On the Merits of Using a 3D-FGAT Assimilation Scheme with an Outer Loop for Atmospheric Situations Governed by Transport

SÉBASTIEN MASSART, BENJAMIN PAJOT, AND ANDREA PIACENTINI

CERFACS/SUC CNRS-URA 1875, Toulouse, France

OLIVIER PANNEKOUCKE

CNRM/GAME, Météo-France/CNRS, Toulouse, France

ABSTRACT

Three-dimensional variational data assimilation (3D-Var) with the first guess at appropriate time (FGAT) appears to be an attractive compromise between accuracy and overall computing time. It is computationally cheaper than four-dimensional (4D)-Var as the increment is not propagated back and forth in time by a model, yet the comparison between the model and the observations is still computed at the right observation time. An interesting feature of the 4D-Var is the iterative process known as the outer loop. This outer-loop approach can also be used in conjunction with 3D-FGAT. But it requires the application of the 3D-FGAT analysis increment at the beginning of the assimilation window.

The pros and cons of using this unusual 3D-FGAT variant are illustrated in this paper on two applications focused on the transport, one of the main phenomena governing the atmospheric evolution. The first one is the one-dimensional advection of a passive tracer. By three representative situations, it shows the benefits of the outer loop, except for practical situations driven by very rapid dynamics such as a zonal wind of 50 m s^{-1} on the earth's great circle, when the assimilation window has a size of 3 h. The second application is the 3D-FGAT assimilation of true ozone measurements into a chemical–transport model. It confirms the previous results, showing that the 3D-FGAT analysis with the outer loop produces an overestimation of the ozone increment in regions where the wind speed is high compared to the time length of the assimilation window.

1. Introduction

In geosciences, data assimilation designates the process in which different kinds of observations distributed in time and space are merged together with a dynamical numerical model (Talagrand 1997). Its aim is to determine as accurately as possible the state of the modeled phenomenon. Data assimilation was initially related to the necessity of defining the initial state of numerical weather prediction (NWP) models dealing with strong nonlinearities. In its beginning, data assimilation was considered to be a sequence of corrections performed at successive observation times (Gandin 1963; Kalman 1960). At every observation time, the corrections were computed according to the available observations and were added to the model solution. This approach has changed with the

variational methods. The idea of applying variational methods in meteorology dates back to the 1960s (Sasaki 1958, 1970). Variational methods and in particular four-dimensional variational data assimilation (4D-Var) provided an alternative approach that globally adjusts the model solution to fit the complete set of the observations available in a given time interval called the assimilation window. Since the early 1980s, this approach has taken advantage of the use of an adjoint equation to compute the gradient of the cost function and then iteratively solve the variational minimization problem. The variational approach was then successfully applied in the mid-1980s to data assimilation problems in the context of nonlinear dynamical models (Talagrand and Courtier 1987; Courtier and Talagrand 1987). The methodology was later rapidly exported both to operational meteorological centers and to other geophysical domains, such as oceanography (Thacker and Long 1988; Bennett 2002) and atmospheric chemistry (Fisher and Lary 1995; Elbern et al. 1997).

Corresponding author address: Sébastien Massart, CERFACS 42, avenue Gaspard Coriolis, 31057 Toulouse, France.
E-mail: massart@cerfacs.fr

In the 1990s, one main objective for operational implementations was to reduce 4D-Var calculations to an affordable level of computational time. For example, Courtier et al. (1994) report that, in 1994 at the European Centre for Medium-Range Weather Forecasts (ECMWF), 4D-Var assimilation on a 24-h window, with 30 minimizer iterations, required as much computational time as the integration of 100 days of the direct model. This period was dedicated to finding low-cost assimilation solutions such as three-dimensional variational data assimilation (3D-Var; Courtier et al. 1998). The 3D-Var with the first guess at appropriate time (3D-FGAT) also appeared and was shown to be superior to 3D-Var (Fisher and Andersson 2001). Contrary to 3D-Var, it propagates in time the background (or prior estimate) state with the complete forecast model in order to compare this state to the observations at the observation time, as in 4D-Var. Another lower-cost solution proposed by Courtier et al. (1994) is an incremental approach based on a 4D-Var in terms of increments. In the 4D-Var incremental approach, the minimization process uses an increment that is propagated with a linear model instead of the complete forecast model. This version requires the linearization of the assimilation operators, which is carried on in the vicinity of the first guess. This first guess is usually the prior estimate (or the background) state. But the solution of the linearized optimization problem can be used as a new first-guess point where to linearize the assimilation operators (instead of the prior estimate state). This gives raise to a new linear problem to solve. This operation can be iteratively repeated: this procedure refers to introducing use of an outer loop. At each iteration of this outer loop, one can use a different simplified linear propagation model. Because a simple linear model may spectacularly reduce the cost of the assimilation, the complexity of the model is usually increased with the number of the iteration.

Simplified linear models applied with the incremental 4D-Var are commonly obtained by omitting some physical process or using a lower spatial resolution (Courtier et al. 1994). The simplest linear model is the identity matrix. The choice of the identity matrix instead of a linearized forecast model to propagate the increment is the basis of the 3D-FGAT, which is of particular interest because the tangent linear and adjoint model integrations explain most of the total computing cost of 4D-Var (Gauthier et al. 2007). This makes 3D-FGAT a hybrid method between 4D-Var and 3D-Var. 3D-FGAT is a useful method since it does not need the adjoint of the linearized propagation model. It is a nonnegligible advantage since, for many applications, the adjoint of the linearized propagation model is not immediately or easily available. 3D-FGAT is in addition more precise than 3D-Var (Lorenc and Rawlins 2005; Laroche et al. 2007). Within 3D-FGAT, the misfit between the model and the observations is computed at the real observation times while 3D-Var uses a single synoptic observation time per assimilation window. Being ordinarily a 4D-Var feature, the outer loop can also be a 3D-FGAT one. In such a case, the 3D-FGAT analysis increment is necessarily added at the beginning of the assimilation window, which results in a 3D-FGAT variant. All the more, the 3D-FGAT analysis increment is normally added at the middle of the assimilation window.

Even if computers are more and more powerful, having a cheaper assimilation method than 4D-Var and a more precise one than 3D-Var is still pertinent. As an example, 3D-FGAT is relevant for providing reanalyses of the chemical atmospheric composition over long periods, which is useful for trends estimation and to detect the signature of the climate change on the composition of the atmosphere. As the transport is one of the main physical processes that control the atmosphere and in particular its chemical composition, this study is dedicated to phenomena whose evolution is governed by the dynamics. It aims to illustrate with two models, the pros and cons in this context, of the 3D-FGAT variant using the outer loop. The first model simulates the one-dimensional advection of a passive tracer. Section 3 of this paper presents this model and its use to demonstrate the particular behavior of the 3D-FGAT with the outer loop in three specific situations: the first one is the assimilation using a nonlinear observation operator, the second one concerns the selection of the assimilated observations, and the last one is a situation that presents high advection velocities. The second model is a comprehensive Chemistry–Transport Model (CTM). Section 4 describes a situation encountered while assimilating real data in this CTM using a 3D-FGAT method with the outer loop. Section 2 provides the theoretical background by introducing the variational data assimilation methods with emphasis on the outer loop. Finally, the conclusions are presented in section 5.

2. Variational data assimilation

Variational assimilation aims at globally adjusting a model solution to all the observations available over a given time interval named assimilation window. Most variational algorithms define a scalar function that, for any model solution over the assimilation window, measures a weighted distance between that solution and the available observations, and between that solution and a prior estimate of the true solution. That so-called objective or cost function is the foundation of 4D-Var.

Variational methods exist in a variety of formulations. We present hereafter the variant form of the 4D-Var cost function in terms of increments, the outer loop, and the FGAT formulations.

