Tellus (1993), 45A, 525–538
Printed in Belgium – all rights reserved

TELLUS
ISSN 0280–6495

Treatment of on/off switches in the adjoint method: FDDA experiments with a simple model

By JIAN-WEN BAO*,[1] and THOMAS T. WARNER[1,2], [1]*Department of Meteorology, 503 Walker Building*; [2]*Earth System Science Center, The Pennsylvania State University, University Park, PA 16802, USA*

(Manuscript received 2 November 1992; in final form 7 June 1993)

ABSTRACT

The purpose of this study is to illustrate the feasibility of using the adjoint method in the data assimilation procedure, where the assimilation model involves on/off switches associated with physical processes. Although the assimilation model used in this study is very simple, it has the on/off switches that are typical in more sophisticated and realistic assimilation models. The calculus of variations is used to confirm that variational FDDA using the adjoint method allows the equations of the assimilation model to have a finite number of first-order discontinuous points. These points represent the on/off switches for which the Jacobian matrix of the model equation may not exist. In practice, when the on/off switches are involved, the gradient of the functional can be obtained by assuming that an infinitesimally small initial perturbation does not change the time interval within which the switching takes place. The time when the corresponding switching in the adjoint equations is performed is the same as that in the forward integration of the model equations, i.e., the switching time is determined by the basic state. Numerical experiments are then conducted using the aforementioned ideas in a simple one-dimensional convection model involving on/off switches associated with latent heat release. The implication of the numerical experiments is that there is no difficulty, in theory, in treating the on/off switches in the adjoint of the assimilation model. Rather, the success of the adjoint method is dependent on the numerical aspects of the implementation, such as the proper scaling of the variables.

1. Introduction

Most primitive-equation numerical models simulate, explicitly or through parameterization, many processes of the hydrological cycle in atmospheric circulations. When such models are integrated from observed initial conditions, for data assimilation or forecasting, the hydrological variables (in addition to the dynamic variables) should be initialized. Lack of initialization for the hydrological variables results in the so-called spin-up problem (Girard and Jarraud, 1982; Heckley, 1985; Kasahara et al., 1988; Krishnamurti et al., 1988) because of a large hydrological imbalance in the early period of model integration. The presence of the hydrological spin-up time has a severe impact on the very-short-range forecasts of mesoscale systems. Therefore, initializing hydrological variables and, thus, reducing the hydrological spin-up time are an important problem in mesoscale initialization. Variational four-dimensional data assimilation (FDDA), especially the one using the adjoint method, is expected to be very promising in addressing the spin-up problem (Daley, 1991).

Research on variational FDDA using the adjoint method has been mainly concerned with the objective analysis and initialization of model variables such as wind, temperature and pressure (or geopotential). Although variational FDDA using the adjoint method is regarded as one of the "next-generation" FDDA methods, little research has been done regarding the treatment of the

* Present address: NCAR/MMM/MPS, P.O. Box 3000, Boulder, CO 80307-3000, USA.

moisture variables and the hydrological cycle. Nevertheless, water vapor plays an important role in a wide spectrum of atmospheric motions, from the general circulation to severe storms. The major problem associated with the application of variational FDDA using the adjoint method to the assimilation of moisture variables is the fact that the model equations are not differentiable with respect to the model state at the time when condensation or evaporation occurs, i.e., at this time, the derivatives of the model equations with respect to the model variables do not exist because of the existence of the 1st-order discontinuity of the model equations with respect to the model state associated with condensation or evaporation. This type of discontinuity is represented by on/off switches in the computer program. However, it is not clear whether this differentiability is a restriction on the model equations in the application of variational FDDA using the adjoint method to initialization of the moisture variable (Derber, 1989).

In this study, the calculus of variations is used to justify that variational FDDA using the adjoint method allows the existence of the first-order discontinuities of the numerical model equations in the time interval during which FDDA is performed. The theoretical arguments are presented in Section 2, and the interpretation of the discrete form of the problem can be found in Section 3. A simple one-dimensional convection model, with a parameterized condensation process, and its adjoint are used to illustrate the feasibility of variational FDDA using the adjoint method to initialize moisture variables. Numerical results obtained by using this model are presented in Section 4. The implication of the theoretical analysis and the practical illustration to the further application of variational FDDA to initialize moisture variables in more realistic models is summarized in Section 5.

2. Variational FDDA subject to discontinuous constraints

2.1. Background

The mathematical basis of variational FDDA using the adjoint method involves specifying the initial state for a numerical model subject to dynamical constraints, so as to minimize the functional J which is dependent upon the initial state of the model:

$$J(\boldsymbol{x}_0) = \int_0^T D[\boldsymbol{x}(\boldsymbol{x}_0, t), t]\, \mathrm{d}t, \tag{1}$$

where $D(\boldsymbol{x}, t)$ is the measure of the discrepancy between the model state and the observation at t, $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{x}_0, t)$ is the model state at t, $\boldsymbol{x}_0$ is the initial state, and T is the assimilation window. The dynamical constraints are the physical laws that the numerical model equations describe. Assume that $D[\boldsymbol{x}(\boldsymbol{x}_0, t), t]$ is quadratic as follows

$$D[\boldsymbol{x}(\boldsymbol{x}_0, t), t] = [\boldsymbol{x}(\boldsymbol{x}_0, t) - \boldsymbol{x}_{\mathrm{obs}}(t)]^{\mathrm{tr}} \times \boldsymbol{W}(t)[\boldsymbol{x}(\boldsymbol{x}_0, t) - \boldsymbol{x}_{\mathrm{obs}}(t)], \tag{2}$$

where the subscript, obs, denotes the observation, the superscript, tr, denotes the transpose of a vector or a matrix, and $\boldsymbol{W}(t)$ is a positive-definite diagonal (thus symmetric) matrix whose elements are the weighting coefficients of observed variables. It is clear that the functional J may be regarded as the time integral of the squared "distance" between the model state and the observations. It should be noted that there are an infinite number of ways to define the form of D because, mathematically, "distance" is represented by a norm which can be defined in many ways. The reason for the quadratic form being chosen here is that it is defined in the normed inner product space which is commonly used in meteorology. The quadratic functional is also mathematically convenient for minimization, because its minimum is determined by a system of linear equations if the numerical model equations are linear.

Without losing generality, the numerical model used in data assimilation can be written as an ordinary differential equation of vector form:

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \boldsymbol{M}(\boldsymbol{x}, t), \tag{3}$$

$$\boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{4}$$

where $\boldsymbol{M}$ is a nonlinear function of $\boldsymbol{x}$ and t.

Two general approaches are used to minimize (1) subject to the constraints (3). One is to multiply a vector of the Lagrangian multipliers by (3), add the product to (2), and replace the sum with

D in eq. (1) to form a so-called Lagrangian (or augmented) functional J'. Taking the first-variation of J' and letting it be zero yields the so-called Euler-Lagrange equations. In principle, by solving the Euler-Lagrange equations, the optimal initial model state can be determined. Unfortunately, the Euler-Lagrange equations resulting from application of this procedure to operational data assimilation are usually very difficult to solve. The second approach, referred to as the adjoint method, is to directly minimize J by an iterative numerical algorithm using first-order derivative information of the functional. The essence of this approach is, in addition to the calculation of the functional value, the computation of the first-order derivative of the functional with respect of the model state, which is accomplished by integrating the adjoint equations derived from the numerical model equations.

