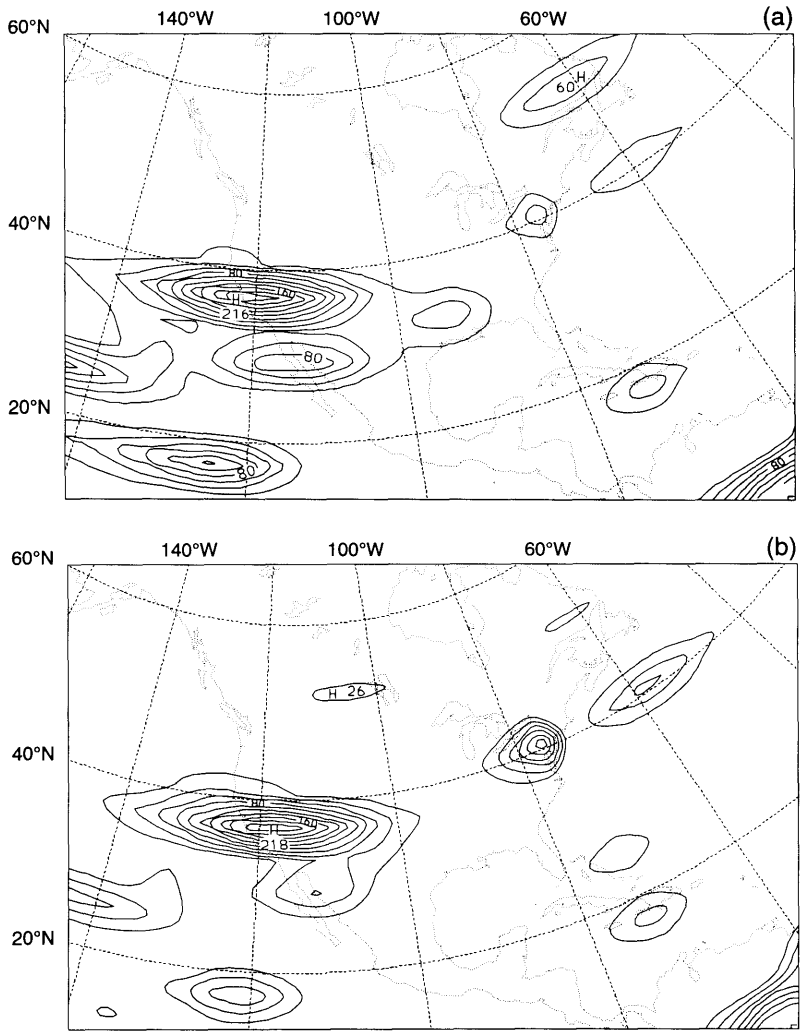

Fig. 7. The total 6-h precipitation forecasted from the initial state with (a) or without (b) the "on-off" processes. Contour interval is 2.0×10^{-2} m and values on the isolines are scaled by 10^3.

that the initial condition obtained by the variational data assimilation including the "on-off" processes produce larger amounts of precipitation and greater area coverage over most of the map. However, the amount of precipitation near Lake Erie is smaller than the forecast from the initial condition obtained without the inclusion of the physical processes. Since we do not have verifying data, we cannot conclude which result is better. We have however demonstrated that differences near meteorologically active areas may result from the two respective minimizations.

6. Conclusions

In this paper, a variational data assimilation scheme including diabatic effects of large-scale precipitation and the cumulus convection has been presented. An 18 level T40 version of the NMC global spectral model was used to demonstrate that the "on-off" operations encountered in the physical routines do not hinder the performance of the adjoint technique in the case of a fit to analysis. The limited-memory quasi-Newton method of Liu and Nocedal (1989) successfully minimizes the

cost function that defines the distance between observations and prediction.

From our initial experiment using the two operationally analyzed data states, we found that the "on-off" physical processes included in variational data assimilation impact only marginally on the retrieved initial state. The effects of these differences may, however, be significant near frontal areas. The most significant result of this work is the ability to perform 4-D data assimilation experiments with meteorologically realistic precipitation processes included in the adjoint model. Further work on assimilating real data, especially when they are sparsely distributed, and obtaining a better assessment of the beneficial effect of including the physical processes remains to be carried out.