a. 4D-Var in terms of increments

4D-Var aims at providing the best estimate $\mathbf{x}^a$ of the unknown true state $\mathbf{x}^t$ at t_0, the initial time of the assimilation window. This estimate, referred to as analysis, is a particular initial condition of the propagation (or forecast) model $\mathcal{M}$. It is computed using the observations $\mathbf{y}_o$ available over the time interval $[t_0, t_P]$, where t_P is the final time of the assimilation window. The model solution $\mathbf{x}(t_i)$ at the observation time t_i is compared to the observations $\mathbf{y}_o(t_i)$ via the observation operator $\mathcal{H}_i$. The 4D-Var cost function is

$$\mathcal{J}[\mathbf{x}(t_0)] = \frac{1}{2}[\mathbf{x}(t_0) - \mathbf{x}_b]^\mathrm{T}\mathbf{B}_0^{-1}[\mathbf{x}(t_0) - \mathbf{x}_b]$$
$$+ \frac{1}{2}\sum_{i=0}^{P}[\mathbf{y}_o(t_i) - \mathcal{H}_i\mathbf{x}(t_i)]^\mathrm{T}\mathbf{R}_i^{-1}[\mathbf{y}_o(t_i) - \mathcal{H}_i\mathbf{x}(t_i)],$$
$$(1)$$

where the background (or prior estimate) field $\mathbf{x}_b$ is an estimate of the true state with an error derived from the realization of an unbiased random variable of covariance $\mathbf{B}_0$. The instrumental and representativeness observation errors are included in the covariance matrices $\mathbf{R}_i$ (Lorenc 1986). The model state $\mathbf{x}(t_i)$ at the observation time t_i is the propagation of the initial state $\mathbf{x}(t_0)$ via the operator $\mathcal{M}_i$:

$$\mathbf{x}(t_i) = \mathcal{M}_i\mathbf{x}(t_0). \quad (2)$$

Courtier et al. (1994) proposed a formulation of the 4D-Var cost function in terms of increments. The increment $\delta\mathbf{x}(t_0)$ is the difference between the initial and the background states, $\delta\mathbf{x}(t_0) = \mathbf{x}(t_0) - \mathbf{x}_b$. Their formulation assumes that the tangent linear operators $\mathbf{M}_i$ and $\mathbf{H}_i$ of $\mathcal{M}_i$ and $\mathcal{H}_i$, respectively, satisfy

$$\mathcal{M}_i[\mathbf{x}(t_0) + \delta\mathbf{x}(t_0)] \approx \mathcal{M}_i\mathbf{x}(t_0) + \mathbf{M}_i\delta\mathbf{x}(t_0), \quad (3)$$

$$\mathcal{H}_i[\mathbf{x}(t_i) + \delta\mathbf{x}(t_i)] \approx \mathcal{H}_i\mathbf{x}(t_i) + \mathbf{H}_i\delta\mathbf{x}(t_i), \quad (4)$$

for acceptable magnitude of the perturbations $\delta\mathbf{x}(t_0)$ and $\delta\mathbf{x}(t_i)$. Under the linearity hypothesis whose validity is discussed in Tremolet (2004), Eqs. (3) and (4) are exact. When the assimilation operators are at most weakly nonlinear during the assimilation window, assuming Eqs. (3) and (4) are suitable approximations, one can express the model equivalent of the observations as a function of the following increment:

$$\mathcal{H}_i\mathcal{M}_i\mathbf{x}(t_0) = \mathcal{H}_i\mathcal{M}_i\mathbf{x}_b + \mathbf{H}_i\mathbf{M}_i\delta\mathbf{x}(t_0). \quad (5)$$

Under this linear approximation, the nonlinear cost function of Eq. (1) is approximated by the linear cost function:

$$\mathcal{J}[\delta\mathbf{x}(t_0)] = \frac{1}{2}\delta\mathbf{x}(t_0)^\mathrm{T}\mathbf{B}_0^{-1}\delta\mathbf{x}(t_0)$$
$$+ \frac{1}{2}\sum_{i=0}^{P}[\mathbf{d}(t_i) - \delta\mathbf{y}(t_i)]^\mathrm{T}\mathbf{R}_i^{-1}[\mathbf{d}(t_i) - \delta\mathbf{y}(t_i)].$$
$$(6)$$

Here $\delta\mathbf{y}(t_i) = \mathbf{H}_i\mathbf{M}_i\delta\mathbf{x}(t_0)$ and $\mathbf{d}(t_i)$ is the innovation vector at observation time t_i, which represents the distance between the observations and the background state propagated at t_i with the nonlinear model, and projected in the observation space with the nonlinear observation operator:

$$\mathbf{d}(t_i) = \mathbf{y}_o(t_i) - \mathcal{H}_i\mathcal{M}_i\mathbf{x}_b. \quad (7)$$

The solution of the assimilation problem is obtained by computing the increment $\delta\mathbf{x}_a(t_0)$ that minimizes the cost function of Eq. (6). The analysis field $\mathbf{x}_a(t_0)$ is the sum of this increment and the background state, $\mathbf{x}_a(t_0) = \mathbf{x}_b + \delta\mathbf{x}_a(t_0)$. In practical applications, the minimum of Eq. (6) is computed by a minimization algorithm that generally requires the gradient of the cost function. The gradient is usually computed using adjoint methods. Nevertheless, the minimum of Eq. (6) can be expressed analytically by looking at the solution that nullifies the gradient. One can demonstrate that the solution is

$$\mathbf{x}_a(t_0) = \mathbf{x}_b + \mathbf{K}^{4D}[\mathbf{y}_o - \mathcal{H}\mathcal{M}\mathbf{x}_b], \quad (8)$$

with

$$\mathbf{K}^{4D} = \mathbf{B}\mathbf{M}^\mathrm{T}\mathbf{H}^\mathrm{T}[\mathbf{H}\mathbf{M}\mathbf{B}\mathbf{M}^\mathrm{T}\mathbf{H}^\mathrm{T} + \mathbf{R}]^{-1}, \quad (9)$$

where $\mathbf{y}_o$ is the generalized observation vector that contains all the observations available over the assimilation interval; $\mathbf{B}$ and $\mathbf{R}$ are the generalized covariance matrices of background and observation errors, respectively. In the same way, $\mathcal{M}$ and $\mathcal{H}$ are the generalization of $\mathcal{M}_i$ and $\mathcal{H}_i$, with their tangent linear form $\mathbf{M}$ and $\mathbf{H}$, respectively. Equations (8) and (9) establish the link between the linear form of 4D-Var in terms of increments and the statistical estimation approach, as it also represents the analysis step of the extended Kalman filter (Maybeck 1979).

b. The outer loop

The 4D-Var in terms of increments requires us to linearize the assimilation operators $\mathcal{M}$ and $\mathcal{H}$. The first

time, the linearization is computed in the vicinity of the background. It leads to the quadratic cost function of Eq. (6). The minimization of this function results in a first analysis. In a second step, this analysis can be used as a new guess in the vicinity of which to linearize the assimilation operators. This leads to a new cost function to minimize, which results in a new analysis. This iterative process is commonly denoted the "outer loop." Its aim is to approach the minimum of the nonlinear 4D-Var cost function with a succession of quadratic cost functions minimizations.

For the outer-loop formulation, the solution $\mathbf{x}^n(t_0)$ of the iteration n is searched as an increment $\delta\mathbf{x}^n(t_0)$ to be applied to the analysis $\mathbf{x}_a^{n-1}(t_0)$ resulting from the previous iteration:

$$\mathbf{x}^n(t_0) = \mathbf{x}_a^{n-1}(t_0) + \delta\mathbf{x}^n(t_0). \qquad (10)$$

The analysis of Eq. (10) is then propagated in time by the nonlinear forecast model in order to be compared to the observations. In terms of increments, the model equivalent of the observations is expressed as

$$\mathcal{H}_i\mathcal{M}_i\mathbf{x}^n(t_0) = \mathcal{H}_i\mathcal{M}_i\mathbf{x}_a^{n-1}(t_0) + \mathbf{H}_i^n\mathbf{M}_i^n\delta\mathbf{x}^n(t_0), \qquad (11)$$

where $\mathbf{H}_i^n$ and $\mathbf{M}_i^n$ denote, respectively, the linear observation operator and the linear model used for the iteration n. The linear model can differ from the tangent linear of $\mathcal{M}$ and can differ from one iteration to another. This is the basis of incremental variational methods. In practical applications, $\mathbf{M}^n$ is the tangent linear of $\mathcal{M}^n$ where $\mathcal{M}^n$ is a version of $\mathcal{M}$, which can have a coarser horizontal resolution and could reflect a less detailed description of the physical laws.

At each iteration n, observations are compared to the best current estimate of the true state. This is another key point of the outer-loop process. Actually, at iteration n, observations are compared to the first guess $\mathbf{x}_a^{n-1}$. This current guess is the best estimate of the true state given the linear estimation of the previous iteration $n-1$; thus, it is updated at each iteration. The innovation vector at time t_i is then the difference between the observations and the model equivalent of the state $\mathbf{x}_a^{n-1}$ given by Eq. (11):

$$\mathbf{d}^n(t_i) = \mathbf{y}_o(t_i) - \mathcal{H}_i\mathcal{M}_i\mathbf{x}_a^{n-1}(t_0). \qquad (12)$$

As previously described and as explained by Courtier et al. (1994), the outer loop was initially used to reduce the cost of the assimilation process. But the outer loop presents other interesting characteristics. It can be seen as a way to iteratively solve a nonlinear problem by a succession of (simpler) linear problems: successive

linearizations of the assimilation operators are recomputed around an updated trajectory that gets closer to the truth at every iteration. The current analysis increment becomes progressively smaller and therefore compliant with the linearity hypothesis. For these reasons, we aim at using the outer loop in conjunction with a cheaper method than 4D-Var.

c. 3D-FGAT with the outer loop

In the incremental approach proposed by Courtier et al. (1994), the linear model $\mathbf{M}$ can differ from the tangent linear of $\mathcal{M}$. The FGAT approach represents an extreme case where the linear model $\mathbf{M}$ is replaced by the identity matrix in Eq. (5), which implies

$$\mathcal{H}_i\mathcal{M}_i(\mathbf{x}_b + \delta\mathbf{x}) = \mathcal{H}_i\mathcal{M}_i\mathbf{x}_b + \mathbf{H}_i\delta\mathbf{x}, \qquad (13)$$

and the increment $\delta\mathbf{x}$ is no longer propagated in time. The 3D-FGAT method loses the dynamical aspect of the 4D-Var and is thus classified as a 3D variational method. The background state is still compared to observations at the observation time (i.e., the appropriated one) by Eq. (7). As the linear propagator is the identity matrix, the 3D-FGAT produces a static increment. According to Fisher and Andersson (2001), the valid time of the increment obtained by the classic 3D-FGAT (without the outer loop) is the central time of the window (not the initial time as in 4D-Var). Applying roughly the 3D-FGAT (or the 3D-VAR) analysis increment at the central time may cause a chock response of the model. To avoid this chock, the increment may be distributed around the central time for example by digital filters (Fillion et al. 1995), incremental digital filters (Gauthier and Thépaut 2001), or the incremental analysis update approach (Bloom et al. 1996).