In the formalism of FDDA using the adjoint method, it is required that the so-called Gâteaux derivatives of the model equation, and the functional J with respect to the model state, exist (Cacuci, 1981). The Gâteaux derivative is related to the Gâteaux differential which is the most general definition of the variation of a nonlinear functional (Sagan, 1969; Cacuci, 1981). The Gâteaux differential of $J(\boldsymbol{x})$ at $\boldsymbol{x}_0$ in the direction $\boldsymbol{r}$ is defined as:

$$\delta J = J(\boldsymbol{x}_0 + \boldsymbol{r}) - J(\boldsymbol{x}_0) + \Delta(\boldsymbol{r}), \tag{5}$$

where $\Delta(\ )$ is such that, when α is a real number,

$$\Delta(\alpha \boldsymbol{r})/\alpha \to 0 \qquad \text{as} \quad \alpha \to 0. \tag{6}$$

In some of the literature, δJ is referred to as the weak variation of J. Following Sagan (1969), we will refer to the Gâteaux differential as the first variation. It is important to note that J need not be continuous in $\boldsymbol{x}$ for δJ to exist. When $J(\boldsymbol{x})$ satisfies a weak Lipschitz condition at $\boldsymbol{x}_0$, and

$$J(\boldsymbol{x}_0 + \alpha \boldsymbol{r}_1 + \alpha \boldsymbol{r}_2) - J(\boldsymbol{x}_0 + \alpha \boldsymbol{r}_1) - J(\boldsymbol{x}_0 + \alpha \boldsymbol{r}_2) + J(\boldsymbol{x}_0) = \theta(\alpha), \tag{7}$$

where $\theta(\alpha) \to 0$ as $\alpha \to 0$, δJ is linear in $\boldsymbol{r}$. That is, $\delta J = [\partial J(\boldsymbol{x}_0)/\partial \boldsymbol{x}]^{\mathrm{tr}} \boldsymbol{r}$, where $\partial J(\boldsymbol{x}_0)/\partial \boldsymbol{x}$ is called the Gâteaux derivative (Saaty, 1981; Cacuci, 1981). For convenience, we will call $\partial J(\boldsymbol{x}_0)/\partial \boldsymbol{x}$ the gradient of J with respect to $\boldsymbol{x}$ at $\boldsymbol{x}_0$. It should be pointed out that this gradient need not be continuous at $\boldsymbol{x}_0$.

The numerical procedure of minimizing J subject to the constraints (3) has been discussed by a number of authors (e.g., Lewis and Derber, 1985; Talagrand and Courtier, 1987; Courtier and Talagrand, 1987; Thacker and Long, 1988). However, in all of the documented experiments, the gradients of the functional were computed using numerical models and their adjoint that only involved simple physics. That is, forcing terms involving on/off switches were not taken into account. An example of this kind of forcing term is the latent heal release due to water-vapor condensation. Since the adjoint of the tangent linear version of the numerical model is integrated backward to estimate the gradient of J with respect to $\boldsymbol{x}_0$, the numerical procedure using the model with simple physics can not be applied without theoretical justification of the situations where on/off switches are involved in the model and its adjoint. This is because the first-order derivatives of the model equations with respect to the model state may not exist at the time when the switches are turned on/off. Although, practically, this problem may not be serious, as commented by Derber (1989), it is still unclear how to deal with the switches in practice. In the following subsection, the calculus of variations is applied to derive the adjoint method that can be used to compute the gradient of the functional in the FDDA procedure where the numerical model has on/off switches.

2.2. *Computation of the functional gradient using the adjoint with on/off switches*

In the situations where M is not smooth within the interval $[0, T]$ due to on/off switches of some forcing terms, it is reasonable to believe that M is still continuous and piece-wise smooth within that interval if the solution of (3) and (4) is assumed to be unique with the same interval. That is, the number of so-called corner points, where the derivatives of $\boldsymbol{M}$ with respect to the model state, $\boldsymbol{x}(t)$, do not exist, is finite. The continuousness guarantees that the left and right derivatives of $\boldsymbol{M}$ with respect to $\boldsymbol{x}(t)$ exist at the corner points.

Without losing generality, assume that there is one corner point, t_c, within $[0, T]$, which is specified by the criterion condition

$$\boldsymbol{C}(\boldsymbol{x}(t_c), t_c) = 0. \tag{8}$$

Furthermore, we suppose that one set of numerical model equations,

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}=\boldsymbol{M}_{(1)}(\boldsymbol{x},t), \tag{9}$$

is applied for $t<t_c$, and another set of numerical model equations is applied for $t>t_c$, namely,

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}=\boldsymbol{M}_{(2)}(\boldsymbol{x},t). \tag{10}$$

Although the following discussion will be confined to the single-corner-point situation, the extension to a finite number of corner points, where criterion conditions like (8) must be met, is straightforward.

Let $\boldsymbol{x}^*(t)$ be the known state with the specific initial state $\boldsymbol{x}_0^*$. Introduce a vector of the adjoint variables, $\boldsymbol{a}(t)$. By adjoining J defined by (1) to the differential constraints (9) and (10) with $\boldsymbol{a}(t)$, one can get a new functional J', called the Lagrangian functional:

$$\begin{aligned}J'(\boldsymbol{x}_0^*)=&\int_0^{t_c^*}\left\{D(\boldsymbol{x}^*,t)+\boldsymbol{a}(t)^{\mathrm{tr}}\right.\\&\left.\times\left[\frac{\mathrm{d}\boldsymbol{x}^*(t)}{\mathrm{d}t}-\boldsymbol{M}_{(1)}(\boldsymbol{x}^*,t)\right]\right\}\mathrm{d}t\\&+\int_{t_c}^{T}\left\{D(\boldsymbol{x}^*,t)+\boldsymbol{a}(t)^{\mathrm{tr}}\right.\\&\left.\times\left[\frac{\mathrm{d}\boldsymbol{x}^*(t)}{\mathrm{d}t}-\boldsymbol{M}_{(2)}(\boldsymbol{x}^*,t)\right]\right\}\mathrm{d}t,\end{aligned} \tag{11}$$

where t_c^* is the corner point corresponding to $\boldsymbol{x}_0^*$ and specified by (8). It should be noted that $J=J'$, because the model equations are strong constraints.

It is easy to verify that the first variation of J with respect to the variation in $\boldsymbol{x}_0^*$ has the following form:

$$\begin{aligned}\delta J=&\int_0^{t_c^*}\left\{\nabla_x^{\mathrm{tr}}D(\boldsymbol{x}^*,t)\,\delta\boldsymbol{x}+\boldsymbol{a}(t)^{\mathrm{tr}}\right.\\&\left.\times\left[\frac{\mathrm{d}\,\delta\boldsymbol{x}(t)}{\mathrm{d}t}-\frac{\partial\boldsymbol{M}_{(1)}(\boldsymbol{x}^*,t)}{\partial\boldsymbol{x}}\delta\boldsymbol{x}\right]\right\}\mathrm{d}t\\&+\left\{D(\boldsymbol{x}^*,t_c^{*-})+\boldsymbol{a}(t_c^{*-})^{\mathrm{tr}}\right.\\&\left.\times\left[\frac{\mathrm{d}\boldsymbol{x}^*(t_c^{*-})}{\mathrm{d}t}-\boldsymbol{M}_{(1)}(\boldsymbol{x}^*,t_c^{*-})\right]\right\}\delta t_c\\&-\left\{D(\boldsymbol{x}^*,t_c^{*+})+\boldsymbol{a}(t_c^{*+})^{\mathrm{tr}}\right.\\&\left.\times\left[\frac{\mathrm{d}\boldsymbol{x}^*(t_c^{*+})}{\mathrm{d}t}-\boldsymbol{M}_{(2)}(\boldsymbol{x}^*,t_c^{*+})\right]\right\}\delta t_c\\&+\int_{t_c^*}^{T}\left\{\nabla_x^{\mathrm{tr}}D(\boldsymbol{x}^*,t)\,\delta\boldsymbol{x}+\boldsymbol{a}(t)^{\mathrm{tr}}\right.\\&\left.\times\left[\frac{\mathrm{d}\,\delta\boldsymbol{x}(t)}{\mathrm{d}t}-\frac{\partial\boldsymbol{M}_{(2)}(\boldsymbol{x}^*,t)}{\partial\boldsymbol{x}}\delta\boldsymbol{x}\right]\right\}\mathrm{d}t,\end{aligned} \tag{12}$$

