NOTES AND CORRESPONDENCE

The Role of Time Step Size in Numerical Stability of Tangent Linear Models

JIANG ZHU

Meteorological Research Institute, Ibaraki, Japan, and
Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

MASAFUMI KAMACHI

Meteorological Research Institute, Ibaraki, Japan

12 July 1999 and 13 October 1999

ABSTRACT

It is found that some stable time-integration schemes for some nonlinear models do not guarantee stable integrations of the associated tangent linear models with the same time step size. These problems usually occur when the nonlinear models describe vertical diffusion processes and are numerically implemented by semi-implicit time-integration schemes that are unconditionally stable. The direct linearization procedure performed on such numerical schemes of nonlinear models can be interpreted as some conditionally stable numerical schemes of the underlying linearized equations.

Numerical experiments using a simple, illustrative model and a realistic ocean mixed layer model and their tangent linear models showed instabilities in the tangent linear models. Several methods are tried to reduce the nonphysical noise caused by the numerical instabilities. This study suggests that reducing time step size can give good results compared to some other methods that either are not accurate enough or change too much from the original nonlinear model.

1. Introduction

A tangent linear model (TLM), which determines the evolution of errors in a forecast model, plays an important role in extended Kalman filtering and variational data assimilation. Construction of a TLM is an intermediate step in construction of the adjoint model of a nonlinear model. A TLM also serves as the forward model in the incremental approach to variational data assimilation (Courtier et al. 1994). Other use of TLMs includes stability analysis (e.g., Ehrendorfer and Errico 1995).

The ultimate goal in developing a TLM is to achieve an accuracy that it can well describe behavior of perturbations that correspond to ones in the nonlinear model as the perturbation size increase to the size of *actual uncertainties*, rather than in the infinitesimal limit (Errico and Reader 1999). However, there are three major causes of inaccuracy in some TLMs. One is due to finite magnitude of the perturbations in initial conditions, boundary conditions, or model parameters. The second is due to strong nonlinearities, discontinuities in nonlinear models, which is often created by some embedded parameterizations of fast physical processes. The third is due to numerical instabilities in some TLMs. Recently reported by Errico et al. (1993), Mahfouf (1999), Errico and Reader (1999), and Janiskova et al. (1999) that large noises exist and grow in the TLM solution in some atmospheric models with some physical processes. We also found this problem during development of the TLM of an ocean mixed layer (OML) model of the Mellor–Yamada type (Mellor and Yamada 1982). With a resolution 1 m in vertical direction and a time step size of 30 min, the nonlinear model always runs smoothly, but its TLM integration with basic state updates each time step is not stable, which usually is blown up after several days of integration.

Some possible mechanisms of instability in some TLMs were discussed by Mahfouf (1999). By analyzing the underlying partial differential equations (PDEs), which are somehow simplified, he illustrated that positive feedback in linearizations of physical parameterizations can lead to an exponential growth of initial perturbations, that the lack of nonlinear effects will prevent from saturating with time. This argument can explain instabilities of numerical schemes of TLMs when the underlying PDE is unstable, but may not be convincing for explaining instabilities of numerical schemes of TLMs when underlying PDE is stable.

The methods to reduce the noise in TLMs are dis-

Corresponding author address: Dr. Jiang Zhu, Meteorological Research Institute, 1-1 Nagamine, Tsukuba, 305-0052 Ibaraki, Japan.
E-mail: jzhu@mri-jma.go.jp

cussed and developed by several groups of researchers. These methods include smoothing and removing some types of discontinuities (e.g., Verlinde and Cotton 1993; Zupanski and Mesinger 1995; Tsuyuki 1996; Janiskova et al. 1999), omitting some terms in TLMs (e.g., Errico et al. 1993; Mahfouf 1999) and an approximated Jacobian method (Errico and Reader 1999) that uses nonlinear perturbations to calculate elements of the Jacobian of a parameterization.

The goals of the study are: first, looking for genesis of numerical instabilities existing in some TLM integrations, especially for TLMs of nonlinear vertical diffusion processes; second, to propose some alternative approach in order to reduce the nonphysical noise in some TLMs, based on the generating mechanism of numerical instabilities. This note is arranged as follows: in section 2, the underlying PDEs of some TLMs of vertical diffusion models are discussed and numerical implementations of these TLMs are discussed. Stability for their numerical schemes is addressed in section 3. Some numerical results are given in section 4 using a simple, illustrative example and an ocean mixed layer model based on the Mellor–Yamada turbulence closure scheme. Section 5 is discussions and conclusions.

2. Underlying PDE of tangent linear models

a. Vertical diffusion process for velocity components

Consider the following one-variable PDE, which describes vertical diffusive process for a horizontal component of the velocity

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial z}\left(K\frac{\partial u}{\partial z}\right), \tag{1}$$

where K is the eddy viscosity coefficient, which depends on u. We assume that K is parameterized locally by gradient Richardson number Ri. Then it is also a function of density and velocity shear. To simplify our analysis, we assume in this subsection that K is only a function of $\partial u/\partial z$, that is $K = K(\partial u/\partial z)$. This assumption is only valid for some vertical diffusion process regimes where $\partial u/\partial z$ is the dominating factor of diffusion. If we write $x = \partial u/\partial z$, then the following two conditions should hold for K:

- $K(x)$ is a nondecreasing function for $x \geq 0$; (2)
- $K(x) = K(-x) > 0$. (3)

Condition (2) describes the fact that the eddy viscosity coefficient is increasing with the velocity shear. Condition (3) says that the eddy viscosity coefficient is determined by the magnitude of the velocity shear.

If $K(x)$ is differentiable or only has some isolated nondifferentiable points, by combining (2) and (3), one has

$$xK'(x) \geq 0 \quad \text{for all } x. \tag{4}$$

Here $K'(x)$ is the derivative of $K(x)$ with respect to x.

The tangent linear equation (TLE) of (1) reads

$$\frac{\partial \delta u}{\partial t} = \frac{\partial}{\partial z}\left[G\left(\frac{\partial u}{\partial z}\right)\frac{\partial \delta u}{\partial z}\right] \tag{5}$$

where

$$G\left(\frac{\partial u}{\partial z}\right) = K\left(\frac{\partial u}{\partial z}\right) + \frac{\partial u}{\partial z}K'\left(\frac{\partial u}{\partial z}\right). \tag{6}$$

From (2)–(4), $G(\partial u/\partial z) \geq 0$, which means that Eq. (5) is also a diffusion PDE with positive diffusion coefficient. Hence, the TLE is a well-defined parabolic PDE, and its solution will never grow from initial conditions if there is no boundary flux or forcing according to the maximum principle (see Dautray and Lions 1988). When there is forcing or boundary flux, the solutions are bounded by the norms of the forcing and boundary flux over any span of time. So the genesis of numerical instabilities in TLMs cannot be found in their underlying PDEs.

b. Vertical diffusion process for temperature

Now we consider the vertical diffusion process for temperature that can be simplified as

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left(K\frac{\partial T}{\partial z}\right), \tag{7}$$

where K is the diffusion coefficient, which depends on T only. Here we assumed that K is parameterized locally by the gradient Richardson number. That is, $K = K(\partial\rho/\partial z)$. In the atmosphere and ocean mixed layers, the density ρ is not linearly dependent on T. As a consequence, the diffusion coefficient K cannot be expressed as $K = K(\partial T/\partial z)$ as in the case for velocity component. We here write $K = K(T)$.