To perform an outer loop with the 3D-FGAT, the increment has to be applied at the initial time, so to compute the misfit between the new first guess $\mathbf{x}_a^{n-1}(t_0)$ and the observations at the observation time by Eq. (12). Using Eq. (13), we rewrite the cost function of Eq. (6), so the 3D-FGAT cost function at the iteration n is

$$\begin{aligned} \mathcal{J}^{\mathrm{FGAT}}(\delta\mathbf{x}^n) = {} & \frac{1}{2}[\delta\mathbf{x}^n + \delta\mathbf{x}_b^n]^{\mathrm{T}}\mathbf{B}_0^{-1}[\delta\mathbf{x}^n + \delta\mathbf{x}_b^n] \\ & + \frac{1}{2}\sum_{i=0}^{P}[\mathbf{d}^n(t_i) - \mathbf{H}_i^n\delta\mathbf{x}^n]^{\mathrm{T}} \\ & \times \mathbf{R}_i^{-1}[\mathbf{d}^n(t_i) - \mathbf{H}_i^n\delta\mathbf{x}^n], \qquad (14) \end{aligned}$$

where $\delta\mathbf{x}_a^n = 0$ for the first iteration and $\delta\mathbf{x}_a^n = \delta\mathbf{x}_a^{n-1}(t_0) - \mathbf{x}_b$ afterward. The minimization of the cost function of Eq. (14) produces the analysis increment $\delta\mathbf{x}_a^n(t_0)$ and therefore an analysis $\mathbf{x}_a^n(t_0) = \mathbf{x}_a^{n-1}(t_0) + \delta\mathbf{x}_a^n(t_0)$. This 3D-FGAT variant with the correction time placed at the

beginning of the assimilation window will hereafter be simply referred to as 3D-FGAT. This variant makes the aforementioned filtering techniques that distribute the increment around the central time not appropriate as well. Moreover, as the data distribution is obviously more likely to be about the middle time, the expected consequence for this 3D-FGAT variant is that it will be penalized by the observations localized at the end of the window. Indeed, the time inconsistency between the propagated guess and the observations may introduce large errors in the innovation vector. This paper does not detail the differences between the classic 3D-FGAT and the 3D-FGAT with the increment applied at the beginning of the assimilation window. Nevertheless, their results should not be too different under conditions of weak enough impact from flow dependence over the assimilation window. Moreover, both would be deficient compared to 4D-Var under conditions of fast flow.

In the general nonlinear case, it is not possible to have an analytic expression of both 4D-Var and 3D-FGAT analyses as a function of the iteration of the outer loop. But, in the linear case, the 4D-Var solution is given by Eq. (8). The solution is found at the first iteration and does not change with the next iterations. However, this does not hold for 3D-FGAT (except if the model is the identity) as its solution depends on the iteration. The solution at the nth iteration is

$$\mathbf{x}_a^n = \mathbf{x}_b + \mathbf{K}^{\text{FGAT}}[\mathbf{y}_o - \mathbf{H}(\mathbf{x}_b - \mathbf{x}_a^{n-i}) - \mathbf{HMx}_a^{n-1}], \quad (15)$$

where $\mathbf{K}^{\text{FGAT}}$ is a gain matrix that does not contain the linear propagator $\mathbf{M}$:

$$\mathbf{K}^{\text{FGAT}} = \mathbf{BH}^{\text{T}}(\mathbf{HBH}^{\text{T}} + \mathbf{R})^{-1}. \quad (16)$$

Equation (15) shows that the convergence value of the iterative 3D-FGAT is not the 4D-Var solution even in the linear context. The first guess $\mathbf{x}_a^{n-1}$ is propagated by the model $\mathbf{M}$, while the increment $\mathbf{x}_b - \mathbf{x}_a^{n-1}$ is not propagated. It indeed shows there are two differences between 4D-Var (in terms of increments) gain of Eq. (9) and 3D-FGAT gain of Eq. (16). One is that background errors are flow dependent in 4D-Var since they are represented by $\mathbf{MBM}^{\text{T}}$. The second one is that, in 4D-Var, the gain matrix included a backward propagation in time thanks to $\mathbf{M}^{\text{T}}$. Neither of these flow-dependent effects occurs in 3D-FGAT. When the methods are compared, it is difficult to identify which of the two aspects predominates.

In the next section, we propose an example where only the second aspect occurs. The illustration focuses on the outer loop and shows that, even if the iterative 3D-FGAT analyses differ from the 4D-Var analyses, the

use of the outer loop with 3D-FGAT is relevant. Actually, we illustrate that the interesting properties of the outer loop may be advantageous for the 3D-FGAT analysis. But, the outer loop can also cause a deficiency in the 3D-FGAT analysis in situations driven by rapid dynamics compared to the temporal size of the assimilation window.

3. Illustration on a circle

Some practical applications of using the outer loop with the 3D-FGAT are illustrated using a circular one-dimensional domain where the forecast model simulates the advection with a velocity $\mathbf{u}$ of a passive tracer with the concentration $\mathbf{x}(\eta, t)$, along the one-dimensional axis referred to herein as the coordinate η. This forecast model is based on the following equation:

$$\frac{\partial \mathbf{x}}{\partial t} + \mathbf{u}(\eta)\frac{\partial \mathbf{x}}{\partial \eta} = 0. \quad (17)$$

It aims at reproducing one of the major processes governing the atmosphere and its chemical composition.

To evaluate the merits of the outer loop, we set up a twin experiment in which the assimilated observations are synthetic (computed starting from a perturbation of a model state) and the true state is known. Three applications are illustrated in the sections following the descriptions of the numerical experiments.

a. Description of the numerical experiments

The numerical model described by Eq. (17) is used in the particular case of a constant velocity and for a circular domain of radius 6380 km, which corresponds to the earth's great circle. The circle is divided into 445 evenly spaced grid points. In such a situation, Eq. (17) is reduced to a simple translation. The length of the time step is adjusted with the velocity so the field is translated over an integer number of grid points. This limits the numerical diffusion due to the numerical scheme and grants interesting properties of the propagator $\mathbf{M}$ like $\mathbf{M}^{-1} = \mathbf{M}^{\text{T}}$ (see the appendix for further details).

The background $\mathbf{x}_b$ of the concentration of the passive tracer has a sinusoidal shape with a period equal to the half of the domain length, a mean equal to unity, and an amplitude of 0.1 (Fig. 1). A true state is built by adding to $\mathbf{x}_b$ a random Gaussian noise with a zero mean and a covariance matrix equal to the covariance matrix of the background errors. The standard deviation of the background errors is constant in space with its value set to 0.1. The correlations of the background errors are Gaussian with a length scale of 500 km. The knowledge of the true state allows us to build synthetic observations. The
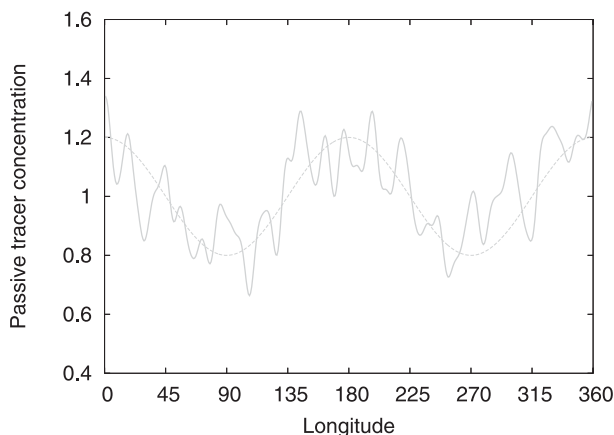
FIG. 1. Concentration of the passive tracer along the equatorial circle at the initial time: true state (solid gray line) and background state (dotted gray line).

observation operator selects the value of the tracer's concentration at the grid point where the observation is available. The observations are built by applying the observation operator to the true state and by adding a Gaussian distributed random noise with zero mean and a standard deviation also set to 0.1. There is no correlation applied between observation errors.