where, δt_c is the variation of t_c^* with respect to the variation in $\boldsymbol{x}_0^*$. Integrating by parts, we have

$$\begin{aligned}\delta J=&\boldsymbol{a}(T)^{\mathrm{tr}}\,\delta\boldsymbol{x}(T)-\boldsymbol{a}(0)^{\mathrm{tr}}\,\delta\boldsymbol{x}(0)\\&+\boldsymbol{a}(t_c^{*-})^{\mathrm{tr}}\,\delta\boldsymbol{x}(t_c^{*-})-\boldsymbol{a}(t_c^{*+})^{\mathrm{tr}}\,\delta\boldsymbol{x}(t_c^{*+})\\&+\left\{D(\boldsymbol{x}^*,t_c^{*-})+\boldsymbol{a}(t_c^{*-})^{\mathrm{tr}}\right.\\&\left.\times\left[\frac{\mathrm{d}\boldsymbol{x}^*(t_c^{*-})}{\mathrm{d}t}-\boldsymbol{M}_{(1)}(\boldsymbol{x}^*,t_c^{*-})\right]\right\}\delta t_c\\&-\left\{D(\boldsymbol{x}^*,t_c^{*+})+\boldsymbol{a}(t_c^{*+})^{\mathrm{tr}}\right.\\&\left.\times\left[\frac{\mathrm{d}\boldsymbol{x}^*(t_c^{*+})}{\mathrm{d}t}-\boldsymbol{M}_{(2)}(\boldsymbol{x}^*,t_c^{*+})\right]\right\}\delta t_c\\&+\int_0^{t_c^*}\left\{\left[-\frac{\mathrm{d}\boldsymbol{a}(t)^{\mathrm{tr}}}{\mathrm{d}t}-\boldsymbol{a}(t)^{\mathrm{tr}}\frac{\partial\boldsymbol{M}_{(1)}(\boldsymbol{x}^*,t)}{\partial\boldsymbol{x}}\right.\right.\\&\left.\left.+\nabla_x^{\mathrm{tr}}D(\boldsymbol{x}^*,t)\right]\delta\boldsymbol{x}(t)\right\}\mathrm{d}t\\&+\int_{t_c^*}^{T}\left\{\left[-\frac{\mathrm{d}\boldsymbol{a}(t)^{\mathrm{tr}}}{\mathrm{d}t}-\boldsymbol{a}(t)^{\mathrm{tr}}\frac{\partial\boldsymbol{M}_{(2)}(\boldsymbol{x}^*,t)}{\partial\boldsymbol{x}}\right.\right.\\&\left.\left.+\nabla_x^{\mathrm{tr}}D(\boldsymbol{x}^*,t)\right]\delta\boldsymbol{x}(t)\right\}\mathrm{d}t.\end{aligned} \tag{13}$$

As illustrated by Fig. 1, the following relations hold

$$\begin{aligned}\delta\boldsymbol{x}(t_c)&=\delta\boldsymbol{x}(t_c^{*-})+\frac{\mathrm{d}\boldsymbol{x}^*(t_c^{*-})}{\mathrm{d}t}\delta t_c\\&=\delta\boldsymbol{x}(t_c^{*+})+\frac{\mathrm{d}\boldsymbol{x}^*(t_c^{*+})}{\mathrm{d}t}\delta t_c.\end{aligned} \tag{14}$$

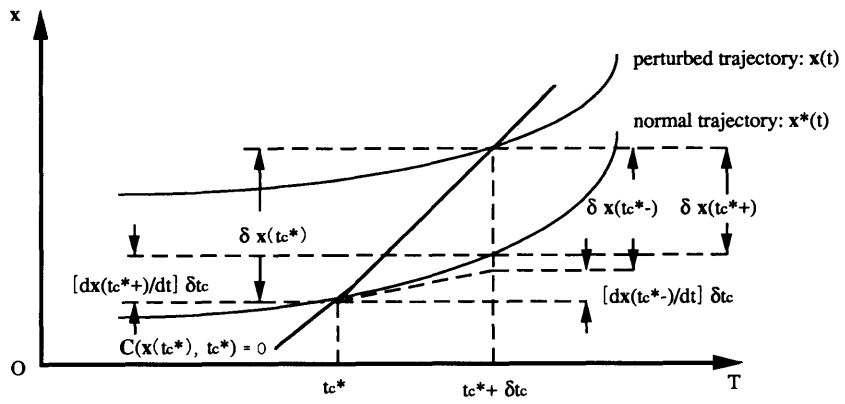

Fig. 1. Relationships between $\delta x(t^*)$, $\delta x(t^{*-})$, $\delta x(t^{*+})$ and δt_c.

Using (14) to eliminate $\delta x(t_c^{*-})$ and $\delta x(t_c^{*+})$ in eq. (13) yields

$$\begin{aligned}\delta J = {} & \boldsymbol{a}(T)^{\mathrm{tr}}\,\delta\boldsymbol{x}(T) - \boldsymbol{a}(0)^{\mathrm{tr}}\,\delta\boldsymbol{x}(0) \\ & + (\boldsymbol{a}(t_c^{*-})^{\mathrm{tr}} - \boldsymbol{a}(t_c^{*+})^{\mathrm{tr}})\,\delta\boldsymbol{x}(t_c) \\ & + \{[D(\boldsymbol{x}, t_c^{*-}) - \boldsymbol{a}(t_c^{*-})^{\mathrm{tr}}\,\boldsymbol{M}_{(1)}(\boldsymbol{x}, t_c^{*-})] \\ & - [D(\boldsymbol{x}, t_c^{*+}) - \boldsymbol{a}(t_c^{*+})^{\mathrm{tr}}\,\boldsymbol{M}_{(2)}(\boldsymbol{x}, t_c^{*+})]\}\,\delta t_c \\ & + \int_0^{t_c^*} \left\{ \left[-\frac{\mathrm{d}\boldsymbol{a}(t)^{\mathrm{tr}}}{\mathrm{d}t} - \boldsymbol{a}(t)^{\mathrm{tr}}\,\frac{\partial \boldsymbol{M}_{(1)}(\boldsymbol{x}^*, t)}{\partial \boldsymbol{x}} \right.\right. \\ & \left.\left. + \nabla_x^{\mathrm{tr}}\,D(\boldsymbol{x}^*, t) \right] \delta\boldsymbol{x}(t) \right\} \mathrm{d}t \\ & + \int_{t_c^*}^{T} \left\{ \left[-\frac{\mathrm{d}\boldsymbol{a}(t)^{\mathrm{tr}}}{\mathrm{d}t} - \boldsymbol{a}(t)^{\mathrm{tr}}\,\frac{\partial \boldsymbol{M}_{(2)}(\boldsymbol{x}^*, t)}{\partial \boldsymbol{x}} \right.\right. \\ & \left.\left. + \nabla_x^{\mathrm{tr}}\,D(\boldsymbol{x}^*, t) \right] \delta\boldsymbol{x}(t) \right\} \mathrm{d}t. \end{aligned} \tag{15}$$

