SPECIAL
Mathematical Advances in Data Assimilation
COLLECTION

Chemical Source Inversion Using Assimilated Constituent Observations in an Idealized Two-Dimensional System

ANDREW TANGBORN*

Global Modeling and Assimilation Office, Goddard Space Flight Center, Greenbelt, Maryland

ROBERT COOPER+

Department of Physics, Williams College, Williamstown, Massachusetts

STEVEN PAWSON

Global Modeling and Assimilation Office, Goddard Space Flight Center, Greenbelt, Maryland

ZHIBIN SUN#

Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, Maryland

(Manuscript received 29 August 2008, in final form 24 December 2008)

ABSTRACT

A source inversion technique for chemical constituents is presented that uses assimilated constituent observations rather than directly using the observations. The method is tested with a simple model problem, which is a two-dimensional Fourier–Galerkin transport model combined with a Kalman filter for data assimilation. Inversion is carried out using a Green's function method and observations are simulated from a true state with added Gaussian noise. The forecast state uses the same spectral model but differs by an unbiased Gaussian model error and emissions models with constant errors. The numerical experiments employ both simulated in situ and satellite observation networks. Source inversion was carried out either by directly using synthetically generated observations with added noise or by first assimilating the observations and using the analyses to extract observations. Twenty identical twin experiments were conducted for each set of source and observation configurations, and it was found that in the limiting cases of a very few localized observations or an extremely large observation network there is little advantage to carrying out assimilation first. For intermediate observation densities, the source inversion error standard deviation is decreased by 50% to 90% when the observations are assimilated with the Kalman filter before carrying out the Green's function inversion.

1. Introduction

Understanding the terrestrial carbon cycle is of prime importance to predicting the evolution of climate and ecosystems. It is particularly useful to gain knowledge of the fluxes of carbon species between land and atmosphere and ocean and atmosphere; without this knowledge, an understanding of the physical and biological processes that govern the present-day carbon budget cannot be attained, which in turn means that there is little chance of accurately predicting future climate. There are two predominant approaches to deducing these fluxes or source–sink distributions. One of them, the ''bottom-up'' method, uses models of ocean

* Additional affiliation: Joint Center for Earth Systems Technology, University of Maryland, Baltimore County, Baltimore, Maryland.

+ Current affiliation: Department of Molecular Biology, Princeton University, Princeton, New Jersey.

Current affiliation: Atmospheric and Ocean Sciences Program, Princeton University, Princeton, New Jersey.

Corresponding author address: Andrew Tangborn, Global Modeling and Assimilation Office, Goddard Space Flight Center, Code 610, Greenbelt, MD 20771.
E-mail: andrew.v.tangborn@nasa.gov

DOI: 10.1175/2009MWR2775.1

© 2009 American Meteorological Society

biogeochemistry or land ecosystems along with data constraints (meteorological analyses and relevant biophysical parameters, such as leaf-area index deduced from satellite data). Examples of such bottom-up approaches include Tucker et al. (1986) and Randerson et al. (1997). In contrast, the ''top-down'' approach uses atmospheric concentration measurements in conjunction with transport fields (winds, cloud mass fluxes, and diffusivity) deduced from atmospheric analyses or models. Both approaches are subject to uncertainty associated with model error, analysis uncertainty, and characteristics of the various types of observations. Limitations in the observations include sparse sampling of inhomogeneous quantities and the inherent averaging involved in deducing quantities of physical relevance (e.g., concentrations) from measurements (e.g., radiances).

A number of inverse modeling studies have used surface concentration measurements from a sparse global network to deduce fluxes for a small number (about 12) of continental- or basin-sized regions. For example, Gurney et al. (2005) examine some uncertainties in this method by analyzing differences between deduced fluxes among inverse models that employed different wind fields. Although the continental-scale flux estimates were in reasonable agreement in regions with a few data sources, there was much uncertainty in unconstrained regions, as would be expected. Pétron et al. (2002) used synthesis inversion to estimate time-dependent CO fluxes using ground Climate Monitoring and Diagnostics Laboratory [CMDL; now called National Oceanic and Atmospheric Administration/Earth System Research Laboratory (NOAA/ESRL)] surface station data. A number of other studies (e.g., Rayner and O'Brien 2001) have considered the utility of trace gas constraints derived from space-based instruments, which offer a vastly enhanced data coverage—potentially thousands of soundings per day, compared to tens of observations from in situ instruments.

Inverse methods for estimating chemical sources and sinks generally use either differential (deterministic) or integral (Bayesian) methods. Differential methods use a mass balance to solve for the chemical sources and therefore require constituent observations on a regular grid. Bayesian methods involve the minimization of a cost function and can employ Green's functions (Tarantola 1987; Enting 2000; Pétron et al. 2004; Arellano et al. 2004), adjoint methods (Kaminski et al. 1999; Rayner et al. 2005; Kopacz et al. 2009), or ensemble Kalman filter methods (Peters et al. 2005). Green's functions are defined as the set of observed constituent values that would be expected given a unit source at a single region (or grid point) using a chemical transport model (which includes estimated sources and sinks). The actual observations are then used to invert the resulting system to calculate a new source–sink estimate. In global models, there are generally too many grid points to define a Green's function for each one, so synthesis inversion is used in which the sources are defined in terms of larger emissions regions (or source pattern). The inversion then solves for the magnitude of each source region. Adjoint methods compute the new source estimate using the adjoint of the model (i.e., the transpose of the Jacobian) and apply it to the difference between the observed and modeled tracer values.

Data assimilation and inverse modeling of atmospheric constituents are fundamentally interrelated methodologies, so much so that the terms are often used interchangeably within the chemical inversion community. Both involve the use of transport models and observations of chemical constituent concentrations. They also have in common the use of Bayesian formalism and require an estimate of model and observation error covariances. However, they differ in that data assimilation is generally concerned with obtaining the best possible estimate of the state of the atmosphere (where the state refers to the space–time distribution of the chemical species), whereas chemical inversion is concerned with estimating surface sources and sinks of the species. The question arises as to whether these differences in purpose result in an equivalent extraction of information from the observations. The answer to this will depend in part on exactly which assimilation and inversion techniques are used.

The Kalman filter (Kalman 1960) produces an optimal estimate of the state of a system in the minimum error sense when certain conditions are met. These include assumptions of unbiased forecast and observation errors, Gaussian error statistics, and linear dynamics. Each of these requirements is difficult to achieve in atmospheric data assimilation applications, but they can often be good approximations to real systems. For linear state estimation problems, the Kalman filter gives a minimum variance solution by minimizing a cost function that gives weights to the forecast and observations according to their relative covariances (Cohn 1997). The forecast error covariance, $\mathbf{P}^f$, is evolved by the linearized dynamics and therefore contains current information on error variance and the correlations between different locations. In carrying out assimilation, nonzero correlations are used to spread the corrections to the forecast to grid points near the observations. The resulting analysis error covariance, $\mathbf{P}^a$, then includes the current error variance and correlation lengths for the analysis field. This approach is only valid for linear systems, but the extended Kalman filter (EKF) can be applied to nonlinear systems (Gelb 1974; Jazwinski 1970).