In our particular case of a constant velocity and a homogeneous covariance matrix of the background errors, according to Eq. (A3), the gain of Eq. (9) associated with 4D-Var becomes

$$\mathbf{K}^{4D} = \mathbf{B}\mathbf{M}^T\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}. \tag{18}$$

The relation $\mathbf{B}\mathbf{M}^T = \mathbf{M}^T\mathbf{B}$ can be derived from Eqs. (A2) and (A3) and leads to the equivalent expression of the following gain:

$$\mathbf{K}^{4D} = \mathbf{M}^T\mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}. \tag{19}$$

According to Eq. (16), expressing the gain in the 3D-FGAT case, there is a direct link with this 4D-Var gain under the previously described assumptions:

$$\mathbf{K}^{4D} = \mathbf{M}^T\mathbf{K}^{FGAT}. \tag{20}$$

This relation shows that in our pure advection case with a homogeneous **B** matrix, the analysis increment is the same for the 4D-Var and 3D-FGAT methods, but with a shift due to the adjoint (or inverse) model.

b. Using a nonlinear observation operator

A justification of using the outer loop with a 3D-FGAT assimilation scheme is the potential to relinearize each operator around a better estimate of the truth.

In our case where the forecast model is linear, we illustrate the effect of the outer loop while using a nonlinear observation operator. The illustration simulates the assimilation of the radiance. According to the Beer–Lambert law, the radiation absorbed by the total atmospheric column of a chemical species with a concentration $C(z)$ at the altitude z is proportional to the optical length τ:

$$\tau = \tau_0 \exp\left[-\frac{1}{\sigma}\int_0^\infty C(z)\,dz\right]. \tag{21}$$

Here $\int_0^\infty C(z)\,dz$ is the total column of the chemical species and σ is the absorption cross section of the species. As an example, choosing the ozone as the gas, τ represents the UV absorption.

Since there is no vertical component in our one-dimensional model, we simulate the total column by the state vector **x**. Assuming that $\tau_0 = 1$ and $\sigma = 1$ leads to the nonlinear observation operator:

$$\mathcal{H}_i: \mathbf{x}(t_i) \rightarrow \mathbf{y}(t_i) = e^{-\mathbf{x}(t_i)}. \tag{22}$$

In this experiment, we randomly choose 89 observations evenly spaced (every fifth grid point). The observations are assumed to be all measured at the end of a 3-h assimilation window (Fig. 2a). The observation operator consists here in (i) a projection of the state with the operator of Eq. (22) onto a space we will name observation space for simplicity, and (ii) a selection at the grid point where the observation is available. The velocity is constant with a value of 50 m s^{-1}.

We performed 10 iterations of the outer loop, refreshing the linearization of the observation operator at every iteration. For each iteration, we computed the standard deviation over the whole assimilation window of the current analysis minus the true state. This standard deviation is normalized by the one computed from the difference between the background and the true states. In this example, the standard deviation decreases with the iteration number and becomes almost constant after six iterations (Fig. 2c). The first analysis already provides a good estimation of the truth but with the increase in the number of outer iterations, the structures of the fields are better positioned and have a better amplitude compared to the truth (Fig. 2b). This illustrates the improvement coming from the update of the linearization origin in the outer loop when the assimilation problem is not linear.

c. The observation screening

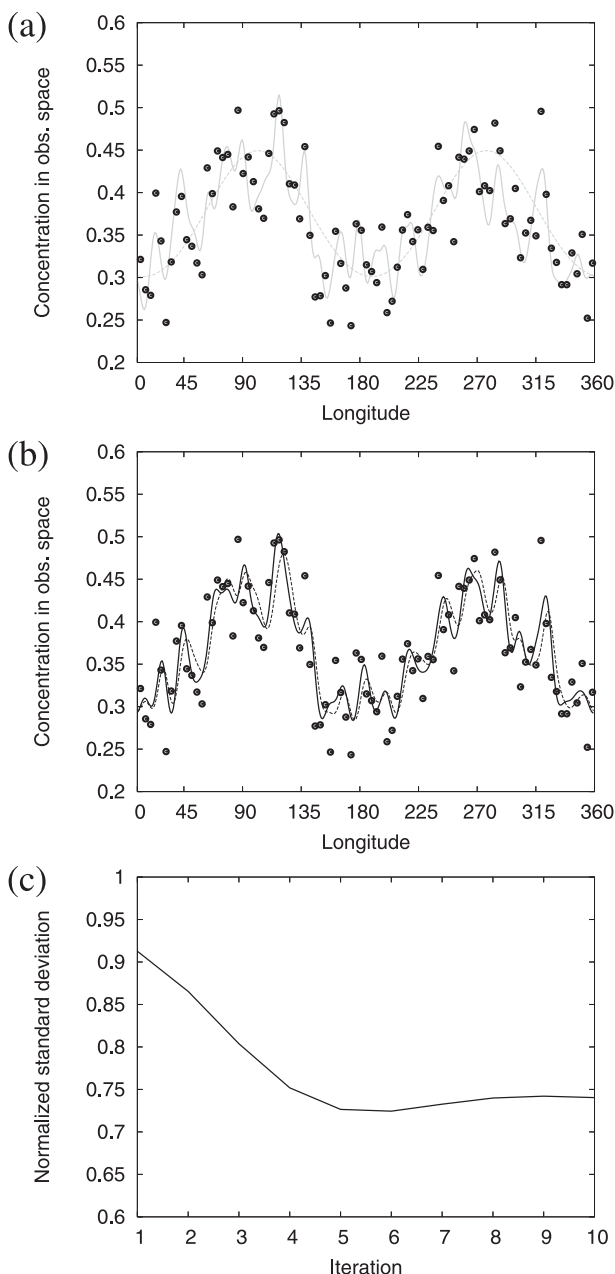As the 3D-FGAT requires the linearization of the assimilation operators, the increment has to be small

FIG. 2. Illustration of the outer loop for the case of a nonlinear observation operator. (a) As in Fig. 1, but in the observation space, at the observation time, and 3 h later after an advection with a velocity of 50 m s^{-1}. The dots represent the observations. (b) As in (a), but for the analyzed field at the first (dotted line) and the last (solid line) iteration of the outer loop. (c) Normalized standard deviation of the difference between the analysis and the truth as a function of the iteration of the outer loop.

enough to stay in the linearization validity region. A useful way to limit the amplitude of the increment is to proceed to a data selection by keeping those that are not too far from the background state (projected in the observation space). This is known as observation screening.

The observation screening is also a mean to detect suspicious measurements and to remove the bad ones from the assimilated dataset.

The observations to be assimilated can be selected first of all when the background is integrated in time with the full propagation model and through the (possibly nonlinear) observation operators. The fact that removed data are far from the background means that some of them are doubtful or that the background is inappropriate. The outer loop produces a new first guess at each iteration. Expecting that the first guess is closer and closer to the truth, it can be used at each iteration to proceed to a new data selection. The removed observations that remain are therefore those that are doubtful, the first guess being no longer inappropriate. The outer loop here offers a chance to inject new observations in the assimilation process while still rejecting the bad ones.

We illustrate this property with the one-dimensional advection model. The variance of the background departure of Eq. (12) can be estimated as a sum of observation and background errors variances, assuming that the observation and the background errors are uncorrelated. In the observation screening, the square of the background departure may be considered as suspect when it exceeds its expected variance. In our example, the observation and background error standard deviations are both equal to 0.1: it makes a departure suspect if its standard deviation exceeds 0.14. To impose a looser criterion, we considered observations with a departure lower than 0.2.

Because of the construction of the twin experiment, most of the synthetic observations satisfy this criteria. We thus decided to keep the previous background state but to compute a new truth by adding to the previous one a random bias with a mean of −0.2 and a standard deviation of 0.1. The 89 computed observations are based on this new truth and are still evenly spaced and measured at the end of a 3-h assimilation window (Fig. 3a). This aims at simulating an atmospheric situation where the forecast model misses something, an important decrease in the concentrations over the whole domain in this particular example. In the proposed illustration, the observation operator exists only in a selection and the velocity is still constant with a value of 50 m s^{-1}.