Taking the first variation of (7) gives

$$\frac{\partial \delta T}{\partial t} = \frac{\partial}{\partial z}\left(K\frac{\partial \delta T}{\partial z}\right) + \frac{\partial}{\partial z}\left[K'(T)\frac{\partial T}{\partial z}\delta T\right], \tag{8}$$

which is no longer a vertical diffusion equation but a general flux-conservative equation. This is also a well-defined PDE. Its solution will remain bounded as time tending to infinite according to the conservative law of flux if the boundary flux and external forcing have finite integration over infinite time. Equation (8) can also be interpreted as an advection-diffusion PDE for the linear perturbation δT where $w = K'(T)\partial T/\partial z$ is the ''vertical velocity'' and the second term in the right-hand side (rhs) of (8) can be regarded as a divergence of vertical flux of the linear perturbation δT. From this observation, any discrete time integration schemes of (8) can only be conditionally stable if $K'(T)\partial T/\partial z \neq 0$.

3. Numerical implementations

Now we look at numerical implementations of TLM for genesis of instabilities.

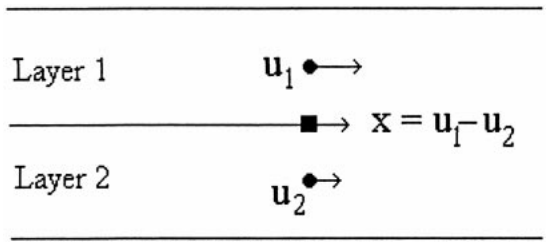

FIG. 1. A schematic of a two-layer diffusive model. Here, u_1 (u_2) is the velocity in the Layer 1 (Layer 2) and x is the vertical shear.

A TLM is usually numerically implemented by running the codes that are written directly from the nonlinear model codes (Giering and Kaminski 1998). Below we discuss stability of some TLMs that are constructed this way.

a. Vertical diffusion for velocity components

1) AN ODE EXAMPLE

First consider the following simple scale ODE

$$\frac{dx}{dt} = -K(x)x, \tag{9}$$

which describes the vertical diffusive process of a two-layer fluid in a highly simplified manner. Figure 1 gives a schematic diagram with $x = u_1 - u_2$ being the vertical shear and $K(x)$ is the eddy viscosity coefficient. Here function K is assumed to satisfy (2) and (3). Though very simple, the example can illustrate well the main idea of the short note.

Now consider the semi-implicit scheme:

$$x_{n+1} = x_n - \Delta t K(x_n)x_{n+1}, \quad \text{or} \tag{10}$$

$$x_{n+1} = \frac{1}{1 + \Delta t K(x_n)}x_n, \tag{11}$$

which is obviously unconditionally stable. Then taking the first variation of (11) gives the TLM

$$\delta x_{n+1} = \frac{[1 + \Delta t K(x_n)] - \Delta t x_n K'(x_n)}{[1 + \Delta t K(x_n)]^2}\delta x_n$$

$$= \frac{1 - \Delta t x_{n+1} K'(x_n)}{1 + \Delta t K(x_n)}\delta x_n, \tag{12}$$

which may not necessarily be unconditionally stable for every function K with properties (2) and (3). Some functions with properties (2) and (3) make (12) unstable for time step size not small enough, that is,

$$\frac{1 - \Delta t x_{n+1} K'(x_n)}{1 + \Delta t K(x_n)} < -1.$$

Unfortunately, functions used in some vertical diffusion parameterization based on turbulence closure schemes (e.g., the Mellor–Yamada scheme) are among these functions. We will show this in the next section in detail.

Above we have shown using a one-variable model that the stability of time integration of a nonlinear model does not necessarily guarantee the stability of time integration of its TLM.

2) THE SEMI-IMPLICIT SCHEME FOR (1)

Now consider the semi-implicit scheme for Eq. (1):

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{1}{\Delta z^2}[K_j^n(u_{j+1}^{n+1} - u_j^{n+1}) - K_{j-1}^n(u_j^{n+1} - u_{j-1}^{n+1})], \tag{13}$$

where

$$K_j^n = K(u_{j+1}^n - u_j^n). \tag{14}$$

The TLM is

$$\frac{\delta u_j^{n+1} - \delta u_j^n}{\Delta t}$$

$$= \frac{1}{\Delta z^2}[K_j^n(\delta u_{j+1}^{n+1} - \delta u_j^{n+1}) - K_{j-1}^n(\delta u_j^{n+1} - \delta u_{j-1}^{n+1})]$$

$$+ \frac{1}{\Delta z^2}[H_j^n(\delta u_{j+1}^n - \delta u_j^n) - H_{j-1}^n(\delta u_j^n - \delta u_{j-1}^n)], \tag{15}$$

where

$$H_j^n = K'(u_{j+1}^n - u_j^n)(u_{j+1}^{n+1} - u_j^{n+1}).$$

Looking at (15) one can find that the time-integration scheme is no longer a semi-implicit scheme. The second term of the right-hand side of (15) makes the TLM a mixture of the semi-implicit and explicit schemes. The stability of (15) is not straightforward. To gain some insights of its stability theoretically, we assumed that $u_{j+1}^n - u_j^n$ are spatially homogeneous, or $u_{j+1}^n - u_j^n = \Delta u$, for all j. Thus K_j^n and H_j^n are also spatially homogeneous, and denoted by K and H, respectively. Then (15) can be rewritten as

$$\frac{\delta u_j^{n+1} - \delta u_j^n}{\Delta t}$$

$$= (K + H)\left\{\frac{\theta}{\Delta z^2}[(\delta u_{j+1}^{n+1} - \delta u_j^{n+1}) - (\delta u_j^{n+1} - \delta u_{j-1}^{n+1})]\right.$$

$$\left. + \frac{1 - \theta}{\Delta z^2}[(\delta u_{j+1}^n - \delta u_j^n) - (\delta u_j^n - \delta u_{j-1}^n)]\right\}, \tag{16}$$

where $\theta = K/(K + H)$ and $0 \le \theta \le 1$. The numerical scheme (16) is the generalized Crank–Nicholson scheme, and its stability condition is as follows (see, e.g., Haltiner and Williams 1980):