On the other hand, from the aforementioned definition of the Gâteaux differential, it is known that the first variation of J with respect to $\delta\boldsymbol{x}(0)$ is

$$\delta J = \left[\frac{\partial J}{\partial \boldsymbol{x}(0)}\right]^{\mathrm{tr}} \delta\boldsymbol{x}(0). \tag{16}$$

By introducing the adjoint equation

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{a}(t)}{\mathrm{d}t} = {} & -\frac{\partial \boldsymbol{M}_{(1)}(\boldsymbol{x}^*, t)^{\mathrm{tr}}}{\partial \boldsymbol{x}}\,\boldsymbol{a}(t) \\ & + \nabla_x\,D(\boldsymbol{x}^*, t) \qquad (\text{when } t < t_c^{*-}), \end{aligned} \tag{17}$$

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{a}(t)}{\mathrm{d}t} = {} & -\frac{\partial \boldsymbol{M}_{(2)}(\boldsymbol{x}^*, t)^{\mathrm{tr}}}{\partial \boldsymbol{x}}\,\boldsymbol{a}(t) \\ & + \nabla_x\,D(\boldsymbol{x}^*, t) \qquad (\text{when } t > t_c^{*+}), \end{aligned} \tag{18}$$

$$\boldsymbol{a}(T) = 0, \tag{19}$$

$$\boldsymbol{a}(t_c^{*-}) = \boldsymbol{a}(t_c^{*+}), \tag{20}$$

and making use of the fact that D and M are continuous within $[0, T]$, i.e.,

$$\begin{aligned} & D(\boldsymbol{x}^*, t_c^{*-}) - \boldsymbol{a}(t_c^{*-})^{\mathrm{tr}}\,\boldsymbol{M}_{(1)}(\boldsymbol{x}^*, t_c^{*-}) \\ & \quad = D(\boldsymbol{x}^*, t_c^{+}) - \boldsymbol{a}(t_c^{*+})^{\mathrm{tr}}\,\boldsymbol{M}_{(2)}(\boldsymbol{x}^*, t_c^{*+}), \end{aligned} \tag{21}$$

one can get from (15) and (16) the gradient of the functional with respect to the initial state,

$$\left[\frac{\partial J}{\partial \boldsymbol{x}(0)}\right]_{\boldsymbol{x}(0)=\boldsymbol{x}_0^*} = -\boldsymbol{a}(0). \tag{22}$$

Eq. (22) simply indicates that the gradient of the cost functional with respect to the initial state can be calculated by integrating the adjoint eqs. (17)–(20) backward (i.e., from T to 0). It is this gradient that is used to minimize J and, thus, to accomplish the FDDA.

It has been shown that, as long as the numerical model equations are continuous at t_c, which is believed to be so, the variational FDDA using the

adjoint method can be applied to the situations where the numerical model equations are piecewise smooth in the model-state space. The numerical procedure is the same as in the situations where on/off switches do not exist, except that t_c has to be determined when integrating the model forward in order to determine which set of numerical model equations (and, correspondingly, the adjoint equation) are used during the time integration. It should be point out that, when the model equations have terms that are not continuous at t_c, the aforementioned formalism is not applicable. As a matter of fact, it is fundamentally important to the validity of the adjoint method that the model equations and the functional satisfy the weak Lipschitz condition and (7) mentioned in Subsection 2.1. In the following section, the practical application of the above theoretical result to the discrete form of the numerical model equations will be discussed. Although in practice one is not likely to deal with the analytical form of the numerical model equations, the above discussion is the theoretical foundation for the forthcoming description of the practical application of the approach.

3. Formulation of discrete form

Suppose that the measure of the discrepancy between the model state and the observation, $D(\boldsymbol{x}, t)$, is discretized at $N+1$ successive times t_n, where $n=0, 1, 2, ..., N$ and $t_{n+1}-t_n=T/N$. The discrete form of eqs. (6) and (7) may be written synthetically as a recursive system,

$$\boldsymbol{x}_{n+1}=\boldsymbol{F}^n_{(1)}(\boldsymbol{x}_n) \quad \text{for} \quad n=0, 1, 2, ..., j-1, \tag{23}$$

$$\boldsymbol{x}_{n+1}=\boldsymbol{F}^n_{(2)}(\boldsymbol{x}_n) \quad \text{for} \quad n=j, j+1, j+2, ..., N, \tag{24}$$

where x_n is the model state specified at t_n. The recursive system is used only for the sake of simplicity. The technique to be presented is applicable to all time difference schemes.

The time when the on/off switch is performed deserves some attention. Take the latent heat release as an example. If, at a given location, the model atmosphere is not supersaturated prior to t_j, the latent heat release does not take place when the model state is integrated from t_0 to t_j. However, if, at time t_j, the atmosphere is supersaturated, there is latent heat release during the model integration from t_j to t_{j+1}. Therefore, the diabatic forcing due to the latent heat release can be thought of as beginning to influence the system at some time within the interval $[t_j, t_{j+1}]$, although computationally it is not important to know exactly when the time is. All this means is that, conceptually, the switch of the numerical model equations from $F_{(1)}$ to $F_{(2)}$ can be regarded as being performed at some time between t_j and t_{j+1}.

In the discrete form, the functional to be minimized and the Lagrangian functional (corresponding to eqs. (1) and (8)) are, respectively,

$$J=\sum_{n=0}^{N} D(\boldsymbol{x}_n) \tag{25}$$

$$J'=\sum_{n=0}^{N} D(\boldsymbol{x}_n)+\sum_{n=1}^{j} \boldsymbol{a}_n^{\mathrm{tr}}[\boldsymbol{x}_n-\boldsymbol{F}_{(1)}(\boldsymbol{x}_{n-1})] + \sum_{n=j+1}^{N} \boldsymbol{a}_n^{\mathrm{tr}}[\boldsymbol{x}_n-\boldsymbol{F}_{(2)}(\boldsymbol{x}_{n-1})], \tag{26}$$

where

$$D(\boldsymbol{x}_n)=(\boldsymbol{x}_n-\boldsymbol{x}_{n,\mathrm{obs}})^{\mathrm{tr}}\,\boldsymbol{W}_n(\boldsymbol{x}_n-\boldsymbol{x}_{n,\mathrm{obs}}),$$

and $\boldsymbol{a}_n$ is a vector of the discrete adjoint variable. The recursive relations (23) and (24) indicate that $\boldsymbol{x}_n$ is dependent upon $\boldsymbol{x}_0$. Thus, in eq. (25), J is dependent upon $\boldsymbol{x}_0$. It is not necessary to determine the exact time when the criterion condition (5) is satisfied in order to specify j. Rather, j is determined in the computer program by checking whether the components of $\boldsymbol{C}(\boldsymbol{x})$ change sign.

