Manuscript prepared for Geosci. Model Dev.
with version 3.2 of the LaTeX class copernicus.cls.
Date: 25 January 2011

Construction of Non-diagonal Background Error Covariance Matrices for Global Chemical Data Assimilation

K. Singh[1], M. Jardak[1,4], A. Sandu[1], K. Bowman, M. Lee[2], and D. Jones[3]

[1]Department of Computer Science, Virginia Polytechnic Institute and State University, 2202 Kraft Drive, Blacksburg, VA 24060, USA
[2]Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[3]Department of Physics, University of Toronto, ON M5S 1A7, CANADA
[4]Center for Ocean-Atmospheric Prediction Studies (COAPS), Florida State University, Tallahassee, FL 32306, USA

Correspondence to: Adrian Sandu (sandu@cs.vt.edu)

Abstract. Chemical data assimilation attempts to optimally use noisy observations along with imperfect model predictions to produce a better estimate of the chemical state of the atmosphere. It is widely accepted that a key ingredient for successful data assimilation is a realistic estimation of the background error distribution. Particularly important is the specification of the background error covariance matrix, which contains information about the magnitude of the background errors and about their correlations. As models evolve toward finer resolutions, the use of diagonal background covariance matrices is increasingly inaccurate, as they captures less of the spatial error correlations. This paper discusses an efficient computational procedure for constructing non-diagonal background error covariance matrices which account for the spatial correlations of errors. The correlation length scales are specified by the user; a correct choice of correlation lengths is important for a good performance of the data assimilation system. The benefits of using the non-diagonal covariance matrices for variational data assimilation with chemical transport models are illustrated.

1 Introduction

Chemical data assimilation attempts to optimally use noisy observations along with imperfect model predictions to produce a better estimate (in some optimal sense) of the chemical state of the atmosphere. This optimally estimated state better defines the spatial and temporal fields of key chemical components in relation to their sources and sinks. This information is critical for improved studies of the atmospheric composition. Chemical data assimilation could also, in principle, improve estimates

of emission inventories, of model boundary conditions, or of important model parameters like wet deposition velocities or photolysis rates.

The close integration of observational data is recognized as essential in weather/climate analysis and forecast activities. Consequently, considerable experience with data assimilation have been accumulated in the field of numerical weather prediction (Daley, 1991; Courtier et al., 1998; Rabier et al., 2000; Kalnay, 2002; Navon, 2009). In this work we focus on chemical data assimilation, i.e., on assimilation of observations of pollutant levels in the atmosphere. Chemical data assimilation poses specific challenges related to the multiphysics nature of the system, the stiffness of chemical kinetic equations, the sparseness of chemical observations, and the uncertainty in the levels of anthropogenic and natural pollutants emitted into the atmosphere.

Previous studies have employed various approaches to assimilating observations of trace gases for improved tropospheric chemistry representations. The base concepts of the variational approach to chemical data assimilation, and the construction of adjoint chemical transport models are discussed in detail in (Sandu et al., 2005; Carmichael et al., 2008). Early work in chemical data assimilation using variational techniques has been reported in (Fisher and Lary, 1995; Elbern and Schmidt, 1999, 2001). Since then there is a growing body of literature with applications of 4D-Var chemical data assimilation. Adjustment of gas phase chemical tracer initial conditions has been studied in (Chai et al., 2007; Zhang et al., 2008). Adjustment of pollutant emissions through 4D-Var chemical data assimilation has been discussed in Chai et al. (2009). Data assimilation studies involving particle measurements to improve aerosol fields have been performed in (Hakami et al., 2005; Henze et al., 2009).

Suboptimal Kalman filters have been employed successfully for chemical data assimilation (Menard et al., 2000; Lamarque et al., 2002; Segers et al., 2005; Clark et al., 2006; Pierce et al., 2007; Parrington et al., 2009). The use of the ensemble Kalman filter (EnKF) in chemical data assimilation has been studied in (Constantinescu et al., 2007b,c,d). Data assimilation has been used to improve initial conditions, emissions, and boundary values. Besides the initial conditions, improvements in boundary values lead to improved air quality forecasts. Comparisons of the the performance of different techniques for chemical data assimilation have been performed (Lahoz et al., 2007; Wu et al., 2008).

It is widely accepted that a key ingredient of successful data assimilation is a realistic estimation of the background error distribution. Particularly important is the specification of the background error covariance matrix, which contains information about the magnitude of the background errors and about their correlations. Background covariance matrices impact how the information from observations is spread both spatially and among the different types of analysis variables.