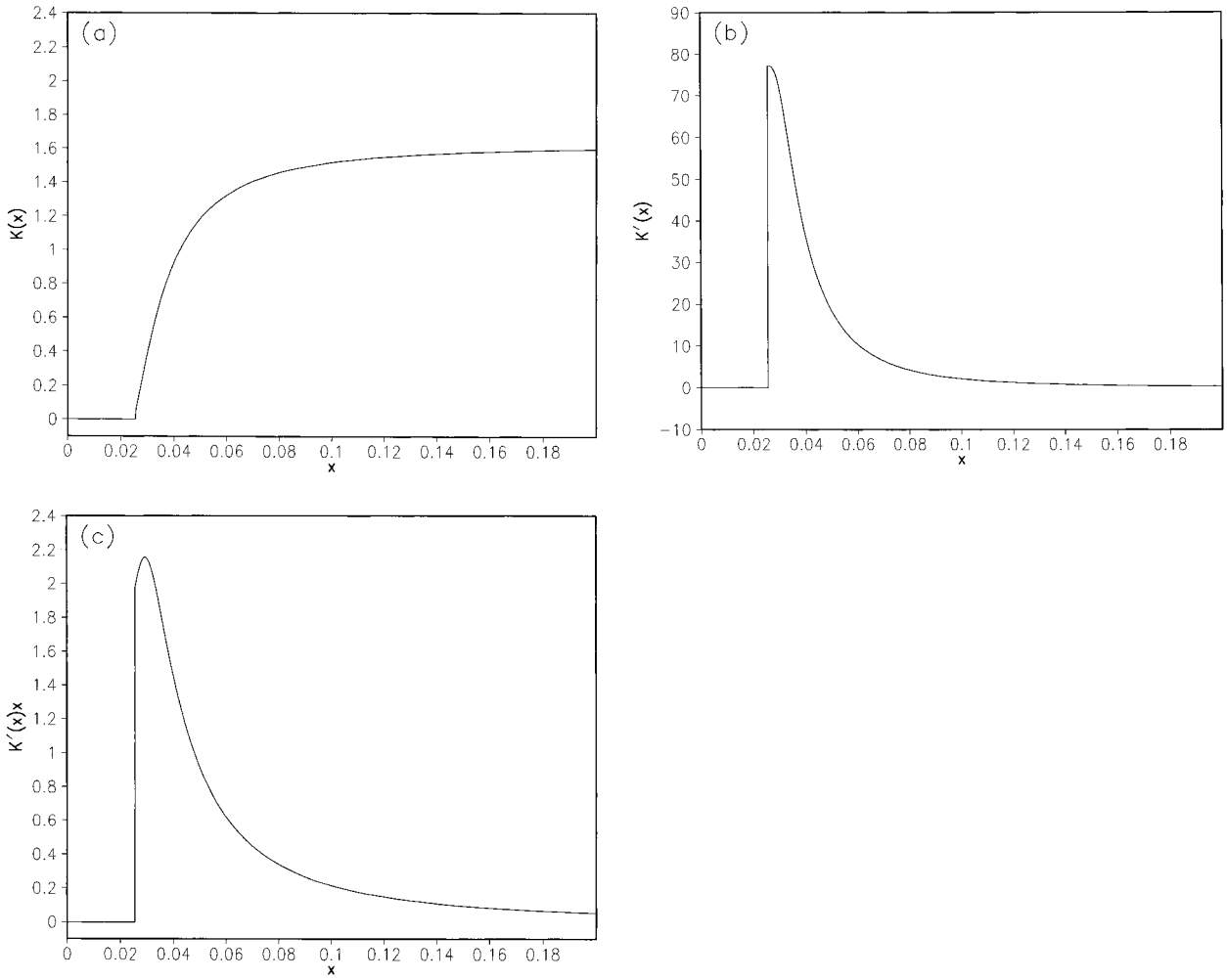

FIG. 2. Function K and its derivative: (a) $K(x)$, (b) $K'(x)$, (c) $xK'(x)$.

$$(K + H)\Delta t/\Delta z^2 \le 1/(2 - 4\theta) \quad \text{if } 0 \le \theta < 1/2$$
$$\text{no restriction on } \Delta t, \quad \text{if } 1/2 \le \theta \le 1. \tag{17}$$

In another word, if $H > K$, (16) is only conditionally stable and the time step size should be restricted by $\Delta z^2/(2H - 2K)$.

Since $H(x) \le xK'(x)$, the condition $xK'(x) \le K(x)$ is important in stability of (15) and (16). We will see in the next section that this condition may not be held in some vertical mixing parameterization of OML. Even in the case of $\frac{1}{2} \le \theta < 1$ in (17), an instability can occur when a gradient boundary condition (Neumann boundary condition) is involved (Roache 1972).

The linearization process changes the nature of the time-integration schemes. The nonlinear model has a time-integration scheme of unconditionally stable, but the linearized model may only have a time-integration scheme of conditionally stable. When the time step used in the nonlinear model is much larger than that required by the stability of TLM, the nonphysical noise in TLM will be generated and will make the linearized model solution meaningless.

b. Vertical diffusion process for temperature

It is shown in the last section that the underlying PDE for TLM is of a different type from that of the nonlinear model. The major difference between a diffusion equation and a general flux-conservative equation is that the (semi-) implicit time-integration scheme for a diffusion equation is unconditionally stable, but only conditionally stable for a general flux-conservative equation (see Press et al. 1992).

Consider the semi-implicit time-integration scheme of (7):

$$\frac{T_j^{n+1} - T_j^n}{\Delta t} = \frac{1}{\Delta z^2}[K_j^n(T_{j+1}^{n+1} - T_j^{n+1}) - K_{j-1}^n(T_j^{n+1} - T_{j-1}^{n+1})], \tag{18}$$