How can this error covariance information be used to improve the estimation of chemical sources? The Kalman filter is generally too computationally expensive for use in global three-dimensional data assimilation systems. There have been, however, some studies that use it on isentropic surfaces in the stratosphere (Ménard et al. 2000; Ménard and Chang 2000; Auger and Tangborn 2004). These studies showed how the error correlation information in $\mathbf{P}^f$ can impact the success of the assimilation. Further investigations have used suboptimal Kalman filters in tropospheric constituent assimilation (Khattatov et al. 2000; Lamarque and Gille 2003). (p. 3)

A direct comparison between inversion for source estimation and data assimilation is difficult because the end product is different. One could, however, devise a way to make a meaningful comparison by adding an extra step to one of the schemes so that both constituent concentration and sources–sinks are estimated. For example, after obtaining a new source–sink estimate using a Bayesian inversion, the model could be rerun to obtain an improved estimate of the constituent concentration state. Alternatively, the analysis concentration field obtained through data assimilation could be used as an input to a source inversion scheme to obtain a new source estimate.

Kalman filtering has previously been used as a technique for inverting for sources and sinks. Hartley and Prinn (1993) defined a vector of source strengths as an extension of the state space; thus, the observation operator is just the linear transport model, and the forecast error variance is then a measure of the uncertainty in the source estimate. This formulation required a perfect model (transport and chemistry) assumption. Gilliland and Abbitt (2001) developed an adaptive iterative Kalman filter for source inversion in which time-integrated emissions are treated as unobserved state variables. In this work they made use of observations that are only available over short time periods and showed how errors in initial concentration estimates can persist during the course of the assimilation.

The value of combining data assimilation and source inversion is most obvious when using a differential inversion method. Assimilation spreads the observation information to nearby grid points, creating the spatial variations needed to calculate spatial derivatives. Law (1999) used spline interpolation to spread the observations, and Dargaville (2000) used a modified interpolation technique to invert CO_2 observations for a variety of regional sources. Neither of these works takes advantage of the covariance propagation or tuning available in current constituent assimilation systems. Furthermore, the mass-balance inversion methods are local, using only nearby grid points, and thus cannot gain any improvement from more distant observations.

This work is motivated by the growth in the quantity of satellite-derived distributions of atmospheric trace gases. Measurement of trace gases in the atmosphere has led to significant increases in efforts to incorporate these measurements into atmospheric transport models with the goal of obtaining improved estimates of their global distribution and of their sources and sinks. State estimation through the combination of observation and model output is generally referred to as data assimilation, whereas source–sink estimation is referred to as inverse modeling.

The present study examines a highly simplified system for top-down, or inverse, modeling. A simple two-dimensional advection model with an analytically specified wind field is used to compute atmospheric tracer concentrations from a specified source–sink distribution. A variety of sampling approaches are then adopted to examine how accurately the original source–sink distribution can be retrieved in the presence of random errors in both observations and source model. An important aspect of the study is the application of data assimilation to produce analyses from the observations; a comparison is made between the source–sink distribution deduced from analyses and direct observations. It is thus a highly idealized observation system simulation experiment (OSSE), which is intended as a prelude to similar experiments using more realistic systems. In section 2 we define the two-dimensional transport model and in section 3 we introduce the Kalman filter for estimating constituent field. This is followed by the Bayesian Green's function inversion procedure for estimating chemical sources in section 4 and the new combined assimilation and inversion scheme in section 5. Section 6 presents the results of the new system followed by the conclusions in section 7.

2. Transport model and observing system

We define the transport model as the solution to the linear two-dimensional convection–diffusion equations

$$\frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} + v\frac{\partial c}{\partial y} = \alpha\left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}\right) + S - Lc, \quad (1)$$

where c is the mixing ratio, (u, v) are the (x, y) components of velocity, α is the diffusivity, S is the rate of production of c, and L is the loss rate frequency of c. We treat this system as nondimensional, so all the variables are unitless. The boundary conditions are periodic in x and y, and the domain is of size $2\pi \times 2\pi$. The numerical model employed is a Fourier–Galerkin scheme π with Crank–Nicolson time-stepping. The numerical solution is then written as

$$\hat{\mathbf{c}}_{k+1} = \mathbf{\Phi}\hat{\mathbf{c}}_k + \hat{\mathbf{S}}, \quad (2)$$

where k is the time step and $\mathbf{\Phi}$ represents the numerical model's system matrix, and the caret indicates that a variable or parameter is in spectral space.

The constituent field is related to its Fourier coefficients by the fast Fourier transform, represented by a matrix operator $\mathbf{F}$, so that

$$\hat{\mathbf{c}}_k = \mathbf{F}\mathbf{c}_k. \tag{3}$$

We define the evolution of the true constituent state as different from the transport model by a random model error, which implies that

$$\hat{\mathbf{c}}^t_{k+1} = \mathbf{\Phi}\hat{\mathbf{c}}^t_k + \hat{\mathbf{S}} + \hat{\mathbf{b}}_k, \tag{4}$$

where $\hat{\mathbf{b}}_k$ is the Fourier coefficient vector of zero mean with Gaussian-distributed random vector $\mathbf{b}_k$. The model error is characterized by its covariance:

$$\mathbf{Q}_k = \langle \mathbf{b}_k \mathbf{b}_k^{\mathrm{T}} \rangle. \tag{5}$$

The diagonal terms of $\mathbf{Q}_k$ are the model variance, $(\sigma^m)^2$ and are constant in time.

The observations ($\mathbf{c}^o$) are taken from the true field, with a spatially uncorrelated random measurement error, $\mathbf{f}_k$. The observations are then

$$\mathbf{c}^o_k = \mathbf{H}_k \mathbf{c}^t_k + \mathbf{f}_k. \tag{6}$$

The observation errors are characterized by the diagonal observation error covariance matrix

$$\mathbf{R}_k = \langle \mathbf{f}_k \mathbf{f}_k^{\mathrm{T}} \rangle, \tag{7}$$

which has an error variance of $(\sigma^o)^2$ along its diagonal and has a characteristic correlation length scale of l_c.

The operator $\mathbf{H}_k$ relates the true constituent field to the actual observation locations. In the next two sections we relate state estimation using Kalman filtering to source–sink estimation using synthesis inversion.

The experiments presented in this paper will make use of synthetic observations obtained from an artificial "nature" run that differs from the model by some difference in the source plus some random errors in the constituent field, $\mathbf{b}_k$. We define this nature run as the "true" state of the system.

The source in the nature run is defined by a constant quadratic function centered at the point (0.47, 0.47) with a peak flux area = 40, as shown in Fig. 1a. The constituent field that results from running the model (starting from a uniformly zero field) for 1000 time steps (unit time of 1.0) is shown in Fig. 2a. In this example the velocity field is $u = 4$, $v = 2$, the diffusivity $\alpha = 0.02$, and the loss coefficient $L = 0.2$.