If eqs. (23) and (24) are consistent with their continuous counterpart (i.e., eqs. (6) and (7)), and their solution is convergent and stable, it is rational to assume that *adding an infinitesimally small perturbation to $\boldsymbol{x}_0$ does not change the time interval within which the switch is performed.* This assumption can be justified by the belief that the original analytic differential equation has a unique

and continuous solution. Therefore, the first variation of the functional is

$$\delta J = \sum_{n=0}^{N} \left[\frac{\partial D(\boldsymbol{x}_n)}{\partial \boldsymbol{x}_n}\right]^{\mathrm{tr}} \delta \boldsymbol{x}_n - \boldsymbol{a}_1^{\mathrm{tr}} \frac{\partial \boldsymbol{F}_{(1)}(\boldsymbol{x}_0)}{\partial \boldsymbol{x}_0} \delta \boldsymbol{x}_0 + \boldsymbol{a}_N^{\mathrm{tr}} \delta \boldsymbol{x}_N + \sum_{n=1}^{j-1} \left[\boldsymbol{a}_n^{\mathrm{tr}} - \boldsymbol{a}_{n+1}^{\mathrm{tr}} \frac{\partial \boldsymbol{F}_{(1)}(\boldsymbol{x}_n)}{\partial \boldsymbol{x}_n}\right] \delta \boldsymbol{x}_n + \sum_{n=j}^{N-1} \left[\boldsymbol{a}_n - \boldsymbol{a}_{n+1}^{\mathrm{tr}} \frac{\partial \boldsymbol{F}_{(2)}(\boldsymbol{x}_n)}{\partial \boldsymbol{x}_n}\right] \delta \boldsymbol{x}_n. \quad (27)$$

Introduction of the discrete adjoint equations,

$$\boldsymbol{a}_n = \left[\frac{\partial \boldsymbol{F}_{(2)}(\boldsymbol{x}_n)}{\partial \boldsymbol{x}_n}\right]^{\mathrm{tr}} \boldsymbol{a}_{n+1} - \frac{\partial D(\boldsymbol{x}_n)}{\partial \boldsymbol{x}_n}$$

$$\text{for} \quad n = j, j+1, j+2, \ldots, N, \quad (28)$$

$$\boldsymbol{a}_n = \left[\frac{\partial \boldsymbol{F}_{(1)}(\boldsymbol{x}_n)}{\partial \boldsymbol{x}_n}\right]^{\mathrm{tr}} \boldsymbol{a}_{n+1} - \frac{\partial D(\boldsymbol{x}_n)}{\partial \boldsymbol{x}_n}$$

$$\text{for} \quad n = 1, 2, 3, \ldots, j-1, \quad (29)$$

$$\boldsymbol{a}_{N+1} = 0, \quad (30)$$

gives the gradient of the functional with respect to the initial model state:

$$\frac{\partial J}{\partial \boldsymbol{x}_0} = \frac{\partial D(\boldsymbol{x}_0)}{\partial \boldsymbol{x}_0} - \left[\frac{\partial \boldsymbol{F}_{(1)}(\boldsymbol{x}_0)}{\partial \boldsymbol{x}_0}\right]^{\mathrm{tr}} \boldsymbol{a}_1. \quad (31)$$

The theoretical basis of the above derivation is the assumption that the solution of the continuous model equations exists uniquely (which is presumed by all the numerical models), and that the discrete model equations are consistent, convergent and stable. The results from numerical experiments using a simple convection model, to be presented in the next section, will show the feasibility of the idea. The above discussion also suggests that, in order to guarantee the accuracy of the computed gradient, it is preferable to save the model output at every time step during the minimization procedure because the corresponding switching time in the integration of the adjoint equations has to be determined accurately. In some synoptic situations, practical experience suggests that the basic state required by the adjoint operator does not need to be updated at every time step (Errico and Vukicevic, 1992; Ghil and Malanotte-Rizzoli, 1991). However, for situations where the forcing associated with on/off switches is important, it is necessary to update the basic state at every time step.

4. Numerical experiments with a simple 1-D convection model

Latent heat release is a typical physical process in numerical models of the atmosphere. It appears in the FORTRAN program of the model as IF statements associated with on/off switches, while, mathematically, it is represented by first-order discontinuities in thermodynamic and water-substance equations. If condensation occurs in a four dimensional data assimilation period, latent heat release should be taken into account in the assimilation models in order to assimilate hydrological variables. In the variational assimilation procedure using the adjoint method, corresponding terms regarding the latent heat release should be present in the adjoint equations. However, most current experiences with data assimilation using the adjoint method are based upon research using dry and adiabatic models in which on/off switches do not exist. Therefore, successful implementation of the adjoint method in moist conditions requires an understanding of how to deal with the on/off switches associated with latent heat release.

In the last section, general ideas of how to deal with the on/off switches were developed theoretically. However, the feasibility of these ideas remains to be investigated. In this section, the aforementioned ideas are applied to a heuristic example of FDDA, where the assimilation model is a simple one-dimensional convection model with non-precipitating condensation of water vapor. Despite its simplicity, the assimilation model has the on/off switches that are typical in more realistic and sophisticated assimilation models. We would like to point out that, as parallel independent work, Vukicevic and Errico (1993), Zupanski (1993) and Zou et al. (1993) have recently carried out studies of on/off switches in the adjoint of three-dimensional models. Verlinde (1992) also examined similar discontinuous processes in the study of fitting microphysical observations to a two-dimensional cloud model.

4.1. Numerical model

The governing equations of the one-dimensional model, similar to those described by Rogers and Yau (1989), are

$$\frac{\partial w}{\partial t}+w\frac{\partial w}{\partial z}=\frac{T-T_0}{T_0}g-q_c g, \tag{32a}$$

$$\frac{\partial T}{\partial t}+w\frac{\partial T}{\partial z}=-\frac{g}{c_p}w+\frac{L}{c_p}E, \tag{32b}$$

$$\frac{\partial q_v}{\partial t}+w\frac{\partial q_v}{\partial z}=-E, \tag{32c}$$

$$\frac{\partial q_c}{\partial t}+w\frac{\partial q_c}{\partial z}=E, \tag{32d}$$

where w is the vertical velocity (m/s), T is the temperature (K), T_0 is the temperature of the basic state, q_v and q_c (kg/kg) are the mixing ratios of water vapor and cloud water, respectively, and E is the condensation term which can be either positive or negative. Micro-physical details of the condensation (or evaporation) are neglected because the purpose of using this model is only to illustrate the treatment of the on/off switches. We parameterize the rate of condensation such that it is proportional to the amount of super-saturation (sub-saturation) and the intensity of the updraft (downdraft), i.e.,

$$E=0.1(q_v-q_{vs})\,w$$
$$\text{if}\quad w>0,\qquad (q_v-q_{vs})>0, \tag{33a}$$

$$E=-0.1(q_v-q_{vs})\,w$$
$$\text{if}\quad w<0,\qquad (q_v-q_{vs})<0,\qquad q_c>0$$
$$[(q_v-q_{vs})+q_c]>0, \tag{33b}$$

$$E=0.1q_c w$$
$$\text{if}\quad w<0,\qquad (q_v-q_{vs})<0,\qquad q_c>0,$$
$$[(q_v-q_{vs})+q_c]<0, \tag{33c}$$

or

$$E=0\qquad \text{otherwise.} \tag{33d}$$

The model equations are discretized from $z=0$ to $z=9.8$ km with a grid length of 200 m. A simple forward scheme and an "upstream" scheme are used for the time integration and spatial differencing, respectively. The time step used is 0.1 s. At the lower boundary, the updraft and the gradients of the other variables are prescribed as zero. At the highest model level, the boundary condition is such that, if there is outflow at the boundary, then the value of the variables at this point are computed from the equations because the diffusion processes are neglected in the model.

Identical twin experiments are conducted to examine the behavior of the model initialization using the adjoint method. The functional to be minimized in this study is the squared "distance" between the model state and the corresponding observations:

$$J=\sum_{\substack{i=1\\n=1}}^{I,N} W_w^{i,n}(w^{i,n}-w_{\text{obs}}^{i,n})^2 + W_T^{i,n}(T^{i,n}-T_{\text{obs}}^{i,n})^2+W_{q_v}^{i,k}(q_v^{i,k}-q_{v,\text{obs}}^{i,k})^2 + W_{q_c}^{i,k}(q^{i,k}-q_{c,\text{obs}}^{i,k})^2, \tag{34}$$

where i and n are grid point and time level indices, respectively. The quantities W_x (x denotes w, T, q_v, or q_c) are weighting coefficients reflecting the quality of the data, the subscript "obs" denotes observations, and I and N are the total numbers of grid points and time levels. Because the numerical model is very simple, it is not reasonable to use real observed data. Rather, for illustrative purposes, the strategy is to use a control simulation to generate simulated observations. The quantities W_x are chosen to be unity or zero depending upon the assumptions about the data, where a value of unity implies that there are no errors in the observations. With these "perfect" observations, we will conduct numerical experiments that will enable us to examine the treatment of the on/off switches, and to better understand the behavior of the adjoint method in the data assimilation of hydrological variables.