where

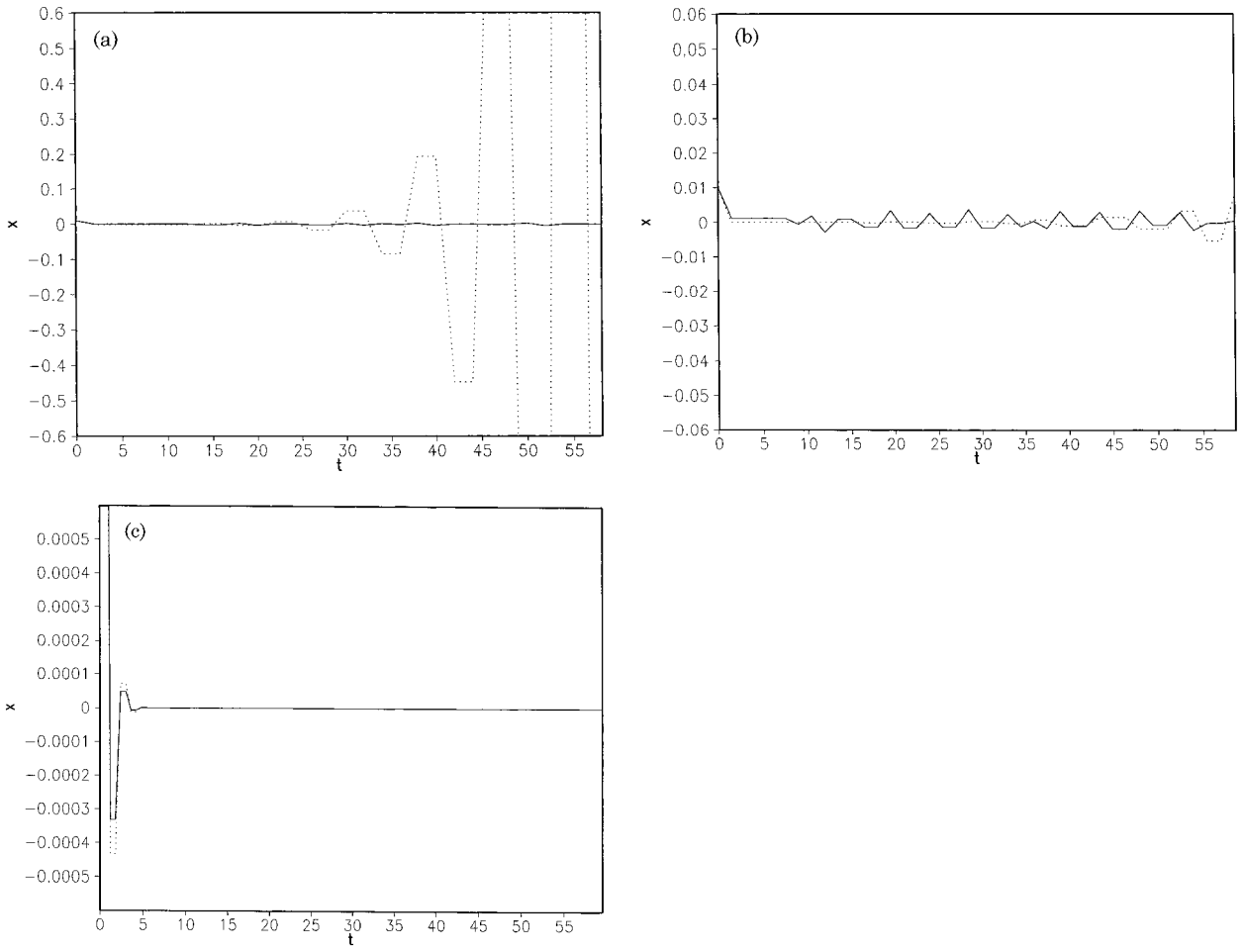

FIG. 3. Accuracy of the TLM solutions for the simple ODE example. The solid lines are for nonlinear perturbation. Dotted lines are for TLM. (a) $\Delta t = 2.0$, (b) $\Delta t = 1.5$, (c) $\Delta t = 0.6$.

$$K_j^n = K(T_j^n, T_{j+1}^n). \tag{19}$$

The associated TLM is

$$\frac{\delta T_j^{n+1} - \delta T_j^n}{\Delta t} = \frac{1}{\Delta z^2}[K_j^n(\delta T_{j+1}^{n+1} - \delta T_j^{n+1}) - K_{j-1}^n(\delta T_j^{n+1} - \delta T_{j-1}^{n+1})] + \frac{1}{\Delta z}\left[\frac{T_{j+1}^{n+1} - T_j^{n+1}}{\Delta z}(F_j^n \delta T_j^n + G_j^n \delta T_{j+1}^n) - \frac{T_j^{n+1} - T_{j-1}^{n+1}}{\Delta z}(F_{j-1}^n \delta T_{j-1}^n + G_{j-1}^n \delta T_j^n)\right], \tag{20}$$

where

$$F_j^n = \left.\frac{\partial K(x, y)}{\partial x}\right|_{x=T_j^n, y=T_{j+1}^n}, \qquad G_j^n = \left.\frac{\partial K(x, y)}{\partial y}\right|_{x=T_j^n, y=T_{j+1}^n}. \tag{21}$$

The numerical scheme (20) can be interpreted as a mixture of a semi-implicit scheme for the diffusion term (first term of rhs) and an explicit scheme for the ''advection'' term (the second term of rhs). This scheme is not unconditionally stable, but some stable conditions can be easily obtained if some assumptions on basic state solution are made such as in the previous subsection. If the time step size in the nonlinear model is larger than that required by the stability of the TLM, the TLM

TABLE 1. Descriptions of experiments.

Case		Wind stress (N m^{-2})	Heat flux (ly/day)	Solar radiation	Integration time
I: Heating	Bas.	0.1	150	No	5 days
	Ptb.	0.1	151		
	TLM	0.0	1		
II: Wind deepening	Bas.	0.5	0	No	1 day
	Ptb.	0.51	0		
	TLM	0.01	0		

Bas.: basic state model run; Ptb.: perturbation model run.

TABLE 2. The comparison of nonlinear perturbation and TLM for heating experiments at the end of 5-day integration with different depth-level K.

Level	Var.	Nonlinear (°C)	TLM (°C)
a. $\Delta t = 1800$ s			
$K = 1$	T*	5.655381525198067e−03	−405.9745834956914
$K = 3$	T	5.632905061450089e−03	−414.0980890641895
$K = 5$	T	5.609360872774971e−03	−436.5769306752263
$K = 7$	T	5.581235746237923e−03	−478.5665910639509
$K = 9$	T	5.545305671020628e−03	−545.5960435821180
b. $\Delta t = 900$ s			
$K = 1$	T	5.471445244438655e−03	−0.9509791165644970
$K = 3$	T	5.452719200661704e−03	−1.122054799023772
$K = 5$	T	5.437165248633136e−03	−1.695313341131168
$K = 7$	T	5.420201114052503e−03	−3.035904034302803
$K = 9$	T	5.398045731727308e−03	−5.802902286262906
c. $\Delta t = 180$ s			
$K = 1$	T	5.008292307692841e−03	4.883128361592444e−03
$K = 3$	T	4.982358104459905e−03	4.853684211797993e−03
$K = 5$	T	4.954189267802889e−03	4.817898143982369e−03
$K = 7$	T	4.923251697654507e−03	4.773982865402045e−03
$K = 9$	T	4.888740840979011e−03	4.719452994456183e−03
d. $\Delta t = 18$ s			
$K = 1$	T	4.905609505506447e−03	4.758155759186547e−03
$K = 3$	T	4.888958669361898e−03	4.728837894564270e−03
$K = 5$	T	4.885584829704470e−03	4.693431941774545e−03
$K = 7$	T	4.900263828822915e−03	4.650314763278436e−03
$K = 9$	T	4.939858925560259e−03	4.597261835531552e−03

* T stands for temperature.

solution will contain large nonphysical noise and even blow up when the integration time is long enough.