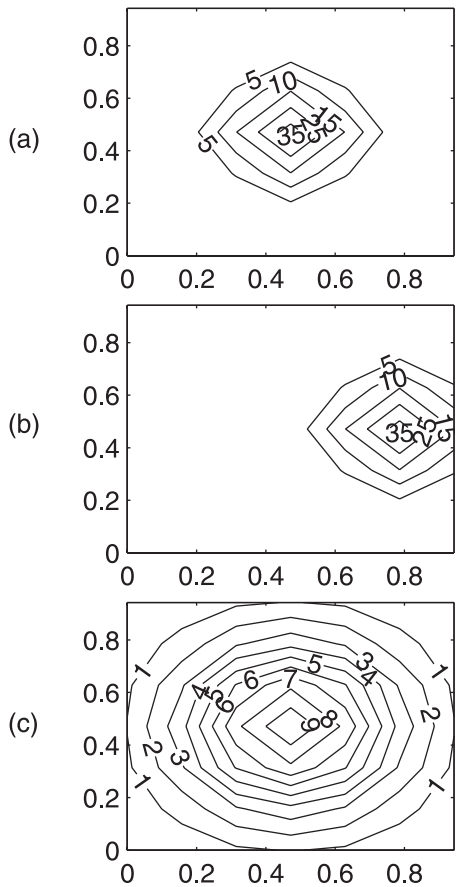

FIG. 1. (a) True source flux, P, which is a quadratic function centered at the position (0.47, 0.47), with a maximum flux area of 40 at its center, and (b),(c) a priori source estimates with (b) *location error* and (c) *spread error.* The source with *location error* is centered about 0.3 units from the true source center and the source with *spread error* has the correct center but twice the diameter as the true source. Note that in each of these plots, only part of the entire domain of $2\pi \times 2\pi$ is shown.

3. The Kalman filter algorithm

The Kalman filter gives the minimum variance solution to the estimation of the state of the system from the model and observations when the errors are unbiased and Gaussian random vectors. It is also assumed that the error variance and correlation lengths for the model, observation, and initial errors are accurately known. Because our system evolves in terms of Fourier coefficients, it is most computationally efficient to evolve the error covariances in the same manner. If the observations are assimilated into the system every m time steps, then the algorithm consists of the following steps.

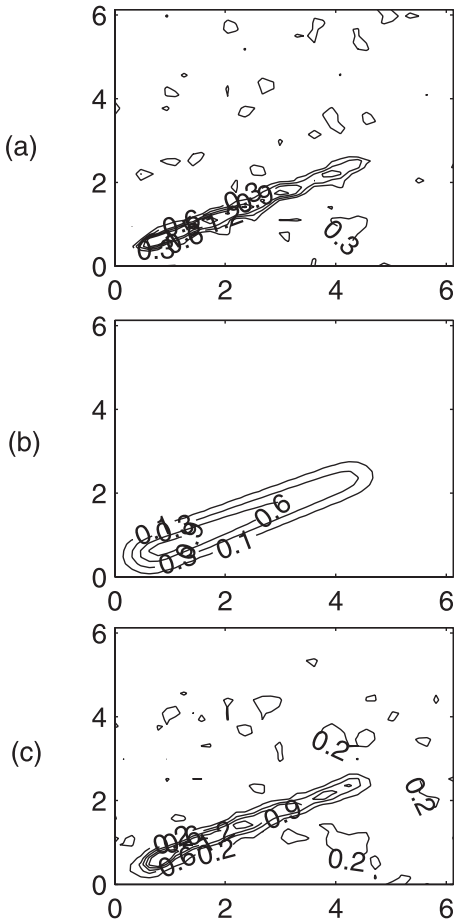

FIG. 2. Concentration field at $t = 1.0$ (after 1000 time steps) that results from (a) the model with the true source, C; (b) the model with *source spread error*; and (c) the assimilation run using the model with spread error, with the satellite observing network. The random part of the field in (a) is due to the time-varying part of the source term P, whereas it is due to the assimilated observations in (c). The center of the source region is represented by a black dot in each panel.

Update of the constituent forecast Fourier coefficients from the previous constituent analysis by m steps

$$\hat{\mathbf{c}}^f_{k+m} = \mathbf{\Phi}^m \hat{\mathbf{c}}^f_k, \tag{8}$$

where $\mathbf{\Phi}^m$ is defined as m applications of the matrix $\mathbf{\Phi}$. The forecast error covariance (in spectral space) is propagated m steps starting from the analysis error covariance by

$$\hat{\mathbf{P}}^f_{k+m} = \mathbf{\Phi}^m \hat{\mathbf{P}}^a_k (\mathbf{\Phi}^m)^{\mathrm{T}} + \hat{\mathbf{Q}}_k, \tag{9}$$

where the covariance matrices have all been transformed to spectral space. The analysis error covariance is determined (in physical space) at assimilation time from

$$\mathbf{P}^a_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\mathbf{P}^f_k. \tag{10}$$

The Kalman gain matrix $\mathbf{K}_k$, which determines the relative weights given to the observations and forecast, is

$$\mathbf{K}_k = \mathbf{P}^f_k \mathbf{H}^{\mathrm{T}} (\mathbf{H}\mathbf{P}^f_k \mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}. \tag{11}$$

Then the new state estimate or analysis update is given by

$$\mathbf{c}^a_k = \mathbf{c}^f_k + \mathbf{K}_k (\mathbf{c}^o_k - \mathbf{H}_k \mathbf{c}^f_k). \tag{12}$$

Cohn (1997) has summarized some of the important properties of the Kalman filter for distributed systems. These include the fact that the error covariances are independent of the observation values but are dependent on the observation locations and errors. This means that as observations are assimilated, information on their impact on the analysis field is included in the analysis error covariances. Because the forecast error covariance is propagated forward starting from the analysis error covariance, it will also contain information on past observation locations and accuracy, insuring that the weighting between forecast and observation takes into account past as well as current information.

4. Synthesis inversion

The terms ''synthesis'' and ''Green's function inversion'' are often used interchangeably, though synthesis inversion is in fact a technique that uses predefined source patterns to reduce the computational cost of the inversion. The technique is based on the Green's function method for solving differential equations through the use of an integral operator. The Green's functions themselves are the resulting set of observations that would be obtained from a unit source at a single point source (or linear combination of sources in the case of synthesis inversion) of unit strength. This is done by running the transport model forward in time from some initial state, for each unit source. Estimates of the sources are obtained by comparing the Green's function with the actual chemical tracer observations and carrying out the inversion.