4.2. Scaling the model variables and verification of the adjoint

Proper scaling of the model variables is critical to the accuracy of the computed gradient of the functional in the situations where the magnitudes of the variables are very different. The magnitudes of the variables arise, of course, simply from the units in which they are expressed. In the one-dimensional convection model used in this study,

the variables vary in magnitude from 10^2 (T in unit of K) to 10^{-3} or 10^{-4} (q_v or q_c in unit of kg/kg). The hydrological variables, q_v and q_c were scaled by 10^5 in this study. Such scaling not only makes the model equations well conditioned with respect to the perturbations of the variables, but also balances the relative "weight" of the variables in the minimization process (Gill et al., 1981; Navon et al., 1992).

The adjoint equations used for the data assimilation experiments and the boundary conditions for the adjoint variables are derived directly from the finite difference form of the model equations using the Lagrangian-multiplier method described in the last section. This method is identical to that suggested by Thacker and Long (1988). It is important to verify that the gradient of the functional thus obtained is numerically correct. One verification method, suggested by Navon (personal communication, also see Navon et al., 1992, and Sun et al., 1991), is to compare the linear approximation of the perturbation value of the functional with the true value. Let $\boldsymbol{x}$ denote an arbitrary state, h a small parameter and $\boldsymbol{r}$ a random vector of unit length. The existence of the first variation (or Gâteaux derivative) of J (see (5)) leads to

$$J(\boldsymbol{x}+h\boldsymbol{r})-J(\boldsymbol{x})=h\boldsymbol{g}(\boldsymbol{x})^{\mathrm{tr}}\,\boldsymbol{r}+\Delta(h\boldsymbol{r}), \tag{35}$$

where $\boldsymbol{g}(\boldsymbol{x})$ is the gradient of the functional, and $\Delta(h\boldsymbol{r})/h \to 0$ as $h \to 0$. A function of h, $F(h)$, may be defined by rearranging eq. (35) as:

$$F(h)=\frac{J(\boldsymbol{x}+h\boldsymbol{r})-J(\boldsymbol{x})}{h\boldsymbol{g}(\boldsymbol{x})^{\mathrm{tr}}\,\boldsymbol{r}}=1+\frac{\Delta(h\boldsymbol{r})/h}{\boldsymbol{g}(\boldsymbol{x})^{\mathrm{tr}}\,\boldsymbol{r}}. \tag{36}$$

If the gradient $\boldsymbol{g}(\boldsymbol{x})$ obtained by using the adjoint method is correct, $F(h)$ should approach unity as h approaches zero. Table 1 shows the value of $F(h)$ versus h. It can be seen that, for the control experiment in which latent heat release is turned on, $F(h)$ approaches unity as h decreases, for h between 10^{-8}–10^{-3} (see the first column in Table 1). The higher values of $F(h)$ for h larger than 10^{-3} are attributed to the significance of the higher order terms due to the nonlinearity. When h is very small ($\leqslant 10^{-9}$), the departure of $F(h)$ from 1 is due to the numerical truncation error of the computer. It is interesting to note in Table 1, by comparing the control experiment with the experiment in which the latent heat release is turned off, that the condensation enhances the nonlinearity of the functional. Without condensation, $F(h)$ is within 8.1% of unity for h as high as 1, whereas in the control experiment, $F(h)$ is larger than 2 for $h=1$.

It should be noted that the safest way of checking the numerical accuracy of the adjoint equations is to compare the numerical output of the tangent linear equations, which are the first variation of the model equations, with that of the adjoint equations (see, e.g., Errico and Vukicevic, 1992). However, performing this check requires the availability of a computer program of the tangent linear equations. As mentioned above, the adjoint equations of the one-dimensional convection model used in this study were developed using the Lagrangian-multiplier method. The tangent linear equations were thus not coded. Therefore, the method mentioned in the proceeding paragraph was used in checking the numerical accuracy of the adjoint equations. It should be pointed out that the effective use of this method requires a certain degree of care. That is, the value of h, below which round-off errors become significant in the evaluation of $F(h)$, needs to be estimated. To make the method effective, h must decrease until the value is reached. As a rule of thumb, the value has approximately the magnitude of $10^{-m/2}$ where m is the number of significant digits of precision. In this study, because m is 15, the value has the magnitude of around 10^{-8} (as shown in Table 1, below this value, the round-off errors are significant).

Table 1. *Verification of the gradient*

$\log_{10}(h)$	$F(h)$ (with condensation/ evaporation)	$F(h)$ (without condensation/ evaporation)
0	2.137621395	1.080197902
−1	1.132705781	1.025976279
−2	1.031433995	1.020549579
−3	1.021299018	1.020006860
−4	1.002028551	1.001995249
−5	1.000201844	1.000199467
−6	1.000020176	1.000019940
−7	1.000002023	1.000001992
−8	1.000000103	1.000000195
−9	1.025512177	0.999999989
−10	1.046721078	0.991896180
−11	0.256721895	0.787615027

4.3. Initialization experiments

Since the purpose of the numerical experiments is to illustrate the feasibility of data assimilation using the adjoint of a numerical model that involves on/off switches, identical twin experiments within the framework of the one-dimensional convection model are used in the initialization problem. Observational data are generated by the control simulation. When being used for FDDA experiments, these data are assumed to be error free and available at each model time-step. Fig. 2 shows the initial state of the control simulation. The profile of temperature is assumed to have a lapse rate of 6° C/km (Fig. 2a). The profile of pressure is deduced using the hydrostatic relation based upon the temperature profile, assuming the surface pressure is 1000 mb. Fig. 2b shows the initial profile of q_v, and Fig. 2c depicts the profile of $(q_v - q_{vs})$, where q_{vs} is the saturation mixing ratio. q_v is specified such that relative humidity is 98 % at every grid point. At the initial time, there is a preexisting up-draft with a parabolic velocity profile shown in Fig. 2d. It is assumed that there is