We chose to compare two experiments: one using the same set of 69 observations selected thanks the background and one with a dynamic screening based on the current first guess of the iteration of the outer loop. As previously, we computed the normalized the standard deviation of the difference between the analysis and the truth. Even if the outer loop reduces this standard deviation when keeping the same number of observations, the reduction is increased when new observations are
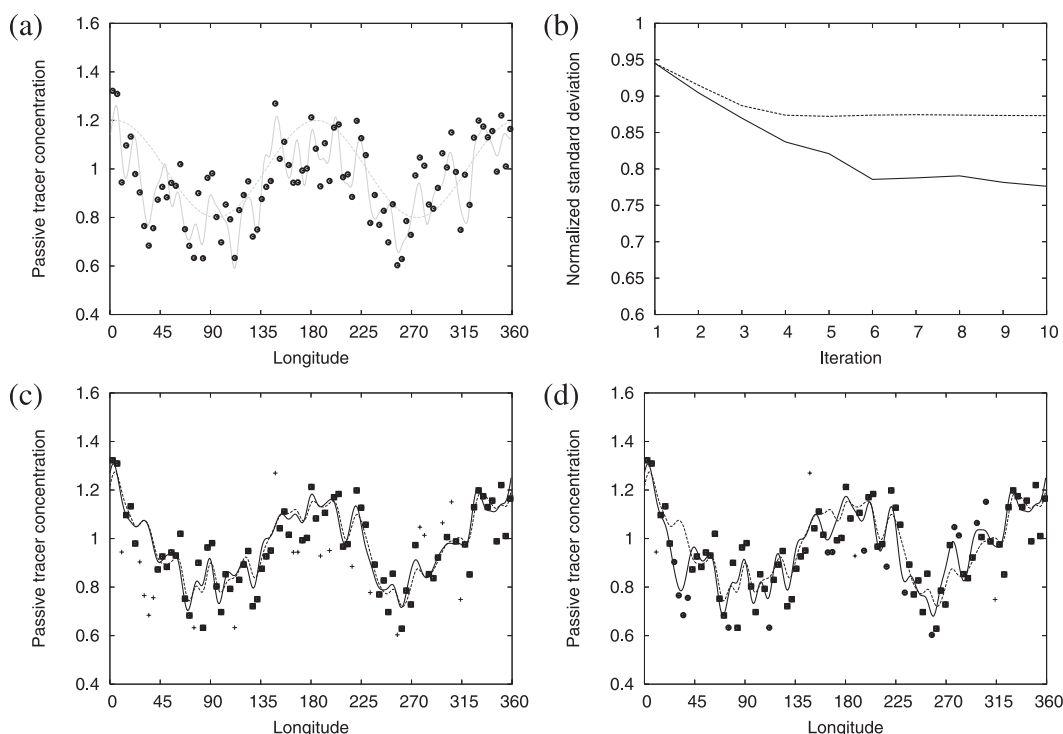
FIG. 3. Illustration of the outer loop for the case of an observation screening. (a) As in Fig. 1, but at the observation time and 3 h later after an advection with a velocity of 50 m s^{-1}. (b) Normalized standard deviation of the difference between the analysis with updated screening (solid line) and with fixed screening (dotted line) and the truth as a function of the iteration of the outer loop. (c) As in (a), but for the analysis with fixed screening at the first (dotted line) and the last (solid line) iteration of the outer loop. (d) As in (c), but for the analysis with updated screening. In (c) and (d), the squares represent the assimilated observations at the first iteration, the dots represent the added assimilated observations at the last iteration and the crosses represent the rejected ones.

added to the assimilation process (Fig. 3b). In this example, the outer loop allows to add progressively 16 more observations over 89 available (Table 1), ignoring as expected the bad ones. For the first iteration of the outer loop, the analysis fields are by construction the same in the cases of a constant screening and an updated one. Most of the rejected observations have a value lower than the background value that is coherent with the negative bias used to build the observations (Figs. 3b,c). But only the case with the updated screening is able to catch these observations after several iterations of the outer loop. This illustrates the improvement coming from the update of the screening in the outer loop when the background state is inappropriate.

d. Situations governed by dynamical effects

In the pure advection case with a constant velocity, the model gets more and more different from the identity with the increase of the velocity. In the presence of high velocities, the 3D-FGAT assumptions reach the limits of their validity. We will show how the use of the outer loop acts in this case. To separate the dynamical effects from

the outer-loop ones, we first begin with an assimilation example without the outer loop. Then the outer loop is added in a second example. Both examples use the previously described experiment framework. They are based on the assimilation of a single observation localized at the end of a 3-h assimilation window. The observation is realized at the longitude 164.2°.

1) ASSIMILATION WITHOUT THE OUTER LOOP

As we only have a single observation, the 3D-FGAT increment distribution has a Gaussian shape due to the correlations of the background errors. The consequence of Eq. (20) is a different location for the maxima of the 3D-FGAT and 4D-Var increments. Within 4D-Var, the adjoint model is used to compute the position of the

TABLE 1. Number of assimilated observations by the 3D-FGAT with the updated screening as a function of the iteration of the outer loop.

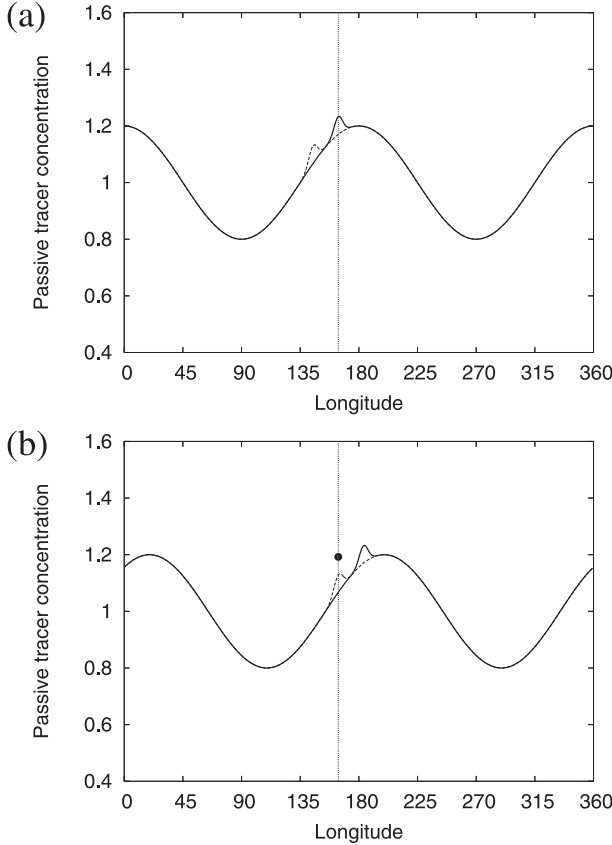| Loop	1	2	3	4	5	6	7	8	9	10
No.	69	81	81	81	81	83	84	85	85	85

FIG. 4. Analyzed concentration of the passive tracer (a) at the initial time and (b) at the observation time, produced by the 3D-FGAT (solid line) and the 4D-Var (dotted line) methods, while assimilating one observation (circle) at the end of a 3-h assimilation window. The velocity is 200 m s^{-1}. The vertical dotted line represents the location of the observation.

correction at the initial time in order for it to be transported at the proper location at the observation time. The correction of the 3D-FGAT is instead added at the observation location at the initial time and is then advected getting beyond the observation location at the observation time. This shows the risk of using the 3D-FGAT method when the dynamics is rapid compared to the size of the assimilation window. It is illustrated in Fig. 4 where the velocity of the forcing field is set to 200 m s^{-1}. This excessive value of the velocity is not relevant for atmospheric winds, but this value was chosen in order to have a strong advection during the assimilation window. In 3 h, a fluid particle at 200 m s^{-1} covers 2160 km, which is more than 4 times the length scale of the correlation of the background errors. This means that the background field and the true state are shifted eastward by about 19.4° in 3 h.

To quantify the 3D-FGAT error, we compared it to the 4D-Var, computing the discrepancy $\mathcal{D}$ between the

analyses produced by the two methods, respectively $\mathbf{x}_a^{\text{FGAT}}$ and $\mathbf{x}_a^{\text{4D}}$, by

$$\mathcal{D} = 100 \times \sqrt{\frac{\|\mathbf{x}_a^{\text{FGAT}} - \mathbf{x}_a^{\text{4D}}\|^2}{\|\mathbf{x}_a^{\text{FGAT}} - \mathbf{x}_b\|^2 + \|\mathbf{x}_a^{\text{4D}} - \mathbf{x}_b\|^2}}, \quad (23)$$

where $\|\cdot\|$ is the L_2 norm defined for a vector $\mathbf{b}$ of dimension n by

$$\|\mathbf{b}\| = \sqrt{\sum_{i=1}^{n} b_i^2}. \quad (24)$$

The discrepancy $\mathcal{D}$ is equal to 0% when the analyses are equal and is equal to 100% when the increments of the analyses are uncorrelated.

We evaluated the discrepancy $\mathcal{D}$ as a function of the velocity. We let the velocity vary from 0.1 to 200 m s^{-1}. By modifying the velocity, we modify the model $\mathbf{M}$. To measure how far is the model from the 3D-FGAT assumption, we built the Frobenius (or Hilbert–Schmidt) norm $\|\cdot\|_F$ of $\mathbf{M} - \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. The Frobenius norm of the $\mathbf{A}$ matrix of dimension n is defined by

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j}^2}, \quad (25)$$

where $a_{i,j}$ is the element of $\mathbf{A}$ on row i and column j. For the different values of the velocity, we computed the discrepancy $\mathcal{D}$ between the 3D-FGAT and the 4D-Var analyses, as well as the norm $\mathcal{N} = \|\mathbf{M} - \mathbf{I}\|_F$. We found that the discrepancy $\mathcal{D}$ is a linear function of the velocity and of the norm $\mathcal{N}$ for low values of the velocity (Fig. 5). At 10 m s^{-1}, a fluid particle covers 108 km in 3 h, which is just more than the size of the grid cell of about 90 km. This means that the concentration in a grid cell weakly influences the concentration in this cell after 3 h. The diagonal of $\mathbf{M}$ is thus close to zero. This explains why the norm $\mathcal{N}$ becomes almost constant when the velocity reaches 20 m s^{-1}. At the same time, the discrepancy $\mathcal{D}$ still grows with the velocity. For velocities of 50 and 200 m s^{-1}, as we use in this paper, the discrepancy $\mathcal{D}$ is respectively 66% and close to 100%.