4. Numerical results

a. A simple scale model

The first numerical example is a simple scale model (9) but with a constant forcing and its numerical scheme is given by

$$x_{n+1} = \frac{1}{1 + \Delta t K(x_n)} x_n + 0.001, \tag{22}$$

which is almost identical to (11) but with a constant forcing. The TLM is given by (12). Here the definition of eddy viscosity coefficient $K(x)$ is based on Mellor–Yamada level-2 closure scheme (Mellor–Yamada 1982) and given in details in the following paragraphs.

Reminding of x in (7) represents the velocity shear. The Richardson number is defined by

$$\mathrm{Ri}(x) = \frac{c}{x^2}, \tag{23}$$

where $c > 0$ is a constant that represents the stratification of the fluid.

The flux Richardson number is then defined by

$$\mathrm{Ri}_f(x) = 0.7231[\mathrm{Ri}(x) + 0.1863 - \sqrt{\mathrm{Ri}(x)^2 - 0.3178\mathrm{Ri}(x) + 0.0347}\,].$$

Then we define

$$S_H(x) = \frac{0.5435 - 2.547\mathrm{Ri}_f(x)}{1 - \mathrm{Ri}_f(x)}, \tag{24}$$

$$S_M(x) = \frac{2.028 - 7.527\mathrm{Ri}_f(x)}{2.7176 - 10.8862\mathrm{Ri}_f(x)} S_H(x). \tag{25}$$

The eddy viscosity coefficient $K(x)$ is then defined by

$$K(x) = \begin{cases} lqS_M(x), & \text{if } lqS_M(x) > 10^{-4} \\ 10^{-4}, & \text{if } lqS_M(x) \le 10^{-4}, \end{cases} \tag{26}$$

where the turbulent length scale l and the turbulent kinetic energy (TKE) q are assumed constants. This function is as the same as in the Mellor–Yamada scheme except that c, q, and l are not fixed with time in the Mellor–Yamada scheme.

The example is treated dimensionlessly. The constants used are $c = 1.41 \times 10^{-4}$, $lq = 4.0$. Figure 2 (a), (b), and (c) show the function $K(x)$, its derivative $K'(x)$ and $xK'(x)$, respectively. There is a discontinuous point near $x = 0.03$ for $K'(x)$. In the neighborhood of $x = 0.03$, $xK'(x) > K(x)$. According to analysis in section 3a, the linearization of (22) is only conditionally stable.

When the time step size is 2.0 for both the nonlinear model and the TLM, the nonlinear model is run twice with different initial conditions $x_0 = 0.04$ and 0.05. The difference of the two solutions is shown in Fig. 3(a) along with the TLM solution, which has the initial solution $\delta x_0 = 0.01$ and the basic states are updated at every time step using the nonlinear solution with $x_0 =$

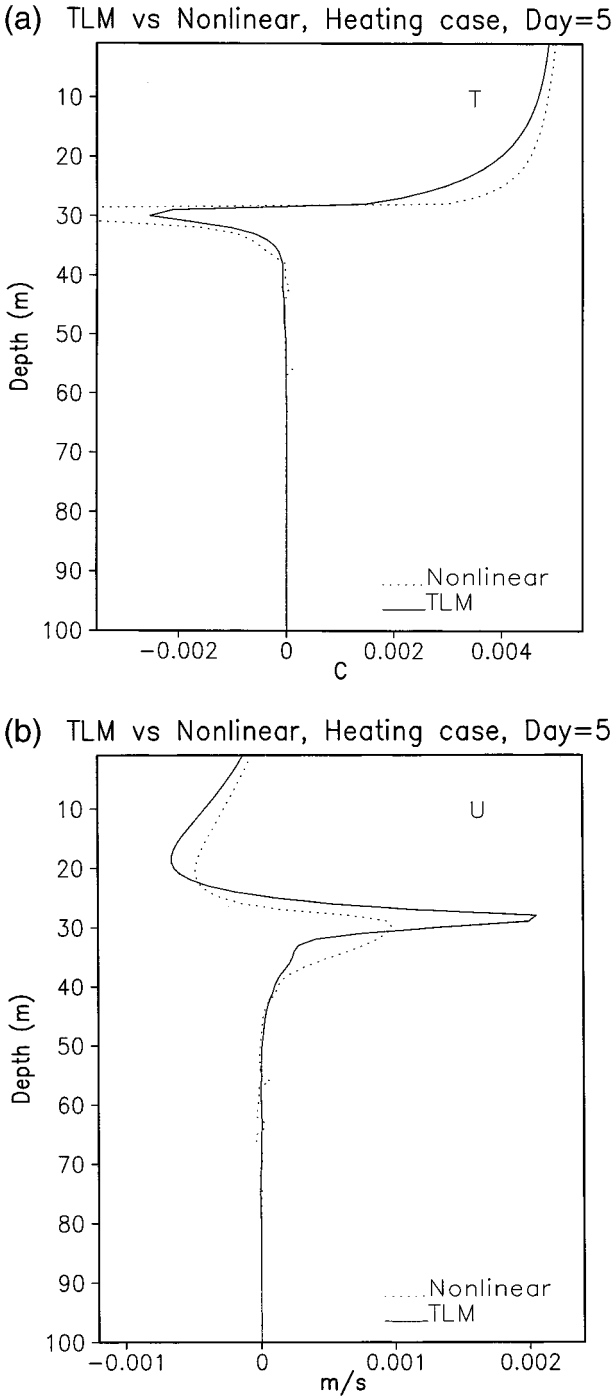

FIG. 4. Vertical profiles of nonlinear perturbation and TLM solution for T and U in heating case. Both nonlinear model and TLM used 180-s time step.

0.04. The time integration of the TLM is not stable, and its solution diverges from the nonlinear perturbation and is becoming unbounded as integration time is getting large.

When the time step size is 1.5 for both the nonlinear model and the TLM, the same nonlinear and TLM runs are made. The time integration of the TLM is still unstable (see Fig. 3b), but provides a better approximation to the nonlinear perturbations than one with time step 2.0 in the time interval (0, 60).