7. Acknowledgments

This research was funded by NSF Grant ATM-9102851. Additional support was provided by National Meteorological Center and the Supercomputer Computations Research Institute at Florida State University, which is partially funded by the Department of Energy through contract No. DE-FC0583ER250000. The first author wishes to thank H. S. Bedi for useful discussions.

8. Appendix

Calculation of the in-cloud profiles of T and q

The actual calculation of specific parcel ascent trajectories begins with the calculation of dew point at the parcel's origin (say for instance at model layer L). First the vapor pressure was obtained from the moisture q of layer L using the formula:

$$VP = \frac{c_3 p_s q \sigma_2}{c_2 + q},$$

$$c_2 = \frac{R/R_v}{1 - R/R_v}, \qquad \text{(A1)}$$

$$c_3 = \frac{100}{1 - R/R_v}.$$

Then the dew point calculation was calculated via a linear interpolation from a T_d versus VP table. The dew point temperature T_d table is precalculated via a Newton iterative method using the following formula:

$$VP = 0.611 \left(\frac{T_0}{T_d}\right)^A e^{(A+B)(1-T_0/T_d)}, \qquad \text{(A2)}$$

where

$$A = \frac{c_l - c_v}{R_v}, \qquad B = \frac{L_0}{R_v T_0}, \qquad \text{(A3)}$$

$$T_0 = 273.16^\circ\text{K},$$

where c_l is the specific heat of liquid water (which enters through the temperature dependence on L).

The value of T_d as a function of VP is obtained by finding the root of $f(x) = 0$, where $f(x)$ is defined as

$$f(x) = A \ln x - (A+B)\, x - [\ln VP - \ln 0.611 - (A+B))], \qquad \text{(A4)}$$

$$x = \frac{T_0}{T_d}.$$

The iteration starts from $x^{(0)} = 1.0$ and ends whenever the following convergence criterion is satisfied:

$$|x^{(n)} - x^{(n-1)}| \leqslant 10^{-8}, \qquad \text{(A5)}$$

where superscript (n) represents the nth iteration. Having the dew point temperature T_d, the temperature and pressure at the lifting condensation level (LCL), T(LCL) and p(LCL), along a dry adiabat are obtained from the formula

$$T(\text{LCL}) = T - (T - T_d) \times (0.954442 + 0.967772 \times 10^{-3} T + (T - T_d) \times (-0.710321 \times 10^{-3} - 0.270742 \times 10^{-5} T)), \qquad \text{(A6)}$$

$$p^\kappa(\text{LCL}) = \frac{T}{T(\text{LCL})}\, p_s^\kappa. \qquad \text{(A7)}$$

The equivalent potential temperature θ_e of a parcel at the LCL is obtained by the Rossby formula:

$$\theta_e = \frac{T}{\left(\frac{p-e_s}{100}\right)^{\kappa}} e^{\varepsilon L e_s/(T(p-e_s))}, \qquad (A8)$$

$$\varepsilon = \frac{R}{R_v c_p}, \qquad \kappa = \frac{R}{c_p},$$

where R and R_v are the gas constants for dry air, e_s is the saturation vapor pressure at T, c_p the specific heat of dry air at constant pressure, L the latent heat of vaporization for water ($L = L_0 + (c_v - c_l)(T - T_0)$), and e_s the saturation vapor pressure which is calculated from an integration of the Clausius–Clapeyron equation:

$$e_s = e_{s_0}\left(\frac{T_0}{T}\right)^{c_l - c_v/R_v} \times e^{((c_l - c_v/R_v) + (L_0/R_v T_0))(1 - T_0/T)}. \qquad (A9)$$

This value of θ_e will be conserved in the specific parcel ascent from the LCL. Based on the θ_e, a new value of temperature, T_c, is calculated as a function of this θ_e and pressure p at different layers along a moist adiabat using the formulation (A8)–(A9). The calculation is completed in the nonlinear model by a bilinear interpolation through a pre-calculated table, which is built by a Newton iterative method applied to the function:

$$f(T_c) = \theta_e - \frac{T_c}{\left(\frac{p-e_s}{100}\right)^{\kappa}} e^{\varepsilon L e_s/T^c(p-e_s)}, \qquad (A10)$$

$$\varepsilon = \frac{R}{R_v c_p}, \qquad \kappa = \frac{R}{c_p},$$

where $e_s = e_s(T_c)$ is calculated using (A9). The iteration starts from $T_c^{(0)} = (1.12\theta_2 - \theta_2)\, p^{\kappa}$ and stops when the following convergence criterion

$$|T_c^{(n)} - T_c^{(n-1)}| \leqslant 10^{-4}. \qquad (A11)$$

is satisfied.