2) ASSIMILATION WITH THE OUTER LOOP

The problem posed by the 3D-FGAT method when the tangent linear of the propagation model is far from the identity matrix is amplified when an outer loop is used. As illustrated above, the 3D-FGAT increment at the observation time may be situated away from the observation location. This means that $\mathbf{HM}\mathbf{x}_a$ does not
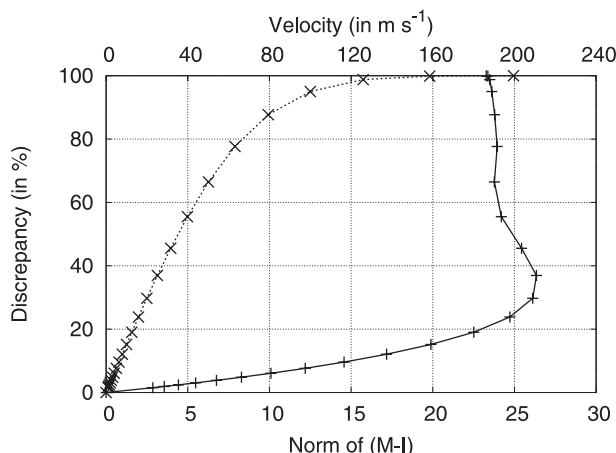
FIG. 5. Discrepancy $\mathcal{D}$ (in %) between the 3D-FGAT and the 4D-Var analyses as a function of the Frobenius norm of $\mathbf{M} - \mathbf{I}$ (solid line) and as a function of the velocity (dotted line). See text for details.

differ much from $\mathbf{HMx}_b$. This case implies that the innovation vector of Eq. (12) for the second iteration of the outer loop is almost the same as in the first iteration. In the observation space, the effect of the first analysis is imperceptible. Assuming that $\mathbf{HMx}_a = \mathbf{HMx}_b$, Eq. (15) becomes

$$\mathbf{x}_a^n = \mathbf{x}_a^1 + \mathbf{K}^{\text{FGAT}}\mathbf{H}(\mathbf{x}_a^{n-1} - \mathbf{x}_b), \qquad (26)$$

where $\mathbf{x}_a^1$ is the analysis obtained at the first iteration. As a consequence, each iteration will produce an increment of the same sign as the increment produced at the previous iteration. This increment is added to the previous one making the analysis differ more from the background (and from the first analysis) when the number of the iteration increases (Fig. 6a). It is possible to determine the convergence of the final increment (difference between the final analysis at the iteration n and the background). In the single observation example, the maximum of the total increment converges toward twice the value of the 4D-Var increment maximum for a velocity of 200 m s^{-1} (Table 2).

When the dynamics is less rapid, the propagation model is closer to the identity matrix but the behavior remains similar. When the discrepancy $\mathcal{D}$ is lower than 100%, the 3D-FGAT increment propagated at the observation time is not null at the observation grid point. This makes the innovation vector change from one outer loop to another. Nevertheless, each iteration of the outer loop adds an artificial supplementary increment to the previous ones, as illustrated by Fig. 6b for a velocity of 50 m s^{-1}. The difference with high velocity values is the fact that the convergence of the total increment maximum is obtained for a lower value (Table 2).
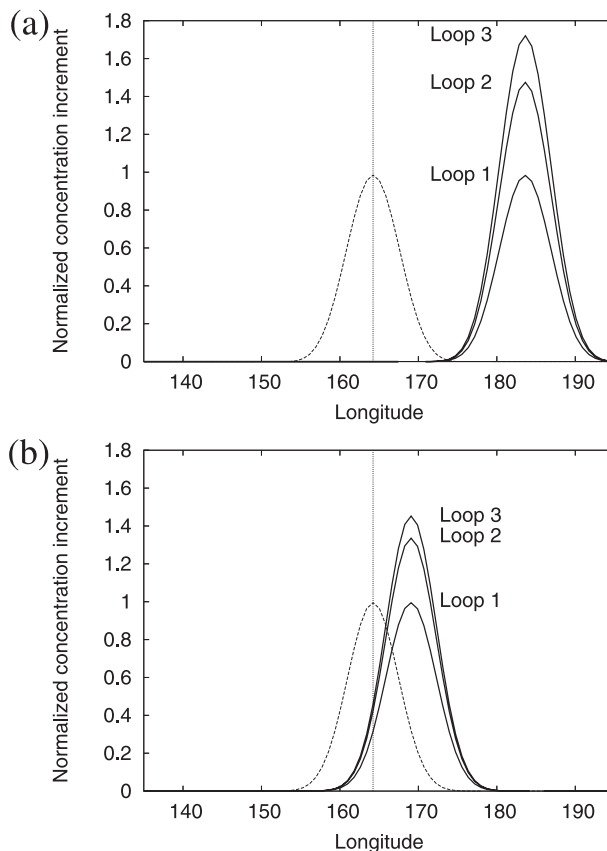

FIG. 6. Local zoom of the increment analysis at the observation time for three outer-loop iterations using the 4D-Var (dashed line) and the 3D-FGAT (solid lines) methods, while assimilating one observation measured at the end of a 3-h assimilation window, with an advection velocity of (a) 200 and (b) 50 m s^{-1}. The values are normalized by the maximum of the 4D-Var increment. The vertical dotted line represents the location of the observation.

4. Illustration on the earth sphere with a CTM

In the previous section, three particular behaviors of the 3D-FGAT using an outer loop were illustrated with a simple one-dimensional advection model. This section describes its behavior in a more complicated model, namely, a three-dimensional CTM. The first section presents the 3D-FGAT assimilation suite used to explore the behavior of the method. The following section describes the experiment conditions. The last section shows and discusses the results of the analysis in a region where the limitations of this 3D-FGAT variant are particularly evident.

a. The assimilation suite

The assimilation system used in this study is based on the European Centre for Research and Advanced Training in Scientific Computation (CERFACS) Valentina

TABLE 2. Value of the maximum of the 3D-FGAT increment for each iteration of the outer loop. The value is normalized by the maximum of the 4D-Var increment of the first outer loop. The assimilation experiments are the cases of one observation and a velocity V of 200 or 50 m s^{-1}.

Iteration	$V = 200$ m s^{-1}	$V = 50$ m s^{-1}
1	1.000	1.000
2	1.502	1.343
3	1.753	1.460
4	1.878	1.501
5	1.940	1.514
6	1.972	1.519
7	1.987	1.521
8	1.995	1.521
9	1.999	1.521
10	2.001	1.521

assimilation framework and the Météo-France comprehensive three-dimensional CTM Mocage. Valentina is an extension of the Mocage–Palm system (Massart et al. 2005) developed jointly by CERFACS and Météo-France in the framework of the FP5 European project Assimilation of *Envisat* Data (ASSET; Lahoz et al. 2007). It is based on the CERFACS Palm software (Buis et al. 2006) that allows it to be coupled with different CTMs. The flexibility offered by Palm makes Valentina a system that offers several assimilation options, like the choice of the variational method (3D-FGAT or 4D-Var), the choice of the representation for covariances of the background errors and the choice of the analyzed domain (regional or global).

The CTM Mocage covers the planetary boundary layer, the free troposphere, and the stratosphere. It provides a number of optional configurations with varying domain geometries and resolutions, as well as chemical and physical parameterization packages. Mocage is currently used for several applications, with recent examples in chemical weather forecasting (Dufour et al. 2004), chemistry–climate interactions (Teyssèdre et al. 2007), and intercontinental transport of ozone and of its precursors (Bousserez et al. 2007).

The first version of Valentina using Mocage, as it was originally implemented for the ASSET project, provided good quality ozone fields as compared with independent measurements with errors of the same order as those produced by several other assimilation systems (Geer et al. 2006). To improve the Valentina assimilation suite, several changes have been recently made in the characterization of the forecast errors (Massart et al. 2007; Pannekoucke and Massart 2008). To increase the consistency of the analysis, several diagnostics were added (Massart et al. 2009) as discussed by Desroziers et al. (2005). Recently, the outer loop has been implemented for the incremental methods.

b. Description of the experiment

In this study, we worked with a 2° by 2° global version of Mocage, with 60 vertical levels (from the surface up to 0.1 hPa). The meteorological forcing fields are provided by the operational ECMWF numerical weather prediction model. To compute the ozone fields, we adopted the linear ozone parameterization developed by Cariolle and Teyssèdre (2007) in its latest version. This parameterization is based on the linearization of ozone production–destruction rates using an altitude–latitude chemical model.