When the time step is further reduced to 0.6 for both the nonlinear model and the TLM, the time integration of the TLM becomes stable and provides much better approximation than ones using larger time steps (see Fig. 3c).

Since the discontinuity of K exists in all above computation, the discontinuity is not responsible for instability. This result verified theoretical discussions in section 3a and a more realistic example is given blow.

b. An ocean mixed layer model

The ocean mixed layer (OML) model is a Mellor–Yamada level 2 closure model. The model equations are

$$\frac{\partial U}{\partial t} = fV + \frac{\partial}{\partial z}\left(K_M \frac{\partial U}{\partial z}\right), \tag{27}$$

$$\frac{\partial V}{\partial t} = -fU + \frac{\partial}{\partial z}\left(K_M \frac{\partial V}{\partial z}\right), \tag{28}$$

$$\frac{\partial T}{\partial t} = \frac{1}{\rho_0 c_p}\frac{\partial I}{\partial z} + \frac{\partial}{\partial z}\left(K_H \frac{\partial T}{\partial z}\right), \quad \text{and} \tag{29}$$

$$\frac{\partial S}{\partial t} = \frac{\partial}{\partial z}\left(K_H \frac{\partial S}{\partial z}\right), \tag{30}$$

where U is the mean eastward velocity component, V the mean northward velocity component, T the mean potential temperature, S the mean salinity, and the first term in the rhs of (29) represents a nonturbulent source flux due to penetrating solar radiation. The diffusion coefficient K_H and eddy viscosity K_M are determined by the flux Richardson number, turbulent kinetic energy (TKE) and the turbulence length scale. The detailed description can be found in Mellor and Yamada (1982). Its performance and comparisons with other OML models can be found in Martin (1985). The time integration schemes for diffusion equations are all semi-implicit. The bottom boundary conditions, temperature, and salinity are nonflux. The slippery bottom boundary conditions are used for velocity components. The method to solve the resulting linear equations is the same as in Princeton Ocean Model (POM) that is described in Mellor (1998) or Richtmyer and Morton (1967). The resolution of the one-dimensional OML model is 1-m. The model has total of 100 vertical levels. The time step size does not have any effects on stability. The choice of time step size only concerns the compromise between accuracy and computational cost. We found that the time step of 1800 s can produce quite reasonable results.

TABLE 3. The comparison of nonlinear perturbation and TLM for wind-deepening experiments at the end of 1-day integration with different depth-level *K*.

Level	Var.	Nonlinear (m s^{-1})	TLM (m s^{-1})
a. Δt = 450 s			
K = 1	U*	1.604474575838025e−03	−1 569 668 875 563.018
K = 3	U	1.534667830172601e−03	−1 415 030 321 614.153
K = 5	U	1.503571445720497e−03	−1 018 106 162 412.618
K = 7	U	1.487390796663574e−03	−316 988 594 054.0956
K = 9	U	1.469436443114050e−03	774 935 857 794.2102
b. Δt = 45 s			
K = 1	U	1.415072087206010e−03	63.80954361758738
K = 3	U	1.364284830611587e−03	63.10950945101569
K = 5	U	1.368634458204038e−03	61.35205278050782
K = 7	U	1.396367654282860e−03	58.34818938208919
K = 9	U	1.430591006803025e−03	53.85840574480086
c. Δt = 15 s			
K = 1	U	1.328471445935370e−03	4.315269605258496e−03
K = 3	U	1.230045798674861e−03	4.061538066533440e−03
K = 5	U	1.145714073422385e−03	3.795266507935753e−03
K = 7	U	1.068762450247104e−03	3.518272624496322e−03
K = 9	U	9.960802672110275e−04	3.233371245996487e−03
d. Δt = 10 s			
K = 1	U	1.325338041016991e−03	1.825246884496160e−03
K = 3	U	1.215019794610278e−03	1.583956842814689e−03
K = 5	U	1.112012819350292e−03	1.348294991818134e−03
K = 7	U	1.012401512808792e−03	1.121479379149642e−03
K = 9	U	9.145618967979219e−04	9.071581250640328e−04

* U stands for eastward component of the velocity.

1) CASE I: HEATING

First, we describe a group of heating experiments in which a constant surface heating flux from the atmosphere is applied. In each experiment two nonlinear model runs and one TLM run are made. Two nonlinear model runs generate a basic state trajectory and a nonlinear perturbation of the basic state trajectory. Experiments differ from each other only in the time step size.

The nonlinear model runs were initialized with a uniform temperature (19°C) and salinity (35 psu) from the surface down to the bottom (100 m deep). The initial field is at rest. The basic state model run was forced by a heat flux of 150 ly day^{-1} (1 ly day^{-1} = 0.484 W m^{-2}). The mixed layer depth was 30 m after a 5-day integration. The perturbation nonlinear model run was forced by a heat flux of 151 ly day^{-1}. The TLM run was forced by the heat flux perturbation of 1 ly day^{-1}. The integration time was 5 days for each nonlinear model and TLM runs. The case represents the stabilizing effect of surface heating. More detailed descriptions of experiments can be found in Table 1.

The results show that the TLM integration is not stable when the time step size is larger, for example, than 900 s. Tables 2a,b,c, and d show near-surface temperature perturbations at the end of 5-day integration for time step size of 18, 180, 900, and 1800 s, respectively. For time step size not larger than 180 s, the TLM produced good approximations to the nonlinear perturbations. Larger time step size (900 or 1800 s) the TLM produced noisy meaningless results. Figure 4 shows the vertical profiles of perturbations of T and U produced by TLM and a nonlinear model at the end of a 5-day integration, respectively, for the time step size of 180 s. The TLM worked well except around 30-m depth where the mixed layer base is located. For even shorter time step size (e.g., 18 s), the relative large discrepancies around the mixed layer base still exist. This suggests that it may not be due to the numerical instability of the TLM. More about this issue is discussed later.

2) CASE II: WIND DEEPENING

The initial conditions of nonlinear model runs for this case are the same as in the heating case except for temperature profile. For this case, the initial stratification was 0.05°C m^{-1}. The difference between the basic state model run and the perturbation model run was the wind stress forcing. Table 1 gives the detailed descriptions.

This case shows the different effects of mixing due to surface heating. The mixing in this case is due to wind-generated turbulence. There is a rapid deepening of the mixed layer during the first day, and a very gradual deepening thereafter. This behavior is characteristic of the shear instability mechanism that the model employs (Martin 1985). The mixed layer depth was 40 m at the end of the first day. It was found that the time step size required by stability of the TLM is much smaller than in the heating case. Table 3 gives comparisons between nonlinear and TLM perturbations at the end of the first day for different time step sizes. Table 3a shows

that for a time step size of 450 s, TLM solutions have huge noise that are of 15 orders of magnitude larger than nonlinear perturbations. Even when the time step size was reduced to 45 s, still huge discrepancies exist (see Table 3b). When the time step size was 15 s, the TLM produced solutions of the same order of the magnitude as the nonlinear perturbations (see Table 3c). The time step size was further reduced to 10 s, the discrepancies were very small (see Table 3d).