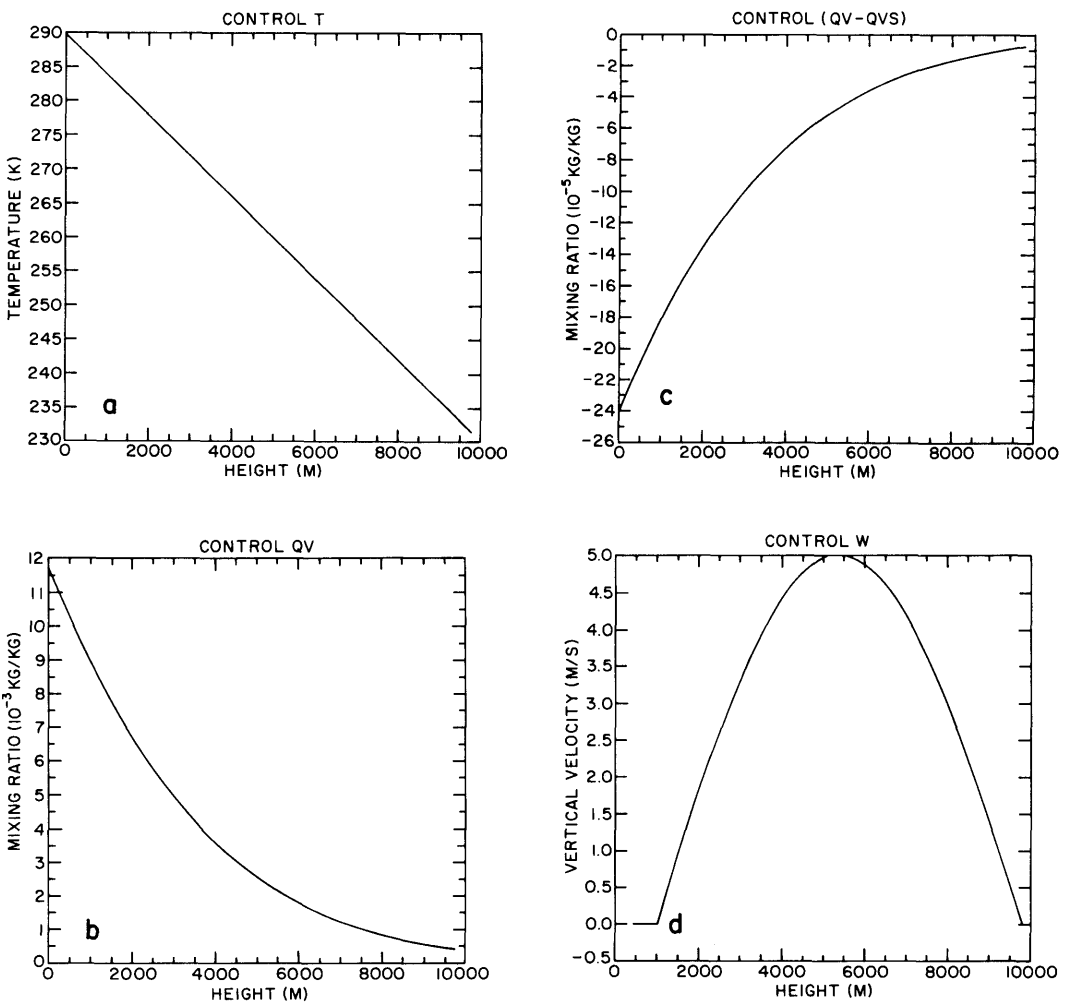

Fig. 2. Initial conditions for the control experiment: (a) profile of T, (b) profile of q_v, (c) profile of $(q_v - q_{vs})$, and (d) profile of w.

no cloud water at the initial time. Fig. 3 shows the results of a 6-min simulation, with a display time interval of 1 min. The dynamics of the model simulation can be described as a moist-buoyant oscillation. As an air parcel ascends, its temperature decreases due to the adiabatic cooling. The condensation takes place when the air parcel cools to saturation and cloud droplets form. Since the parameterized rate of condensation is not large, the latent heat release does not offset the adiabatic cooling. Thus, the air parcel continues cooling. However, the cooling rate is less than it would be without the latent heat release. Since the rising parcel is cooler than its environment, the parcel tends to return to its original level. When the parcel descends, adiabatic warming makes the parcel sub-saturated and, thus, evaporation of the cloud droplets occurs. The evaporation reduces the rate of adiabatic warming, but it is not large enough to offset the adiabatic warming.

It is assumed that the length of the assimilation window is 10 s (100 time steps), and that observa-

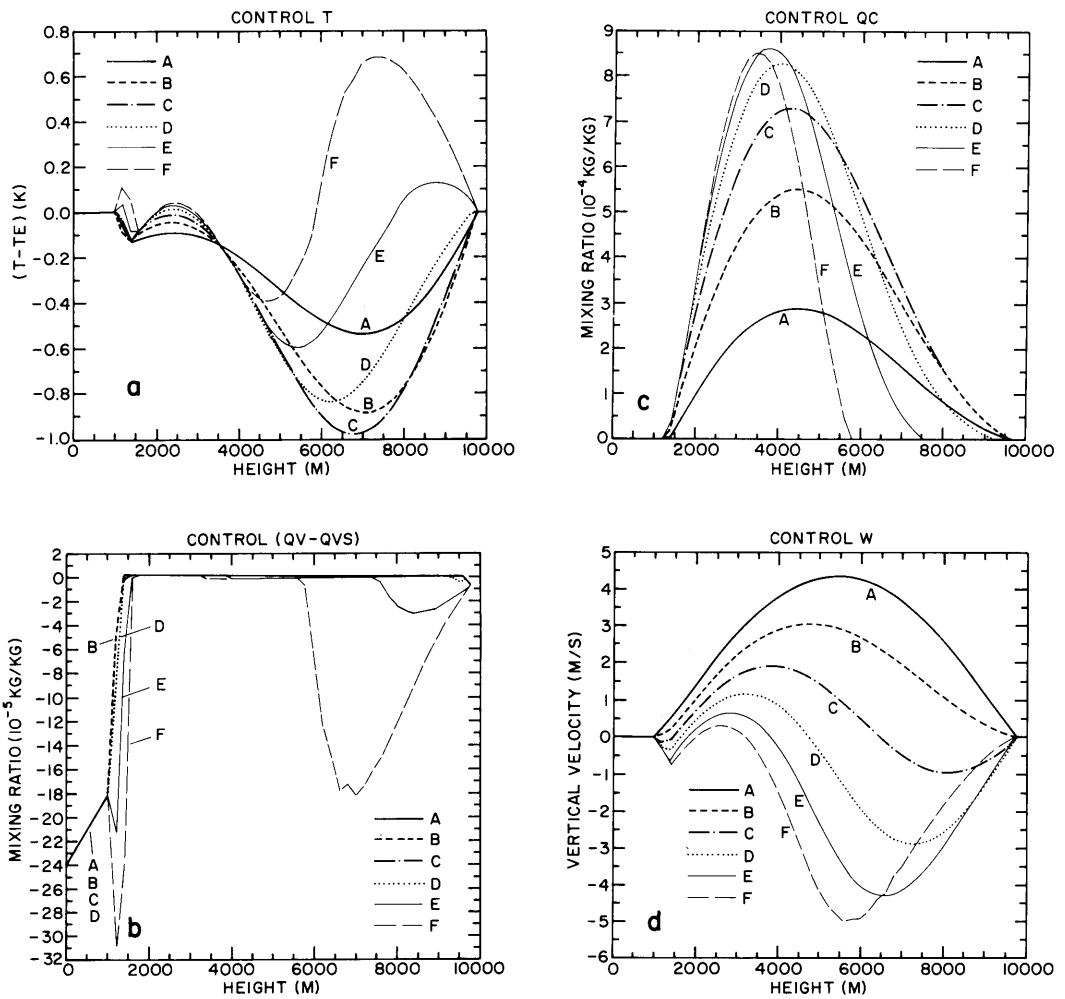

Fig. 3. Results of 6-min simulation with display time interval of 1 min: (a) profile of T, (b) profile of $(q_v - q_{vs})$, (c) profile of q_c, and (d) profile of w. The letters labeling the curves represent the simulation time, i.e., A for 1 min., B for 2 min., ..., F for 6 min.

tions are available at each model time step within this period. The basic state used in the adjoint is updated at every time step. The minimization algorithm used in the experiments is the limited memory quasi-Newton method of Liu and Nocedal (1989). A brief summary of this algorithm can also be found in Navon et al. (1991).

The first guess of the initial conditions is specified such that the vertical velocity (w) is zero uniformly. The temperature (T) has the same distribution as in the control. The relative humidity is 92% everywhere, while in the control simulation it is initially 98% everywhere. It is assumed that observations are available at every grid point and at every time-level for w, T, q_v and q_c. Fig. 4 shows the values of the functional and the norm of the gradient as functions of the iteration number. Small magnitudes of the functional value and the norm of the gradient make it clear that the minimization procedure converges. Therefore, although there are on/off switches involved in the assimilation model, the FDDA procedure using the adjoint method is capable of retrieving the initial state.