Synthesis inversion assumes that surface sources of a particular chemical species will eventually be observed somewhere in the atmosphere and its algorithm requires that the lifetime of the species is long compared to the transport times. If chemical reaction adds or removes a substantial fraction of the species during the time during the time of transport, the Green's functions will not accurately represent the distribution of the species that

results from the surface sources. For this reason, synthesis inversion is generally only used for long-lived species such as CO and CO_2.

The standard nomenclature for chemical source inversion differs from that used in data assimilation. In this paper, we will use the usual inversion notation (Enting 2002) but will relate it to the data assimilation notation to help improve clarity. Green's functions are created by using a source of unit strength at each of the $N_x \times N_y$ grid points for each Green's function and running the transport model forward in time. Thus, each Green's function is the solution to the transport model given a single unit source. The set of all Green's functions ($N_x \times N_y$) is then combined to create a Green's function matrix, $\mathbf{G}$ ($N_xN_y \times N_xN_y$). Given an existing estimate of the sources ($\mathbf{z}$) and a set of observations ($\mathbf{c}^o$), error covariance for the observations ($\mathbf{X}^{-1} = \mathbf{R}$), and error covariance for the source model ($\mathbf{W}^{-1}$), the Green's function inversion yields the new source estimate ($\mathbf{S}_{\text{new}}$) as

$$\begin{aligned}\mathbf{S}_{\text{new}} &= (\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{G} + \mathbf{W})^{-1}(\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{c}^o + \mathbf{W}\mathbf{z}) \\ &= (\mathbf{G}^{\mathrm{T}}\mathbf{X}\mathbf{G} + \mathbf{W})^{-1}(\mathbf{G}^{\mathrm{T}}\mathbf{X}\mathbf{c}^o + \mathbf{W}\mathbf{z}),\end{aligned} \tag{13}$$

where $\mathbf{z}$ is the a priori source estimate and $\mathbf{c}$ is the observational dataset. The error covariance of the estimate $\mathbf{S}_{\text{new}}$ is

$$[\mathbf{G}^{\mathrm{T}}\mathbf{X}\mathbf{G} + \mathbf{W}]^{-1}. \tag{14}$$

5. A combined Kalman filter and synthesis inversion algorithm

In this new approach we combine the two schemes in a way that retains optimal characteristics of the Kalman filter with the formalism of Green's function inversion. The Green's function matrix is formed in the same manner, but instead of using observations directly in the inversion, they are assimilated using the Kalman filter, resulting in analyses that give a new estimate of the state of the atmosphere at each observation time. Then the analyses, $\mathbf{c}^a$, are used at every grid point in place of observations, $\mathbf{c}^o$, with error covariance $\mathbf{X}^{-1} = \mathbf{P}^a$. The new scheme for the inversion is then

$$\mathbf{S}_{\text{assim}} = [\mathbf{G}^{\mathrm{T}}(\mathbf{P}^a)^{-1}\mathbf{G} + \mathbf{W}]^{-1}[\mathbf{G}^{\mathrm{T}}(\mathbf{P}^a)^{-1}\mathbf{c}^a + \mathbf{W}\mathbf{z}], \tag{15}$$

where $\mathbf{S}_{\text{assim}}$ is the new source estimate that uses the assimilated observations. Because this inversion uses the analysis $\mathbf{c}^a$, the inverse of the analysis error covariance replaces $\mathbf{X}$ from (14), and the new estimated error covariance is

$$[\mathbf{G}^{\mathrm{T}}(\mathbf{P}^a)^{-1}\mathbf{G} + \mathbf{W}]^{-1}. \tag{16}$$

The advantages to this approach are that the Kalman filter evolves the error covariance using the linear model. This results in both forecast and analysis error covariances that contain correlations that are affected by transport and diffusion. In particular, information from the source region is transported downstream by advection so that forecast errors should be correlated over greater distances. The estimated source error covariances are discussed further in the next section.

6. Numerical experiments

The experiments presented in this paper make use of synthetic observations that are obtained from an artificial nature run that differs from the model by some difference in the source plus some random errors in the constituent field.

The source in the nature run is defined by a quadratic function centered at the point (0.47, 0.47) with a peak flux = 40 (dc/dt/area), as shown in Fig. 1a. The constituent field that results from running the model (starting from a uniformly zero field) for 1000 time steps (unit time of 1.0) is shown in Fig. 2a. In this example the velocity field is $u = 4$, $v = 2$, the diffusivity $\alpha = 0.02$, and the loss coefficient $L = 0.2$.

We have carried out a series of runs to compare the accuracy of the Green's function inversion by directly using the observation networks with the scheme outlined in section 5, which uses the analysis field instead of the observational input to the inversion scheme. We will refer to these inversions as using direct observations and assimilated observations, respectively. Testing of the algorithms and code includes cases with observations at every point and with only two observations, shown in Figs. 3a–d. In the former case, the source inversion using direct observations (Fig. 3a) and assimilated observations (Fig. 3b) produced identical results, which capture the true source to within the observational error. This implies that when the observations are essentially the entire state, then the assimilation adds nothing to the accuracy of the inversion. In the latter case, the two schemes (Figs. 3c,d) were nearly equally unable to improve on the first guess of the source. This test shows that little or no improvement to the inversion can be made when the observations are too sparse (and the system is not observable).

Our interest is in cases that lie between these two extremes, so we have carried out ensemble experiments with a variety of source model and observing networks, including global (satellite) and ground-based (in situ) observations. The observation networks are shown in

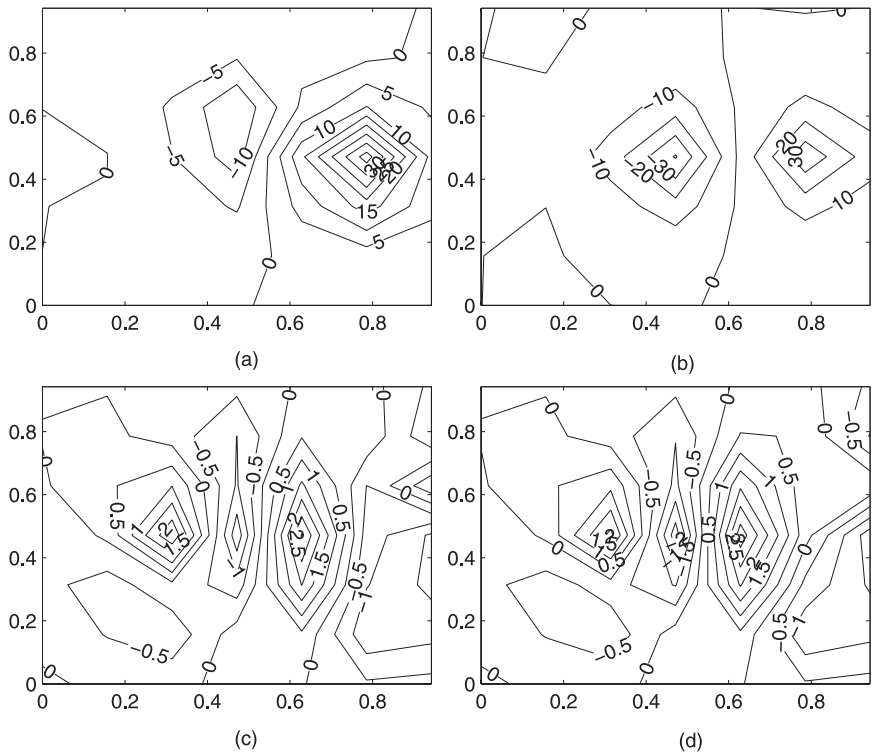

FIG. 3. Mean source errors from the inversion for the extreme cases of (a),(b) two observations and (c),(d) observations at every grid point; (a) and (c) represent direct inversion of the observations whereas (b) and (d) are represent inversion of assimilated observations.