The construction of background covariance matrices is challenging due to poorly characterized background errors, and to the very large dimension of the state space of realistic atmospheric models. As a consequence, many chemical data assimilation studies to date have used diagonal background covariance matrices. An early covariance error modeling approach Hollingsworth and Lonnberg (1986)

partitions differences between observations and the background into background errors and observation errors. A popular approach to approximate the background covariance matrix is the NMC method (Parrish and Derber, 1992), in which the differences between several forecasts verifying at the same time are used to approximate the background error. This method has been successfully applied to chemical data assimilation (Chai etal., 2006). The analysis-ensemble method (Fisher, 2003) runs the analysis system several times for the same period with randomly-perturbed observations. Differences between background fields for different runs provide a surrogate for a sample of background error. Additional methods for covariance modeling include the estimation of background error statistics from innovation statistics, digital filters, and the diffusion operator (Fisher, 2006). The statistical structure of forecast errors is used to construct error covariances in (Ingleby, 2001). Derber and Bouttier (1999) consider errors in the spectral space, split covariances into vertical and horizontal, and construct correlations are homogeneous, isotropic, and non-separable. An alternative approach constructs autoregressive models of background errors based on the short-term linearized model dynamics (Constantinescu et al., 2007a). Multivariate, multidimensional background error covariance matrices that maintain the geostrophic and hydrostatic balance have been proposed in (Akella and Navon, 2009).

A popular ansatz is that the background error correlations decay exponentially in space. This ansatz allows the construction of simple error correlation models, and is the basis of the covariance localization technique used in ensemble Kalman filtering (Gaspari and Cohn, 1999; Ott et al., 2004; Constantinescu et al., 2007b). Experimental studies with chemical transport models support this assumption; for example, in (Chai etal., 2006) it has been shown that ozone error correlations decrease follow, on average, an exponential decay curve. Sub-optimal Kalman filters based on covariance models that impose an exponential decay of correlations with distances have been used in the assimilation of chemical constituents (Khattatov et al., 1999; Pierce et al., 2007).

In the troposphere, ozone is an important greenhouse gas and a major pollutant, which adversely impacts air quality. Its distribution is highly heterogeneous, reflecting the combined influence of atmospheric transport and local chemical sources and sinks. Until recently, observations of the three-dimensional structure of tropospheric ozone have been limited. The Tropospheric Emission Spectrometer (TES) satellite instrument, launched in 2004, produced the first continuous, global profile retrievals of tropospheric ozone. Similar observations are now available from other satellite instruments, such as the Infrared Atmospheric Sounding Interferometer (IASI). Assimilating these data into atmospheric models provides a powerful means to obtain an improved understanding of the processes controlling tropospheric ozone. Parrington et al. (2008) was the first to assimilate the TES ozone profile retrievals, but they did not account for horizontal correlations in the background error.

We propose here a computationally efficient approach for constructing (background) error covariances that account for spatial correlations in both horizontal and vertical directions, and assess its impact on the assimilation of tropospheric ozone profiles from TES. The construction is based but not

restricted to the ansatz of exponential decay of error correlations. The correlation lengths in the latitudinal, longitudinal, and vertical directions can be specified according to the application requirements. Due to the large number of state variables an explicit representation of the full background covariance matrix is impractical. The proposed strategy constructs a multi-dimensional correlation matrix from tensor products of one-dimensional correlation matrices. This avoids the explicit construction and storage of full covariance matrices, and allows the needed linear algebra operations to be performed very efficiently.

The paper is organized as follows. Section 2 reviews variational data assimilation techniques, and Section 3 presents the GEOS-Chem global chemical transport model. The algorithm for constructing multidimensional covariance matrices is discussed in Section 4. Section 5 presents assimilation results of TES ozone profiles with the global chemical transport model GEOS-Chem, and illustrates the benefits of nondiagonal covariances in both three and four dimensional variational data assimilation settings.

2 Variational data assimilation

Variational methods solve the data assimilation problem in an optimal control framework (Sasaki, 1958; Le Dimet and Talagrand, 1986; Courtier and Talagrand, 1987; Lions, 1971). Specifically, they attempt to find the control variable values (e.g., initial conditions) which minimize the discrepancy between the model forecast and observations; the minimization is constrained by the governing dynamic equations. In this discussion, for simplicity of presentation, we focus on discrete models (in time and space) where the initial conditions are the control variables.