The cloud moisture q_c is then calculated by

$$q_c = \frac{0.622 e_s(T_c)}{p - e_s(T_c)}. \qquad (A12)$$

If $T_c \leqslant T_e$ ($q_c \leqslant q_e$), this T_c (q_c) is disregarded.

The full procedure described above was repeated twice, once each for layer $L = 2$ and $L = 3$. The resulting vertical profile is obtained from these two cloud soundings, a composite is selected consisting of the warmest cloud temperature in each layer. The cloud top is considered to be the highest layer at which the parcel is buoyant. The LCL is the cloud base. Non buoyant layers may exist between cloud base and cloud top.

REFERENCES

Anthes, R. A. 1977. A cumulus parameterization scheme utilizing a one dimensional cloud model. *Mon. Wea. Rev.* **105**, 270–286.

Chao, W. C. and Chang, L. P. 1992. Development of a 4-dimensional variational analysis system using the adjoint method at GLA. Part I: Dynamics. *Mon. Wea. Rev.* **120**, 1661–1673.

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

Derber, J. C. 1985. *The variational 4-D assimilation of analysis using filtered models as constraints.* PhD Thesis. University of Wisconsin–Madison, Madison, Wisconsin 53706, 142 pp.

Derber, J. C. 1989. A variational continuous assimilation technique. *Mon. Wea. Rev.* **117**, 2437–2446.

Douady, D. and Talagrand, O. 1990. The impact of threshold processes on variational assimilation. World Meteorological Organisation, *Proceedings from International Symposium on assimilation of observations in meteorology and oceanography*, Clermont-Ferrand, France, 9–13 July, pp. 486–487.

Hoffmann, R. N. 1986. A four-dimensional analysis exactly satisfying equations of motion. *Mon. Wea. Rev.* **114**, 388–397.

Kessler, E. 1969. On the distribution and continuity of water substance in atmospheric circulations. *Met. Monographs* **10**, no. 32, American Meteorological Society, Boston, 84 pp.

Kuo, H. L. 1965. On formulation and intensification of tropical cyclones through latent heat release by cumulus convection. *J. Atmos. Sci.* **22**, 40–63.

Kuo, H. L. 1974. Further studies of the parameterization of the influence of cumulus convection on large-scale flow. *J. Atmos. Sci.* **31**, 1232–1240.

LeDimet, F. X. and Talagrand, O. 1986. Variational algorithms for analysis and assimilation of Meteo-

rological Observations: Theoretical Aspects. *Tellus* **38A**, 97–110.

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

Liu, D. C. and Nocedal, J. 1989. On the limited memory BFGS method for large scale optimization, *Mathematical Programming* **45**, 503–528.

Navon, I. M., Zou, X., Derber, J. and Sela, J. 1992. Variational data assimilation with an adiabatic version of the NMC spectral model. *Mon. Wea. Rev.* **120**, 1433–1446.

NMC, 1988. Documentation of the Research Version of the NMC medium-range forecasting model. National Meteorological Center, W/NMC2, Room 204, Development Division, World Weather Building, 5200 Auth Road, Camp Springs, Maryland 20746, USA.

Talagrand, O. and Courtier, P. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I. Theory. *Q. J. R. Meteorol. Soc.* **113**, 1311–1328.

Thépaut, J. N. and Courtier, P. 1992. Four-dimensional variational data assimilation using the adjoint of a multilevel primitive equation model. *Q J R M S* **117**, 1225–1254.

Vukićević, T. and Errico, R. M. 1993. Linearization and adjoint of parameterized moist diabatic processes, *Tellus* **45A**, 493–510.

Zou, X., Navon, I. M. and LeDimet, F. X. 1992. Incomplete observations and control of gravity waves in variational data assimilation. *Tellus* **44A**, 273–296.

Zou, X., Navon, I. M. and Sela, J. 1993. Control of gravitational oscillations in variational data assimilation. *Mon. Wea. Rev.* **121**, 272–289.

Zupanski, D. 1993. The effects of discontinuities in the Betts–Miller cumulus convection scheme on four-dimensional variational data assimilation, *Tellus* **45A**, 511–524.