The forecast error covariance matrix of the Valentina assimilation suite is split into a correlation matrix and a diagonal matrix filled with the forecast error variances (square of the forecast error standard deviations). The correlation matrix is divided into a horizontal operator and a vertical operator. These operators are both modeled using a diffusion equation, which is a practical way to generate correlation shapes and in particular Gaussian shapes (Weaver and Courtier 2001). For this study, we generated only Gaussian correlations that can be characterized by their length scale. Horizontal correlations of the forecast errors are computed using a constant length scale of 220 km corresponding to a distance of 2° at the equator. Vertical correlations are computed using a constant length scale of 0.35 (in units of the logarithm of the pressure).

The assimilated data come from the Microwave Limb Sounder (MLS) instrument. This instrument has been flying onboard the *Aura* satellite in a sun-synchronous polar orbit since August 2004. Vertical profiles of several atmospheric parameters are retrieved from the millimeter and submillimeter thermal emission measured at the atmospheric limb (Waters et al. 2006). Measurements are performed between 82°S and 82°N, with a long-track resolution that varies from about 165 km to nearly 300 km. For our study we have used the latest version (v2.2) of the MLS ozone product (see Massart et al. 2009 for more details).

The experiment we carried out was a reanalysis of the global atmospheric ozone performed for the fall of 2007. The assimilation scheme used for this was the 3D-FGAT method with the increment applied at the beginning of 3-h windows and three outer-loop iterations.

c. Assimilation results

We present hereafter only the reanalysis results obtained for 27 August 2007 and for the assimilation window that spans from 0300 to 0600 UTC, when a particular behavior of the assimilation was encountered. We focus on the region 60°–15°S, 100°–10°W at the 10-hPa pressure level (which corresponds to an altitude of approximately 30 km). This region of interest is chosen because there is
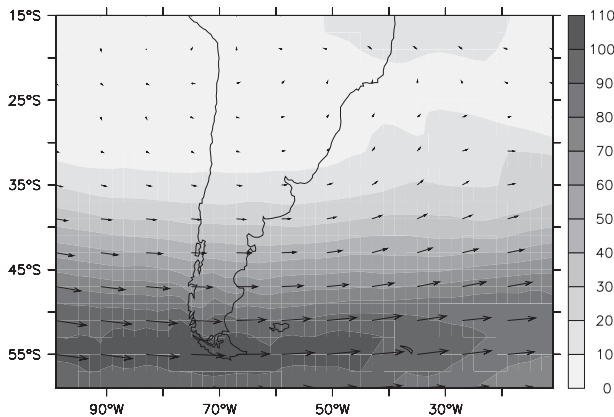
FIG. 7. Regional zoom of the wind amplitude (in m s^{-1}) at 10 hPa for 0300 UTC 27 Aug 2007. The arrows represent the wind direction.

an important gradient in the wind strength for the selected date, with values ranging from 0.6 to 107 m s^{-1} (Fig. 7). Moreover, during the assimilation window, the satellite flies across this region twice, once at the beginning of the window between 0300 and 0307 UTC, and then at the middle of the window between 0435 and 0446 UTC (Fig. 8).

As expected, the 3D-FGAT increments of the first outer-loop iteration are located around the observation location (Fig. 8a). Their spread is due to the correlation of the background errors. As described previously, these increments are added at the initial time of the assimilation window (0300 UTC) to the background state, in order to compute the first analysis. But, with a wind of about 100 m s^{-1} at 60°S, the increment is propagated eastward at approximately 540 km at 1.5 h later, when the second set of data is measured. At 60°S, the distance of 540 km represents a shift of about 10° in longitude. At 45°S, it represents a shift of about 7°. This means that at these latitudes, between 0435 and 0446 UTC, the background and the first analysis are similar at the observation location of the second set of observations. The innovation vector for the second iteration of the outer loop is computed using the first analysis as a first guess, as shown by Eq. (12). As a consequence, for the second set of observations (measured between 0435 and 0446 UTC), in the latitudes between 45° and 60°S, the innovation vector of the second iteration is the same as in the first iteration. This situation is similar to what we saw with the example on the circle when the value of the velocity was 200 m s^{-1} (Fig. 6a). The consequence is that the second iteration of the outer loop computes an increment in the region where the strength of the wind is important of the same order as at the first iteration (Fig. 8b). Where the strength of the wind is lesser, the increments brought by the second outer loop are negligible.
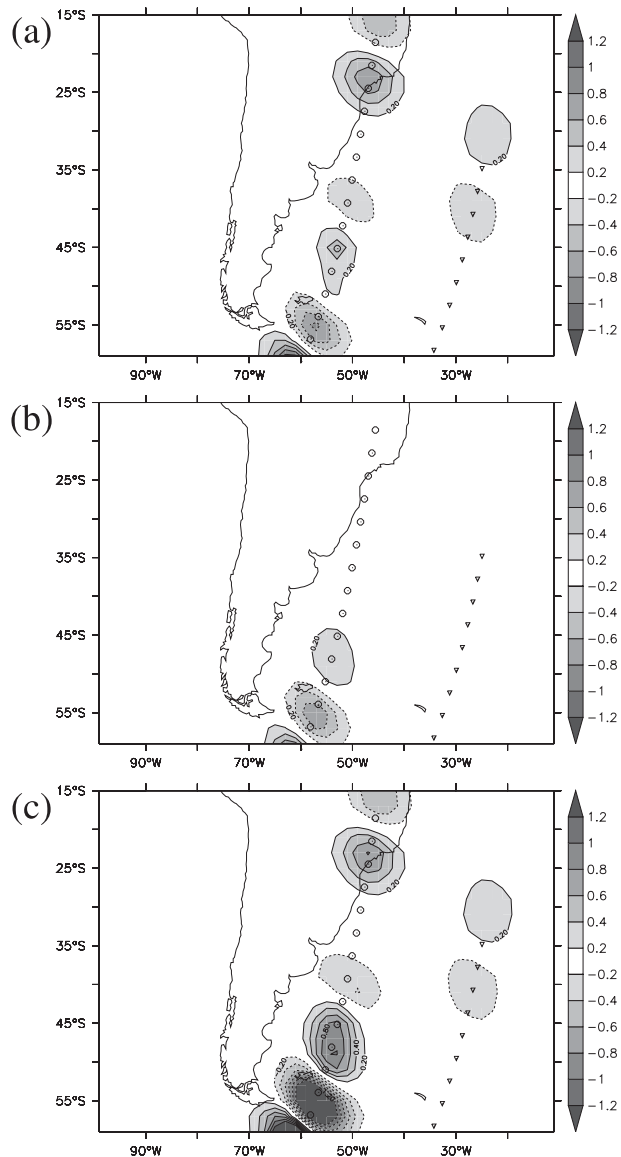

FIG. 8. Regional zoom of the ozone increment (in ppmv) at 10 hPa for 0300 UTC 27 Aug 2007: (a) for the first outer loop, (b) for the second outer loop, and (c) for the final analysis (sum of the increments of the three outer-loop iterations). The triangles represent the MLS data between 0300 and 0307 UTC. The circles represent the MLS data between 0435 and 0446 UTC.

Concerning the first set of observations, as it is measured at the beginning of the assimilation window, the propagation of the first analysis at the time of these observations is close to the identity matrix. The norm $\|\mathbf{M} - \mathbf{I}\|_F$ is low and the 3D-FGAT solution is close to the 4D-Var one. Thus, the analysis computed for the second iteration of the outer loop produces an increment with very small values.

The behavior of the third iteration of the outer loop is similar to the second one. At the latitudes where the

strength of the wind is important compared to the size of the assimilation window, the outer loop amplifies the increment. The consequence is an overestimation of the total correction applied to the ozone concentrations in these regions (Fig. 8c). Moreover, when the analyzed ozone field is compared to the observations (at the observation time), the assimilation process seems not to bring any improvement at these latitudes.

5. Conclusions

3D-FGAT is a useful alternative to 4D-Var as it allows both to avoid the development of the adjoint of the linearized forecast model and to gain computational time. A practical feature of 4D-Var is the outer loop. We investigated the impact of applying the outer loop with 3D-FGAT that does not use the adjoint of the linearized forward model and does not propagate the increment in time. This investigation required the development of a 3D-FGAT variant by imposing the analysis increment to the beginning of the assimilation window instead of at its middle as is usually done. This paper has thus presented the formulation of the 3D-FGAT variant using the outer loop.

To discuss the pros and the cons of this 3D-FGAT variant for atmospheric transport, twin experiments were performed with a forecast model that simulates the advection of a passive tracer on a circular domain. Two experiments showed the benefits of the outer loop. When the assimilation operators are nonlinear the outer loop allows us to update the linearization around a better state and thus increases the quality of the analysis. When a bias exists in the background state, pertinent observations may be rejected by the observation screening. The outer loop here allows us to inject these observations in the assimilated data by updating the observation screening with the first guess refreshed at each iteration. The addition of these observations also increases the quality of the analysis.