Figs. 4a,b, and all of the observations are available at every assimilation time.

The model uses two possible a priori source estimates, which are shown in Fig. 1. Both of these source estimates are unbiased in the sense that the total flux is exactly the same as that in Fig. 1a, but they have an error either in location (Fig. 1b) or in the localization or spread (Fig. 1c). We refer to these source errors as *source location error* and *source spread error*, respectively. These two models also do not account for the random source–sink term in Eq. (4).

For each source model and each observing network, we have carried out 20 twin experiments using perturbed initial conditions. Twin experiments are essentially simulations that are identical in every aspect except for randomly perturbed initial conditions. This allows us to obtain meaningful statistics of the assimilation and inversion results. In each case the model is run for 1000 time steps, which is roughly the time required for constituents to be transported about 2/3 of the way across the domain. The results are presented by comparing the known true source and constituent field with the model output field and assimilated (analysis) fields as well as the resulting chemical source inversion for each case. We compare the source inversion using the observations directly, and by first assimilating every 20 time steps using the Kalman filter as described in the previous section. In all of the experiments, the parameters used are velocities $u = 4$, $v = 2$, and diffusivity $\alpha = 0.02$; the loss rate coefficient is $L = 0.2$. The observation error standard deviation is $\sigma^o = 0.0014$, the model error standard deviation in Eq. (4) is $\sigma^m = 0.01$, and the model correlation length scale is $l_c = 0.1$.

We present detailed results only for the model with *source spread error* and then summarize all the cases at the end of this section. Labels used in the text for each experiment are defined in Table 1.

a. Concentration field

Figure 2b shows the concentration field that results from running the model with *source spread error* for 1000 time steps without assimilation. As one would expect, the impact of the source is wider than in the true state (Fig. 2a) and lacks the small-scale structure that comes from the random source–sink term in Eq. (4). We plot the RMS error for the concentration field as a function of time for this case, as well as for the assimilation cases using the in situ (SIA) and satellite observations (SSA) in Fig. 6. This figure gives an indication of the relative amounts of information in the observing

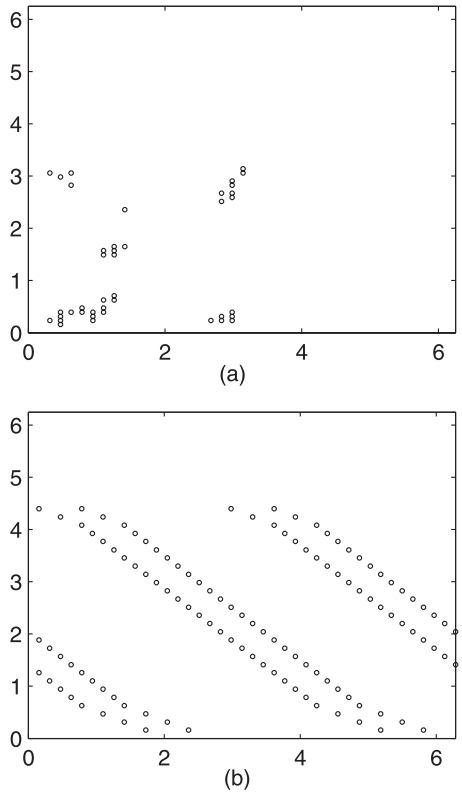

FIG. 4. (a) In situ and (b) satellite observation locations. Observation locations are the same at each analysis time.

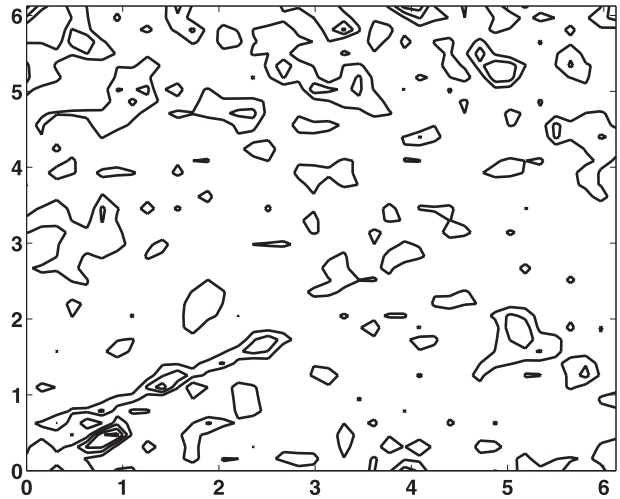

FIG. 5. Contour plot of analysis field minus true field ($\mathbf{c}^a - \mathbf{c}^t$) at the end of 1000 time steps. The contour levels are 0.3. The largest errors are around 0.7 and occur near the source term ($x = 0.47$, $y = 0.47$).

networks, which will be important in the success of the source inversions. With the model alone, the observations have no impact on the constituent field, and the resulting RMS grows continuously as a result of both the local systematic source model error and the random model error. The errors are consistently smaller for the satellite observation network, which has more observations, but fewer in the vicinity of the source. The concentration field obtained from assimilation of satellite observations into the *source spread model* (SSA) is shown in Fig. 2c. The field has narrowed and even contains some of the small-scale features present in the true field. Thus, the assimilation, while not making any correction to the source, changes the downstream structure of the field to more closely resemble the true field. The difference between the assimilation and true final states ($\mathbf{c}^a - \mathbf{c}^t$), shown in Fig. 5, indicates that the analysis field still retains errors on the order of 20%.

b. Source inversion

For each experiment, a source inversion is carried out using the Green's function algorithm, with and without assimilation. The ensemble of twin experiments is used to determine the mean and standard deviation errors relative to the true source. The predicted source inversion error covariances, Eqs. (14) and (16), are valid when the errors are Gaussian and unbiased. We expect that if the model and observation errors are unbiased, then the source inversion should also be unbiased. Figure 7 shows the predicted error variances for the source inversions [Eqs. (14) and (16)] with and without assimilation (SSN and SSA) in a one-dimensional slice through the source region. The predicted errors for the inversion with assimilation are as much as an order of magnitude smaller than the direct inversion errors.