Data assimilation combines the following three sources of information.

1. The apriori, or background state $\mathbf{x}^b$ represents the best estimate of the true state $\mathbf{x}^t$ available before any measurements are taken. This estimate is assumed unbiased, and the random background (estimation) errors ε^b are typically assumed to have a normal probability density with a background error covariance matrix $\mathbb{B}$

$$\varepsilon^b = \mathbf{x}^b - \mathbf{x}^t \in \mathcal{N}(0, \mathbb{B}). \tag{1}$$

2. The model encapsulates our knowledge about physical and chemical laws that govern the evolution of the system. The model evolves an initial state $\mathbf{x}_0 \in \mathbb{R}^n$ at the initial time t_0 to future state values $\mathbf{x}_i \in \mathbb{R}^n$ at future times t_i,

$$\mathbf{x}_i = \mathcal{M}_{t_0 \to t_i}(\mathbf{x}_0). \tag{2}$$

   The size of the state space in realistic chemical transport models is very large. For example, a GEOS-Chem simulation at the $2^\circ x 2.5^\circ$ horizontal resolution has $n \in \mathcal{O}(10^8)$ variables.

3. Observations $\mathbf{x}_i^{\text{obs}} \in \mathbb{R}^m$ of the state are taken at times t_i, $1 = 1, \cdots, N$

$$\mathbf{x}_i^{\text{obs}} = \mathcal{H}(\mathbf{x}_i) + \varepsilon_i^{\text{obs}}. \tag{3}$$

The observation operator $\mathcal{H}$ maps the state space onto the observation space. In many practical situations $\mathcal{H}$ is a highly nonlinear mapping (as is the case, e.g., with satellite observation operators). Usually the observations are sparsely distributed, and the number of observations is small compared to the dimension of the state space, $m \ll n$.

The observations are corrupted by measurement and representativeness errors $\varepsilon_i^{\text{obs}}$. The observation errors at each time are assumed to be independent of background errors, and independent of the observation errors at other times. They are typically assumed to have a normal distribution with mean zero and covariance $\mathbb{R}_i$,

$$\varepsilon_i^{\text{obs}} \in \mathcal{N}(0, \mathbb{R}_i). \tag{4}$$

Based on these three sources of information data assimilation computes the posterior estimate $\mathbf{x}^a$ of the true state; $\mathbf{x}^a$ is called the "analysis".

2.1 Three dimensional variational (3D-Var) data assimilation

In the 3D-Var data assimilation the observations (3) are considered successively at times $t_1, \cdots, t_N$. The background state (i.e., the best state estimate at time t_i) is given by the model forecast, starting from the previous analysis (i.e., best estimate at time t_{i-1}):

$$\mathbf{x}_i^b = \mathcal{M}_{t_{i-1} \to t_i}\left(\mathbf{x}_{i-1}^a\right).$$

The discrepancy between the model state $\mathbf{x}_i$ and observations at time t_i, together with the departure of the state from the model forecast $\mathbf{x}_i^b$, are measured by the 3D-Var cost function:

$$\mathbb{J}(\mathbf{x}_i) = \frac{1}{2}\left(\mathbf{x}_i - \mathbf{x}_i^b\right)^T \mathbb{B}^{-1}\left(\mathbf{x}_i - \mathbf{x}_i^b\right) + \frac{1}{2}\left(\mathcal{H}(\mathbf{x}_i) - \mathbf{x}_i^{obs}\right)^T \mathbb{R}_i^{-1}\left(\mathcal{H}(\mathbf{x}_i) - \mathbf{x}_i^{obs}\right) \tag{5}$$

While in principle a different background covariance matrix should be used at each time, in practice the same matrix is re-used throughout the assimilation window. The 3D-Var analysis is computed as the state which minimizes (5)

$$\mathbf{x}_i^a = \operatorname{argmin}\ \mathbb{J}(\mathbf{x}_i). \tag{6}$$

Typically a gradient-based numerical optimization procedure is employed to solve (6). The gradient $\nabla\mathbb{J}$ of the cost function (5) is

$$\nabla\mathbb{J}(\mathbf{x}_i) = \mathbb{B}^{-1}\left(\mathbf{x}_i - \mathbf{x}_i^b\right) + \left(\mathcal{H}'(\mathbf{x}_i)\right)^T \mathbb{R}_i^{-1}\left(\mathcal{H}(\mathbf{x}_i) - \mathbf{x}_i^{obs}\right) \tag{7}$$

Note that the gradient requires computation of the linearized observation operator $\mathcal{H}'$ about the current state.