For the time step size of 10 s, Fig. 5 shows the vertical profiles of perturbations of T and U produced by TLM and nonlinear models at the end of day one, respectively. The same as for the heating case, the TLM worked well except around 40-m depth where the mixed layer base is located. In the OML model, the diffusion coefficients K_H and eddy viscosity K_M are parameterized differently in the ML and below the ML. Below the ML, TKEs are very small and K_H and K_M are set to the threshold values of 10^{-4} m^2 s^{-1} and 10^{-5} m^2 s^{-1}, respectively. In the ML, their values are much larger than the threshold values. Around the ML base, their values change dramatically and this causes nondifferentiability and strong nonlinearity. Hence, the TLM cannot produce high accuracy around the ML base.

3) Simplified TLM

A simplified TLM scheme by setting $K'_M = 0$ and $K'_H = 0$ was used by Errico et al. (1993) and Mahfouf (1999) in order to obtain a stable TLM for the same step size as in a nonlinear model. Here we tried this scheme and found that the simplified TLM cannot produce good approximations to nonlinear perturbations for our OML model, even though it was stable. Table 4 gives some comparisons of nonlinear perturbation, TLM and the simplified TLM in the wind-deepening case. For eastward components of the near-surface velocity profile, perturbations produced by the simplified TLM are one order of magnitude less than those produced by nonlinear model.

4) Regularization

A regularization method proposed by Janiskova et al. (1999) is to replace the original computation of the Richardson number Ri with a modified form

$$\mathrm{Ri}_{\mathrm{modified}} = \frac{g}{c_p T} \frac{\dfrac{\partial s}{\partial z}}{\left\| \dfrac{\partial \mathbf{v}}{\partial z} \right\|^2 + \dfrac{a}{(\Delta t_{\mathrm{phys}})^2}}, \tag{31}$$

where Δt_{phys} is the physical time step, and a is the tuning parameter of the regularization.

This regularization method was tested in the OML model. Note that (31) is for an atmospheric model, for our OML model, it reads

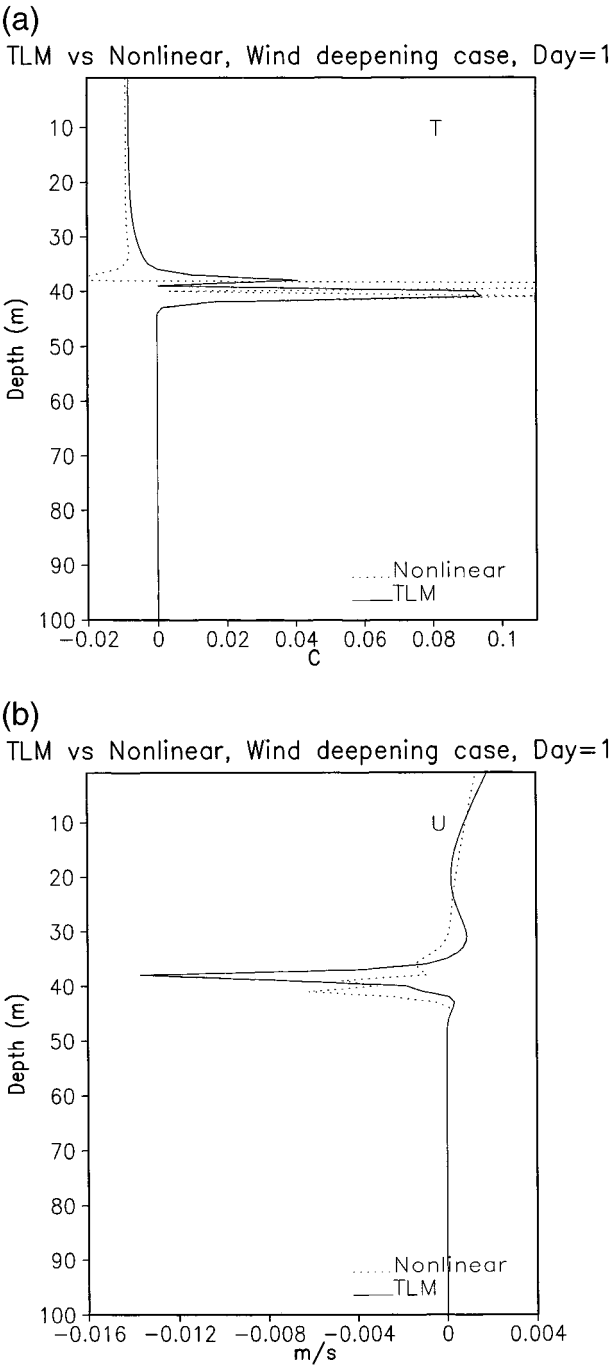

FIG. 5. Vertical profiles of nonlinear perturbation and TLM solution for T and U in wind-deepening case. Both nonlinear model and TLM used 10-s time step.

$$\mathrm{Ri}_{\mathrm{modified}} = -\frac{g}{\rho_0} \frac{\dfrac{\partial \rho}{\partial z}}{\left\| \dfrac{\partial \mathbf{v}}{\partial z} \right\|^2 + \dfrac{a}{(\Delta t_{\mathrm{phys}})^2}}. \tag{32}$$

It was pointed out by Mahfouf (1999) that the main

TABLE 4. The comparison of nonlinear perturbation, TLM, and approximated TLM for wind-deepening experiments at the end of a 1-day integration. The time step size is 900 s.

Level	Var.	Nonlinear (m s^{-1})	TLM (m s^{-1})	Simplified TLM (m s^{-1})
$K = 1$	U	−1.206727650850271e−02	40.93636053248095	2.051219137424033e−03
$K = 3$	U	−1.174131009264068e−02	41.38414096521854	1.818405390515570e−03
$K = 5$	U	−1.101925354277629e−02	42.58641234246276	1.596784276042646e−03
$K = 7$	U	−1.006002585553818e−02	44.92474102786242	1.383887083802127e−03
$K = 9$	U	−8.970310888801958e−03	49.00898655981054	1.177956301076004e−03

disadvantage of this approach lies in the subtle tuning of the smoothing parameters: they need to be small enough in order to provide similar results to the original nonlinear scheme, but they must be large enough to prevent instabilities from growing in the tangent linear scheme. For our OML model, it is very difficult to find such a suitable *a.* For example, in the heating case, we used the time step size of 1800 s ($=\Delta t_{\text{phys}}$) for the nonlinear model and TLM, and tried different values for the smoothing parameter *a.* Table 5 shows some results for perturbation of SST at the end of the 5-day integration. The larger *a* is, the larger the difference of the regularized model from the original model and a better agreement between nonlinear perturbation and TLM. When *a* is 0.130, the TLM is still not stable and produces meaningless results. However, the regularized OML model already differs greatly from the original model. The depth of ML is 47 m, comparing with the 30 m from the original model. When *a* is 0.162, the TLM agrees well with nonlinear perturbation, but the regularized model produces meaningless results, such as a ML of 100 m deep.