The existence of on/off switches makes the functional more nonlinear than it would be without the on/off switches, and thus the minimization would take more iterations. Fig. 5 depicts the reduction in the values of the functional and its gradient versus the number of iterations for an experiment where condensation (or evaporation) is not allowed to occur during the control simulation and the assimilation. Compared with Fig. 4, one can see that the minimization procedure converges much faster than it does for the experiment where condensation (or evaporation) is taken into account. It is known that small errors in the estimation of the gradient of the functional will slow down the convergence. It has been seen in this study that the existence of the on/off switches makes the minimization slow, although the basic state used in the adjoint is updated at every time step. As discussed in Section 3, when the assimilation model involves the on/off switches, errors will be introduced in the gradient of the functional if the basic state used in the adjoint is not updated at every time step. Therefore, the results obtained in this study suggest that the basic state should be saved at every time step in order to make the minimization efficient.

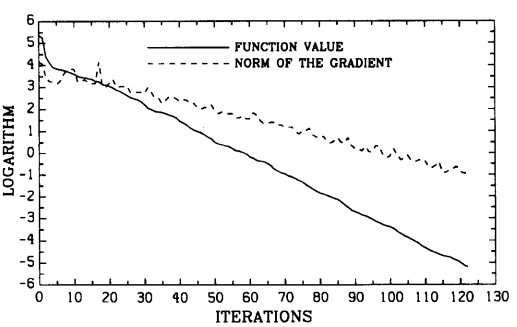

Fig. 4. Values of the functional and the norm of the gradient as a function of iteration number for the experiment with the condensation of water vapor.

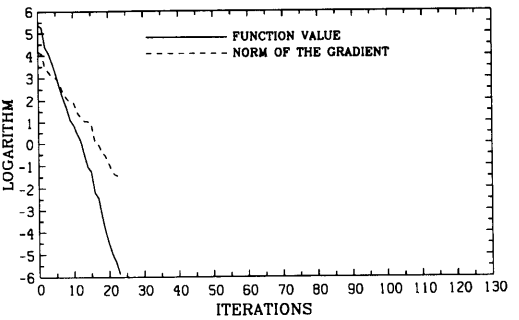

Fig. 5. Same as Fig. 4, but for the experiment without the condensation of water vapor.

5. Summary and conclusions

The calculus of variations has been used to justify that variational FDDA using the adjoint method can be performed successfully, even though the numerical model equations have a finite number of first-order discontinuities. These points represent the on/off switches for which the Jacobian matrix of the model equations does not exist. In practice, when the on/off switches are involved, the gradient of the functional can be obtained by assuming that an infinitesimally small initial perturbation does not change the time interval within which the switching takes place. The time when the corresponding switching in the adjoint equations is performed is the same as that in the forward integration of the model equations. That is, the switching time is determined by the basic state.

Numerical FDDA experiments are then conducted using the aforementioned ideas in a simple one-dimensional convection model that involves the on/off switches associated with latent heat release. The implication of this study to data assimilation using the adjoint method is that there is no difficulty in theory in treating the on/off switches in the numerical model. Rather, the success of the adjoint method is dependent on the numerical aspects of its implementation, such as the proper scaling of the variables. This study also suggests that, in order to guarantee the accuracy of the computed gradient, it is preferable to save the model output at every model time-step during the minimization procedure, because the corresponding switching time in the integration of the adjoint equations has to be determined accurately.

The emphasis of this study is to illustrate the feasibility of using the adjoint method for data assimilation, where the numerical model involves on/off switches associated with latent heat release. Although the assimilation model that we have used in this study is very simple, it has the on/off switches that are typical in more sophisticated assimilation models. Obviously, further research using more realistic three-dimensional models is needed to assess the practical potential of the adjoint method in generating atmospheric circulations and structures whose evolution not only involves latent heating, but also PBL processes, topographic variations, and radiative heating.

6. Acknowledgements

We thank Professor I. M. Navon for providing us with the minimization software used in this study. We are very grateful to two anonymous reviewers for their critical comments. This research is supported by NASA Grant No. NGT 30040, NSF/FAA Grant No. ATM-9203317, ONT Contract No. N00039-92-C-0100, and NSF Grant No. ATM-9116176. Participation at the Workshop on Adjoint Applications in Dynamic Meteorology was funded by NASA and NSF.

REFERENCES

Cacuci, D. G. 1981. Sensitivity theory for nonlinear system. I. Nonlinear functional analysis. *J. Math. Phys.* **22**, 2794–2802.

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

Daley, R. 1991. *Atmospheric data analysis.* Cambridge University Press, 457 pp.

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

Errico, R. M. and Vukicevic, T. 1992. Sensitivity analysis using an adjoint of the PSU/NCAR mesoscale model. *Mon. Wea. Rev.* **120**, 1644–1660.

Ghil, M. and Malanotte-Rizzoli, P. 1991. Data assimilation in meteorology and oceanography. *Adv. Geophys.* **33**, 141–265.

Gill, P. E., Murray, W. and Wright, M. H. 1981. Practical Optimization, Academic Press, 401 pp.

Girard, C. and Jarraud, M. 1982. *Short- and medium-range forecast differences between a spectral and grid point model: an extensive quasi-operational comparison.* Tech. Rep. No. 32, ECMWF, 176 pp. Available from ECMWF, Shinfield Park, Reading, England, RG2 9AX.

Heckley, W. A. 1985. Systematic errors of the ECMWF operational forecasting in the tropics. *Quart. J. Roy. Meteor. Soc.* **111**, 1551–1556.

Kasahara, A., Balgovind, R. C. and Katz, B. B. 1988. Using of satellite radiometric imagery data for improvement in the analysis of divergent wind in the tropics. *Mon. Wea. Rev.* **116**, 866–883.

Krishnamurti, T. N., Bedi, H. S., Heckley, W. and Ingles, K. 1988. Reduction of the spin-up time for evaporation and precipitation in a spectral model. *Mon. Wea. Rev.* **116**, 907–920.

Lewis, J. M. and Derber, J. C. 1985. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus* **37A**, 309–327.

Liu, D. C. and Nocedal, J. 1989. On the limited memory BFGS method for large scale optimization. *Mathematical Programming* **45**, 1177–1194.

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

Navon, I. M., Zou, X., Berger, M., Phua, P. K. H. and Hopgood, M. 1991. Numerical experience with limited-memory quasi-Newton Methods for large-scale unconstrained nonlinear minimization. SCRI Technical Report, FSU-SCRI-91-14, Tallahassee, FL 32306-4052, 25 pp.

Rogers, R. R. and Yau, M. K. 1989. *A short course in cloud physics.* Pergamon Press, 293 pp.

Saaty, T. L. 1981. *Modern nonlinear equations.* Dover Publication, Inc., 471 pp.

Sagan, H. 1969. *Introduction to the calculus of variations.* Dover Publication, Inc., 449 pp.

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

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

Thacker, W. C. and Long, B. R. 1988. Fitting dynamics to data. *J. Geophys. Res.* **93**, 1227–1240.

Verlinde, J. 1992. *Fitting microphysical observations to a numerical model through an optimal control theory technique.* PhD Thesis, Colorado State University, Fort Collins, Colorado 80353, Paper No. 493, 83 pp.

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

Zou, X., Navon, I. M. and Sela, J. 1993. Variational data assimilation with threshold processes using the NMC multilevel primitive-equation model. *Tellus* **45A**, 370–387.

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