The ensemble mean of the inverted source is defined as

$$\mu_{\text{inv}} = \langle S_{\text{inv}} \rangle, \tag{17}$$

where the ensemble is the 20 twin experiments run for each set of parameters. We define the mean inversion error as

$$\mu^{\epsilon}_{\text{inv}} = \mu - S^t. \tag{18}$$

While all of the errors are globally unbiased, the steady source term has a local bias in the sense that over a long period of time the source at one location can be consistently too large or too small. For example, in the model with source spread error, the flux is consistently too low at the center of the source and is too large near the edge of the source. The total from these sources is the same as the true source total, and the random or short term source–sink term also has zero mean. The mean inversion error $\mu^{\epsilon}_{\text{inv}}$, is therefore an indication of the local and global bias, to the extent that they differ from the true source. In all

TABLE 1. Summary of ensemble results for the assimilation and inversion for the different observation and model types, including the mean peak flux, maximum mean error, peak error standard deviation, and distance of the peak flux from the true location. The true peak flux is 40 (at $x = 0.47$, $y = 0.47$) and the model with location error (at $x = 0.7$, $y = 0.4$) is a distance of 0.33 from the true location. The labels identify which model and observation type is used in each set of experiments. All values are nondimensional and the errors presented are absolute.

			Model with spread error		
			In situ observations		
	Label	Mean peak flux	Max. mean error	Error std dev	Distance from true peak
No assim.	SIN	45	25	150	0
With assim.	SIA	12	30	10	0
			Satellite observations		
	Label	Mean peak flux	Max. mean error	Error std dev	Distance from true peak
No assim.	SSN	50	12	50	0
With assim.	SSA	32	9	24	0
			Model with location error		
			In situ observations		
	Label	Mean peak flux	Max. mean error	Error std dev	Distance from true peak
No assim.	LIN	70	50	170	0.05
With assim.	LIA	38	35	9	0.33
			Satellite observations		
	Label	Mean peak flux	Max. mean error	Error std dev	Distance from true peak
No assim.	LSN	38	20	55	0
With assim.	LSA	35	10	22	0

cases, the ensemble mean inverted sources are zero far from the true source, so we only plot in the vicinity of the source ($0 \le x \le 1$; $0 \le y \le 1$).

Figure 8 shows the mean inversion errors that result from using the source model with spread error. In Fig. 8a, the inversion without assimilation and in situ observations (SIN) is seen to have a mean source that is locally overestimated by as much as 50% ($x = 0.45$, $y = 0.3$) and underestimated by up to 50% ($x = 0.6$, $y = 0.45$). When satellite observations are directed used in the inversion (SSN; Fig. 8c), the maximum mean error is also about 50%.

The inversion using assimilated in situ observations (SIA; Fig. 8b) has a particularly large bias at the center of the source (about 70%), whereas the assimilated satellite observations (SSA; Fig. 8d) are significantly closer to the true source (20% maximum mean error). However, the inversion without assimilation using in situ observations (SSN; Fig. 8a) results in two spurious constituent sinks near the source. This can be seen from the negative mean errors around $y = 0.5$.

The mean inversion errors described above only tell us whether there is any systematic difference between the inverted source and true source. The random component of the error is represented by the error standard deviation

$$\sigma_{\text{inv}} = \langle (\epsilon_S - \mu_\epsilon)^2 \rangle^{1/2} = \langle (S_{\text{invert}} - S_{\text{true}})^2 \rangle^{1/2}, \quad (19)$$

where $\epsilon^S = S_{\text{invert}} - S_{\text{true}}$ and $\mu^\epsilon = \langle S_{\text{invert}} - S_{\text{true}} \rangle$. We calculate the error standard deviation at each grid point and plot the results in Fig. 9 using the same source model

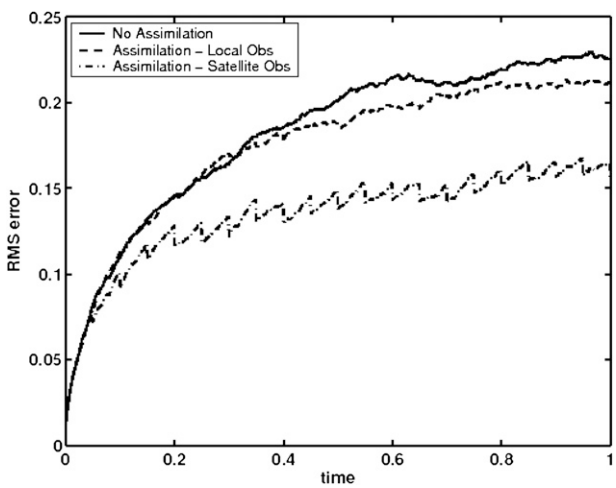

FIG. 6. RMS error in the concentration field relative to the nature run for the model with spread error. The curves shown are for the model only (solid), assimilation of in situ observations (dashed), and assimilation of satellite observations (dashed–dotted).

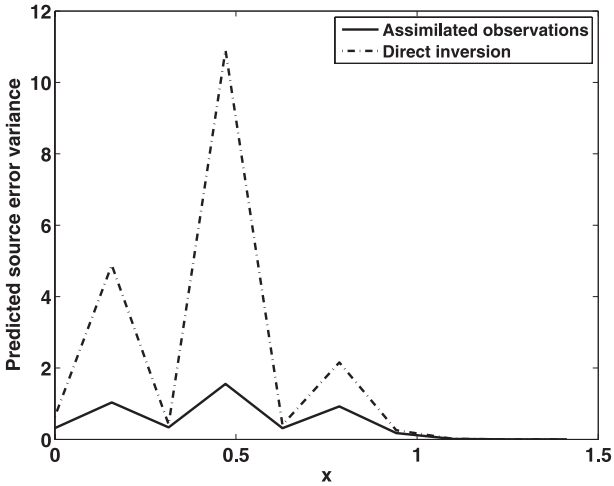

FIG. 7. Predicted error variance for the source inversion for the cases SSA (solid) and SSN (dashed–dotted) along a slice of the source region.

and observations as in Fig. 8. The error standard deviations for inversions without assimilation are consistently larger than those with assimilation, and in some cases the difference can be an order of magnitude. This is important because source inversion is not generally done using ensembles, so that the error standard deviation can be a significant contribution to the inversion error. The difficulty of carrying out ensembles of source inversions when using global models is due to the high computational cost, particularly when many source regions are defined. These results show that the random component is significantly larger than the systematic component for the inversions using the observations directly (Figs. 9a,c). This implies that a single inversion that uses direct observations will have significant uncertainty in the resulting source estimates. Inversions using assimilated observations (Figs. 9b,d) are much smaller than the direct inversion cases.