5. Concluding remarks

The major findings of this study are as follows.

1) A stable time-integration scheme for a nonlinear model does not necessarily guarantee stable time integration of the associated tangent linear model with the same time step size. These problems usually occur when the nonlinear models describe vertical diffusion processes.
2) Though the underlying PDEs for TLMs could be stiffer than the nonlinear ones, the instabilities of TLMs are directly related to their numerical implementations. For nonlinear models with semi-implicit time-integration schemes that are unconditionally stable, the direct linearization procedure performed on such numerical schemes of nonlinear models, known as tangent linear models, can be interpreted as some conditionally stable numerical schemes of the underlying linearized equations.
3) Simply reducing the time step size can eliminate the instabilities of TLMs, though computational cost could be much higher. To reduce computational cost, an alternative but equivalent way is reducing spatial resolution of TLMs that will allow a larger time step size for TLMs.
4) The time step sizes for a stable TLM integration depend on the basic state. For a faster transit episode of the basic state trajectory, a stable integration of TLM requires a shorter time step.

There are two ways to determine a suitable time step size for stable integration of a TLM. One is trying different time step sizes for the TLM, which is used in this study. Another way is via stability conditions [e.g., (17)]. However, using stable conditions requires evaluation of derivatives of some function and is not an easy task. For example, if we want to apply this way to determine a stable time step size for the OML model, evaluations of the partial derivatives K_M and K_H with respect to U, V, T, and S are required for the range of all practically possible values of U, V, T, and S. In addition, since K_M and K_H are related to U, V, T, and S in complicated ways that are not analytically tractable, evaluations of the partial derivatives K_M and K_H require the adjoint codes.

The instability for TLMs could be a potential problem for other nonlinear numerical models. Since the stability of TLM integration has crucial importance in developing an accurate TLM, which can describe nonlinear perturbations of the size of actual uncertainties in data analysis or model forecast, it should be studied carefully. The other factors that influence the accuracy of a TLM include discontinuities and high nonlinearities, which have been paid a lot of attentions previously. Reducing time step size definitely increases computational cost. Another disadvantage is that the reducing time step size cannot remove any discontinuous or nonsmooth terms.

TABLE 5. Perturbation of SST in heating case with respect to different smoothing parameters.

a	Nonlinear (°C)	TLM (°C)	Depth of ML (m)
9.72e−2	2.520473948519708e−02	156.7381160546732	43
1.30e−1	5.393226489033509e−03	243.8217279539280	47
1.62e−1	1.332244386471615e−03	1.292825026916929e−03	100
3.24e−1	1.117906229804788e−03	1.112440211804857e−03	100

Smoothing, simplification, or regularization may reduce the possibility of instabilities in a TLM integration, while improving the accuracy of a TLM. However, they also create discrepancies between the modified nonlinear model and its original model. There is another way to deal with the numerical instability of TLMs. That is to use entirely different time stepping algorithm in the linear models. It is more suitable for the linearized equations. This approach should allow longer time steps in the adjoint computations and the associated adjoint solutions should provide good approximations to the gradients of the cost functions. The combination of reducing time step size and other methods may provide an effective approach to achieve the ultimate goal of developing a TLM.

Acknowledgments. A part of this work has been supported by CREST (Core Research for Evolution Science and Technology) of JST (Japan Science and Technology Corporation). We thank anonymous reviewers for their useful comments.

REFERENCES

Courtier, P., J.-N. Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.,* **120,** 1367–1388.

Dautray, R., and J.-L. Lions, 1988: *Mathematical Analysis and Numerical Methods for Science and Technology,* Vol. 2: *Functional and Variational Methods,* Springer-Verlag, 561 pp.

Ehrendorfer, M., and R. M. Errico, 1995: Mesoscale predictability and the spectrum of optimal perturbations. *J. Atmos. Sci.,* **52,** 3475–3500.

Errico, R. M., and K. Reader, 1999: An examination of the linearization of a mesoscale model with moist physics. *Quart. J. Roy. Meteor. Soc.,* **125,** 169–195.

——, T. Vukicevic, and K. Reader, 1993: Examination of the accuracy of a tangent linear model. *Tellus,* **45A,** 462–477.

Giering, R., and T. Kaminski, 1998: Recipes for adjoint code construction. *ACM TOMS,* **24,** 437–474.

Haltiner, G. J., and R. T. Willams, 1980: *Numerical Prediction and Dynamic Meteorology,* 2d ed. John Wiley and Sons, 477 pp.

Janiskova, M., J.-N. Thepaut, and J-F. Geleyn, 1999: Simplified and regular physical parameterizations for incremental four-dimensional variational assimilation. *Mon. Wea. Rev.,* **127,** 26–45.

Mahfouf, J.-F., 1999: Influence of physical processes on the tangent-linear approximation. *Tellus,* **51A,** 147–166.

Martin, P. J., 1985: Simulation of the ocean mixed layer at OWS November and Papa with several models. *J. Geophys. Res.,* **90,** 903–916.

Mellor, G. L., 1998: *Users Guide for a Three-Dimensional, Primitive Equation, Numerical Ocean Model,* 1998 version. [Available online at http://www.aos.princeton.edu/WWWPUBLIC/htdocs.pom/.]

——, and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.,* **20,** 851–875.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in Fortran 77.* 2d ed. Vol. 1, Cambridge University Press, 933 pp.

Richtmyer, R. D., and K. W. Morton, 1967: *Difference Methods for Initial-Value Problems.* 2d ed. Interscience, 405 pp.

Roache, P. J., 1972: *Computational Fluid Dynamics.* Hermosa Publishers, 446 pp.

Tsuyuki, T., 1996: Variational data assimilation in the Tropics using precipitation data. Part I: Column model. *Meteor. Atmos. Phys.,* **60,** 87–104.

Verlinde, J., and W. R. Cotton, 1993: Fitting microphysical observations of nonsteady convective clouds to a numerical model: An application of the adjoint technique of data assimilation to a kinematic model. *Mon. Wea. Rev.,* **121,** 2776–2793.

Zupanski, D., and F. Mesinger, 1995: Four-dimensional variational assimilation of precipitation data. *Mon. Wea. Rev.,* **123,** 1112–1127.