The limitations of the 3D-FGAT with the outer loop appear when the limitations of the classical 3D-FGAT appear. The absence of backward propagation in 3D-FGAT is one of the two sources of error when compared to the 4D-Var in terms of increments. In our one-dimensional advection study with a constant velocity and a homogeneous covariance matrix of the background errors, it is the only source of error. We showed through the discrepancy $\mathcal{D}$ that errors grow as a function of the velocity. With another twin experiment, we pointed out that when the velocity is too large in the advection case, as illustrated even with a value of 50 m s^{-1} on the earth's great circle, the 3D-FGAT variant increases the increment at each iteration of the outer loop and finally overestimates the amplitude of the correction. We also found this overestimation in the real ozone assimilation experiment with a CTM. The combination of 3D-FGAT and the outer loop must therefore be carefully used, especially when the dynamics of the studied phenomenon is rapid compared to the size of the assimilation window.

APPENDIX

The Forecast Model and Its Properties

The forecast model used in the one-dimensional study (section 3) simulates the evolution of the concentration of a passive tracer along the one-dimensional axis, based on Eq. (17). The corresponding adjoint equation is

$$\frac{\partial \mathbf{x}^*}{\partial t} - \mathbf{u}(\eta)\frac{\partial \mathbf{x}^*}{\partial \eta} = 0, \quad \text{(A1)}$$

where $\mathbf{x}^*$ is the dual of $\mathbf{x}$. This shows that the adjoint of the transport equation is the transport applied in the backward direction (i.e., with the opposite velocity).

In the particular case of an advection with a constant velocity u on the circular domain, the solution of the forecast model is $\mathbf{x}(\eta, t) = \mathbf{x}(\eta - u.t, 0)$. The model can thus be considered as a simple translation τ_ϕ of phase $\phi = u.t, (\tau_\phi \cdot \mathbf{x})(\eta) = \mathbf{x}(\eta - \phi)$. Applying the Fourier transform, we have for the wavenumber n, $(\tau_\phi \cdot \mathbf{x})_n = e^{-\imath n\omega\phi}x_n$, where x_n is the Fourier coefficient at the wavenumber n and $\omega = 2\pi/L$ with L as the perimeter of the circle. This makes $e^{-\imath n\omega\phi}$ an eigenvalue of the translation operator. By the eigendecomposition of τ_ϕ, it can be shown that τ_ϕ^* has the same eigenvalues as τ_ϕ^{-1}, the complex conjugate of $e^{-\imath n\omega\phi}$ (required to compute τ_ϕ^*) being also its inverse (i.e., $e^{\imath n\omega\phi}$). As they share the same eigenvectors and eigenvalues, $\tau_\phi^* = \tau_\phi^{-1}$. Hence, if **M** denotes the matrix associated to the linear transport equation (i.e., a translation in our case) and if **M*** denotes the matrix associated to its adjoint, then **M*** = **M**$^{-1}$. Being in the Cartesian space with the usual scalar product, the matrix associated to the adjoint equation is the transpose of the matrix associated to the direct equation, **M*** = **M**$^{\mathrm{T}}$, which implies

$$\mathbf{M}^{\mathrm{T}} = \mathbf{M}^{-1}. \quad \text{(A2)}$$

Moreover, if **B** is a homogeneous covariance matrix on the circle, then according to the Wiener Khinchin

theorem, its principal axes are the Fourier modes so that **B** is a diagonal in Fourier space (Mallat 2009). Since the translation is also a diagonal operator in Fourier space, **B** and **M** commute, leading to **MB** = **BM**. By using Eq. (A2), it brings

$$\mathbf{MBM}^{\mathrm{T}} = \mathbf{B}. \qquad (A3)$$

The two properties of Eqs. (A2) and (A3) remain true in the discretized case if the model is expressed as a translation operator. This is the way we choose to model the pure advection equation.

REFERENCES

Bennett, A. F., 2002: *Inverse Modelling of the Ocean and Atmosphere*. Cambridge University Press, 260 pp.

Bloom, S. C., L. L. Takacs, A. M. Da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.,* **124,** 1256–1271.

Bousserez, N., and Coauthors, 2007: Evaluation of the MOCAGE chemistry transport model during the ICARTT/ITOP experiment. *J. Geophys. Res.,* **112,** D10S42, doi:10.1029/2006JD007595.

Buis, S., A. Piacentini, and D. Déclat, 2006: PALM: A computational framework for assembling high performance computing applications. *Concurr. Comput.,* **18,** 247–262.

Cariolle, D., and H. Teyssèdre, 2007: A revised linear ozone photochemistry parameterization for use in transport and general circulation models: Multi-annual simulations. *Atmos. Chem. Phys.,* **7,** 2183–2196.

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

——, J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.,* **120,** 1367–1387.

——, and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart. J. Roy. Meteor. Soc.,* **124,** 1783–1807.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J. Roy. Meteor. Soc.,* **131,** 3385–3396.

Dufour, A., M. Amodei, G. Ancellet, and V.-H. Peuch, 2004: Observed and modelled "chemical weather" during ESCOMPTE. *Atmos. Res.,* **74,** 161–189.

Elbern, H., H. Schmidt, and A. Ebel, 1997: Variational data assimilation for tropospheric chemistry modeling. *J. Geophys. Res.,* **102,** 15 967–15 985.

Fillion, L., H. L. Mitchell, H. R. Ritchie, and A. N. Staniforth, 1995: The impact of a digital filter finalization technique in a global data assimilation system. *Tellus,* **47A,** 304–323.

Fisher, M., and D. J. Lary, 1995: Lagrangian 4-dimensional variational data assimilation of chemical-species. *Quart. J. Roy. Meteor. Soc.,* **121,** 1681–1704.

——, and E. Andersson, 2001: Developments in 4D-Var and Kalman filtering. Tech. Memo. 347, ECMWF, 36 pp.

Gandin, L. S., 1963: *Objective Analysis of Meteorological Fields*. Israel Program for Scientific Translation, 242 pp.

Gauthier, P., and J.-N. Thépaut, 2001: Impact of the digital filter as a weak constraint in the preoperational 4DVAR assimilation system of Météo-France. *Mon. Wea. Rev.,* **129,** 2089–2102.

——, M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau, 2007: Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada. *Mon. Wea. Rev.,* **135,** 2339–2354.

Geer, A., and Coauthors, 2006: The ASSET intercomparison of ozone analyses: Method and first results. *Atmos. Chem. Phys.,* **6,** 5445–5474.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *J. Basic Eng.,* **82,** 35–45.

Lahoz, W. A., and Coauthors, 2007: The assimilation of Envisat data (ASSET) project. *Atmos. Chem. Phys.,* **7,** 1773–1796.

Laroche, S., P. Gauthier, M. Tanguay, S. Pellerin, and J. Morneau, 2007: Impact of the different components of 4DVAR on the global forecast system of the Meteorological Service of Canada. *Mon. Wea. Rev.,* **135,** 2355–2364.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **112,** 1177–1194.

——, and F. Rawlins, 2005: Why does 4D-Var beat 3D-Var? *Quart. J. Roy. Meteor. Soc.,* **131,** 3247–3257.

Mallat, S., 2009: *A Wavelet Tour of Signal Processing: The Sparse Way*. 3rd ed. Vol. 1, Academic Press, 832 pp.

Massart, S., D. Cariolle, and V.-H. Peuch, 2005: Vers une meilleure représentation de la distribution et de la variabilité de l'ozone atmosphérique par l'assimilation des données satellitaires. *C. R. Geosci.,* **337,** 1305–1310.

——, A. Piacentini, D. Cariolle, L. El Amraoui, and N. Semane, 2007: Assessment of the quality of the ozone measurements from the Odin/SMR instrument using data assimilation. *Can. J. Phys.,* **85,** 1209–1223.

——, C. Clerbaux, D. Cariolle, A. Piacentini, S. Turquety, and J. Hadji-Lazaro, 2009: First steps towards the assimilation of IASI ozone data into the MOCAGE-PALM system. *Atmos. Chem. Phys.,* **9,** 5073–5091.

Maybeck, P. S., 1979: *Stochastic Models, Estimation, and Control*. Vol. 1. Academic Press, 423 pp.

Pannekoucke, O., and S. Massart, 2008: Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. *Quart. J. Roy. Meteor. Soc.,* **134,** 1425–1438.

Sasaki, Y. K., 1958: An objective analysis based on the variational method. *J. Meteor. Soc. Japan,* **36,** 77–88.

——, 1970: Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev.,* **98,** 875–883.

Talagrand, O., 1997: Assimilation of observations: An introduction. *J. Meteor. Soc. Japan,* **75** (1B), 191–209.

——, and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.,* **113,** 1311–1328.

Teyssèdre, H., and Coauthors, 2007: A new tropospheric and stratospheric chemistry and transport model MOCAGE-Climat for multi-year studies: Evaluation of the present-day climatology and sensitivity to surface processes. *Atmos. Chem. Phys.,* **7,** 5815–5860.

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

Tremolet, Y., 2004: Diagnostics of linear and incremental approximations in 4D-Var. *Quart. J. Roy. Meteor. Soc.,* **130,** 2233–2251.

Waters, J. W., and Coauthors, 2006: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite. *IEEE Trans. Geosci. Remote Sens.,* **44,** 1075–1092.

Weaver, A., and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc.,* **127,** 1815–1846.