We summarize the results of the ensembles of assimilation and inversion calculations in Table 1, which lists the value of the peak mean flux, the maximum mean error $\mu^{\epsilon}_{\text{inv}}$, the error standard deviation σ_{inv}, and the error in the location of the peak mean flux. Overall, the results show that the satellite observations result in substantially better inversion accuracy than the in situ observations (Figs. 8 and 9; Table 1). This is most likely the result of the fact that both assimilation and inversion can make use of the greater number of more distant observations to produce a more accurate source estimate. Comparisons between inversion using

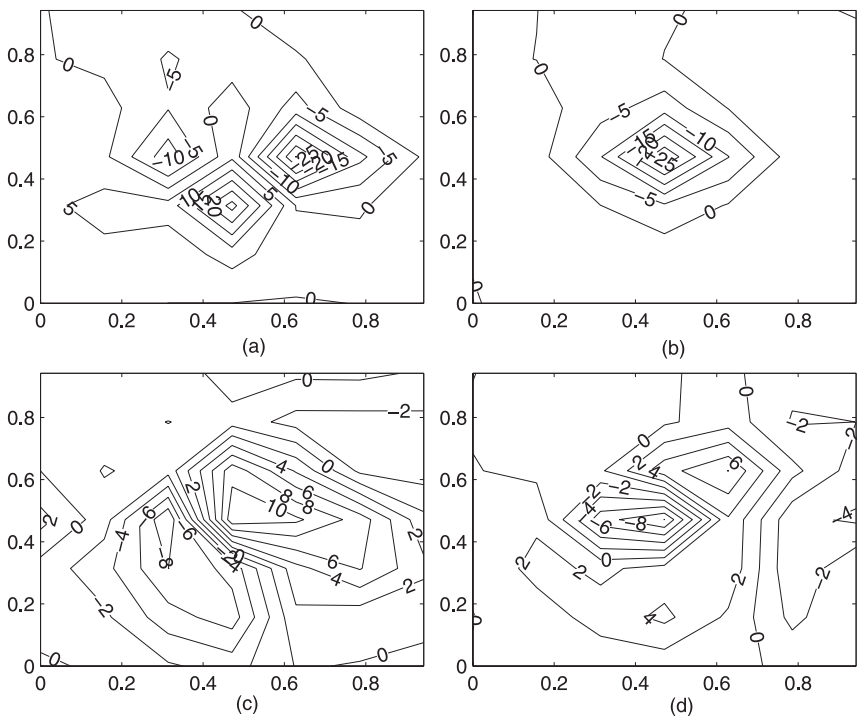

FIG. 8. Ensemble mean errors in source estimates from synthesis inversion using (a) in situ observations (SIN), (b) assimilated in situ observations (SIA), (c) satellite observations (SSN), and (d) assimilated satellite observations (SSA). All cases use the model with spread error.

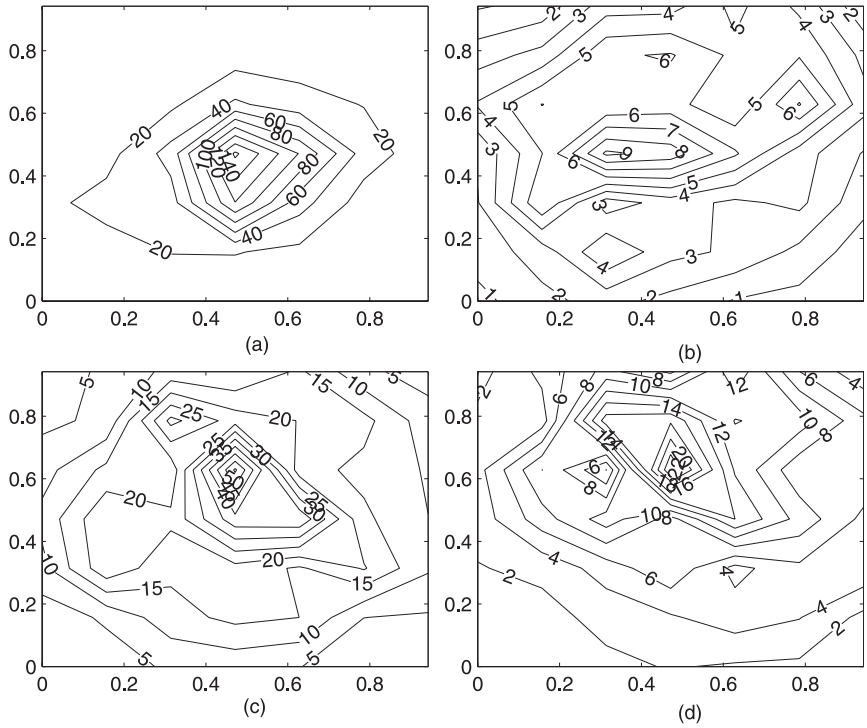

FIG. 9. As in Fig. 8, but for error standard deviation of source estimates from synthesis inversion.

observations directly and those using the assimilated observations are less straightforward. Direct inversion estimates the mean peak flux more accurately when in situ observations are used whereas inversion of the assimilated observations is more accurate when satellite observations are used (Table 1). If we consider the maximum mean error (which is not generally at the same location as the peak), the inversion with assimilated observations is more accurate in three of the four cases (Table 1).

When the model with location error is used, the direct assimilation of observations accurately predicts the peak flux location using either in situ or satellite observations. The inversion using assimilated observations is successful in this regard only when satellite observations are used. Finally, the variability in the solution is far smaller when the observations are first assimilated, as indicated by the large error standard deviations in the direct inversions (Table 1). In addition, the direct inversion created substantial spurious sources and sinks, particularly when using the source model with location error (Figs. 8a,c).

7. Conclusions

We have considered the question of whether assimilating chemical tracer observations into a transport model before carrying out the inversion contributes to the accuracy of the source estimation. The results presented here show that assimilating the observations using a Kalman filter first reduces the random error by factors between 2 and 15 for the cases studied. Improvements to the systematic component of error were less consistent, with decreases in the maximum mean error in most cases but a less accurate prediction of the mean peak flux. The direct inversion of observations results in spurious sources/sinks, whereas the case with assimilation does not. In each case the model or first guess source is globally unbiased but has a local bias.

Because Bayesian source inversion is a statistical weighting of model and observations, the inversion process can never completely overcome any systematic errors. Thus, the actual inversion errors are much larger than the predicted errors (Fig. 7). Additionally, when directly inverting from the observations, the response of the inversion algorithm to these biases is generally to generate spurious sinks in part of the domain while overestimating the source in other parts. When the observations are assimilated first, this tendency is greatly reduced. It is possible that the systematic error in the assimilation could be eliminated using a bias correction scheme (Lamarque et al. 2004).

Most striking is the reduction in the error standard deviation that results from the assimilation. This means

that the accuracy of a single source inversion (as opposed to the ensemble used here) is greatly enhanced by assimilating the observations. The primary reason for this improvement is the more accurate estimate of the error covariance provided by the Kalman filter and the spreading (or smoothing) of observational information. While it is difficult to compare the performance of this simplified system with other inversion systems, Kaminski et al. (2001) showed that errors that result from aggregating source regions in synthesis inversion can be on the order of the emissions themselves. We can therefore state that the reductions found in the present paper are significant in comparison.

The Kalman filter remains a diagnostic tool and is still too computationally expensive for operational data assimilation systems, yet many of its advantages can be translated to other algorithms. Most notably, the ensemble Kalman filter (EnsKF) is being implemented in large-scale atmospheric systems, including trace gas assimilation systems (Arellano et al. 2007). There are also a number of suboptimal Kalman filter algorithms that show some promise for reducing the computational load in evolving error covariances. Finally, even assimilation systems that do not evolve error covariances generally rely on covariance tuning to improve the forecast error estimates. This also acts to improve the inversion computation through improved error statistics.

Acknowledgments. This work is funded by the NASA Modeling, Analysis, and Prediction (MAP) program. R. Cooper was funded by the GSFC Laboratory for Atmospheres summer intern program. The authors gratefully acknowledge the helpful comments of the reviewers.

REFERENCES

Arellano, A. F., P. S. Kasibhatla, L. Giglio, G. R. van der Werf, and J. T. Randerson, 2004: Top-down estimates of global CO sources using MOPITT measurements. *Geophys. Res. Lett.,* **31,** L01104, doi:10.1029/2003GL018609.

——, and Coauthors, 2007: Evaluating model performance of an ensemble-based chemical data assimilation system during INTEX-B field mission. *Atmos. Chem. Phys.,* **7,** 5695–5710.

Auger, L., and A. V. Tangborn, 2004: A wavelet-based reduced rank Kalman filter for assimilation of stratospheric chemical tracer observations. *Mon. Wea. Rev.,* **132,** 1220–1237.

Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan,* **75,** 257–288.

Dargaville, R. J., and I. Simmonds, 2000: Calculating CO_2 fluxes by data assimilation coupled to a three-dimensional mass balance inversion. *Inverse Methods in Global Biogeochemical Cycles, Geophys. Monogr.,* Vol. 14, Amer. Geophys. Union, 256–264.

Enting, I. G., 2000: Green's function methods of tracer inversion. *Inverse Methods in Global Biogeochemical Cycles, Geophys. Monogr.,* Vol. 114, Amer. Geophys. Union, 19–31.

——, 2002: *Inverse Problems in Atmospheric Constituent Transport*. Cambridge University Press, 392 pp.

Gelb, A., 1974: *Applied Optimal Estimation*. MIT Press, 374 pp.

Gilliland, A., and P. J. Abbitt, 2001: A sensitivity study of the discrete Kalman filter (DKF) to initial condition discrepancies. *J. Geophys. Res.,* **106,** 17 939–17 952.

Gurney, K. R., Y.-H. Chen, T. Maki, S. R. Kawa, A. Andrews, and Z. Zhu, 2005: Sensitivity of atmospheric CO_2 inversions to seasonal and interannual variations in fossil fuel emissions. *J. Geophys. Res.,* **110,** D10308, doi:10.1029/2004JD005373.

Hartley, D., and R. Prinn, 1993: Feasibility of determining surface emissions of trace gases using an inverse method in a three-dimensional chemical transport model. *J. Geophys. Res.,* **98,** 5183–5197.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory*. Academic Press, 376 pp.

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

Kaminski, T., M. Heimann, and R. Giering, 1999: A coarse grid three-dimensional global inverse model of the atmospheric transport. 1. Adjoint model and Jacobian matrix. *J. Geophys. Res.,* **104,** 18 535–18 554.

——, P. J. Rayner, M. Heimann, and I. G. Enting, 2001: On aggregation errors in atmospheric transport inversions. *J. Geophys. Res.,* **106,** 4703–4716.

Khattatov, B. V., J.-F. Lamarque, L. V. Lyjak, R. Ménard, P. Levelt, X. Tie, G. P. Brasseur, and J. C. Gille, 2000: Assimilation of satellite observations of long-lived chemical species in global chemistry transport models. *J. Geophys. Res.,* **105,** 29 135–29 144.

Kopacz, M., D. J. Jacob, D. Henze, C. L. Heald, D. G. Streets, and Q. Zhang, 2009: Comparison of adjoint and analytical Bayesian inversion methods for constraining Asian sources of carbon monoxide using satellite (MOPITT) measurements of CO columns. *J. Geophys. Res.,* **114,** D04305, doi:10.1029/2007JD009264.

Lamarque, J.-F., and J. C. Gille, 2003: Improving the modeling of error variance evolution in the assimilation of chemical species: Application to MOPITT data. *Geophys. Res. Lett.,* **30,** 1470, doi:10.1029/2003GL016994.

——, and Coauthors, 2004: Application of a bias estimator for the improved assimilation of Measurements of Pollution in the Troposphere (MOPITT) carbon monoxide retrievals. *J. Geophys. Res.,* **109,** D16304, doi:10.1029/2003JD004466.

Law, R. M., 1999: CO_2 sources from a mass-balance inversion: Sensitivity to the surface constraint. *Tellus,* **51B,** 254–265.

Ménard, R., and L.-P. Chang, 2000: Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part II: χ^2-validated results and analysis of variance and correlation dynamics. *Mon. Wea. Rev.,* **128,** 2672–2686.

——, S. E. Cohn, L.-P. Chang, and P. M. Lyster, 2000: Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part I: Formulation. *Mon. Wea. Rev.,* **128,** 2654–2671.

Peters, W., and Coauthors, 2005: An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations. *J. Geophys. Res.,* **110,** D24304, doi:10.1029/2005JD006157.

Pétron, G., C. Granier, B. Khattatov, J.-F. Lamarque, V. Yudin, J.-F. Müller, and J. Gille, 2002: Inverse modeling of carbon monoxide surface emissions using Climate Monitoring and Diagnostics Laboratory network observations. *J. Geophys. Res.,* **107,** 4761, doi:10.1029/2001JD001305.

——, ——, ——, V. Yudin, J.-F. Lamarque, L. Emmons, J. Gille, and D. P. Edwards, 2004: Monthly CO surface sources

inventory based on the 2000–2001 MOPITT satellite data. *Geophys. Res. Lett.,* **31,** L21107, doi:10.1029/2004GL020560.

Randerson, J. T., M. V. Thompson, T. J. Conway, I. Y. Fung, and C. B. Field, 1997: The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide. *Global Biogeochem. Cycles,* **11,** 535–560.

Rayner, P. J., and D. M. O'Brien, 2001: The utility of remotely sensed CO_2 concentration data in surface source inversions. *Geophys. Res. Lett.,* **28,** 175–178.

——, M. Scholze, W. Knorr, T. Kaminski, R. Giering, and H. Widmann, 2005: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Global Biogeochem. Cycles,* **19,** GB2026, doi:10.1029/2004GB002254.

Tarantola, A., 1987: *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*. Elsevier, 614 pp.

Tucker, C. J., I. Y. Fung, C. D. Keeling, and R. H. Gammon, 1986: Relationship between atmospheric CO_2 variations and a satellite-derived vegetation index. *Nature,* **319,** 195–199.