Preconditioning is often used to improve convergence of the numerical optimization problem (6). A change of variables is performed, for example, by shifting the state and scaling it with the square root of covariance:

$$\widehat{\mathbf{x}}_i = \mathbb{B}^{1/2}\left(\mathbf{x}_i - \mathbf{x}_i^b\right), \tag{8}$$

The optimization is then carried out on the new variables $\widehat{\mathbf{x}}_i$.

2.2 Four dimensional variational (4D-Var) data assimilation

In strongly-constrained 4D-Var data assimilation all observations (3) at all times $t_1,\cdots,t_N$ are simultaneously considered. The control parameters are the initial conditions $\mathbf{x}_0$; they uniquely determine the state of the system at all future times via the model equation (2).

The discrepancy between model predictions and observations at all future times $t_1,\cdots,t_N$, together with the departure of the initial state from the background state, are measured by the 4D-var cost function:

$$\mathbb{J}(\mathbf{x}_0) = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^b\right)^T \mathbb{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^b\right) + \frac{1}{2}\sum_{i=1}^{N}\left(\mathcal{H}(\mathbf{x}_i) - \mathbf{x}_i^{obs}\right)^T \mathbb{R}_i^{-1}\left(\mathcal{H}(\mathbf{x}_i) - \mathbf{x}_i^{obs}\right) \tag{9}$$

Note that the departure of the initial conditions from the background is weighted by the inverse background covariance matrix, $\mathbb{B}^{-1}$, while the differences between the model predictions $\mathcal{H}(\mathbf{x}_i)$ and observations $\mathbf{x}_i^{obs}$ are weighted by the inverse observation error covariances, $\mathbb{R}_i^{-1}$.

The 4D-Var analysis is computed as the initial condition which minimizes (9) subject to the model equation constraints (2)

$$\mathbf{x}_0^a = \text{argmin } \mathbb{J}(\mathbf{x}_0) \quad \text{subject to} (2). \tag{10}$$

The model (2) propagates the optimal initial condition (9) forward in time to provide the analysis at future times, $\mathbf{x}_i^a = \mathcal{M}_{t_0 \to t_i}\left(\mathbf{x}_0^a\right)$.

The optimization problem (10) is solved numerically using a gradient-based technique. The gradient of (9) reads

$$\nabla\mathbb{J}(\mathbf{x}_0) = \mathbb{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^b\right) + \sum_{i=1}^{N}\left(\frac{\partial \mathbf{x}_i}{\partial \mathbf{x}_0}\right)^T \left(\mathcal{H}'(\mathbf{x}_i)\right)^T \mathbb{R}_i^{-1}\left(\mathcal{H}(\mathbf{x}_i) - \mathbf{x}_i^{obs}\right) \tag{11}$$

The 4D-Var gradient requires not only the linearized observation operator $\mathcal{H}'$, but also the transposed derivative of future states with respect to the initial conditions. The 4D-Var gradient can be obtained effectively by forcing the adjoint model with observation increments, and running it backwards in time. The construction of an adjoint model requires considerable effort, time, and know-how.

3 GEOS-Chem

In this paper we specifically consider GEOS-Chem (http://geos-chem.org), a global three-dimensional chemical transport model (CTM) driven by assimilated meteorological fields from Goddard Earth Observing System(GEOS-4) at the NASA Global Modeling and Assimilation Ofice (GMAO). The model was first described in Bey et al. (2001). GEOS-Chem accounts in detail for emissions from both natural and anthropogenic sources, for gas phase chemistry, aerosol processes, long range transport of

pollutants, troposphere-stratosphere exchanges, etc. GEOS-Chem is widely used by research groups world-wide for performing global atmospheric chemistry studies.

The GEOS-Chem Adjoint system (http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-Chem_Adjoint) has been developed through a joint effort of groups at Virginia Tech, University of Colorado, Caltech, Jet Propulsion Laboratory, and Harvard (Henze et al., 2007; Singh et al., 2009a,b; Eller et al., 2009). The system can perform adjoint sensitivity analyses and 4D-Var chemical data assimilation. Inverse modeling studies with GEOS-Chem adjoint are presented in (Kopacz et al., 2009; Henze et al., 2009).

4 Construction of the background error covariance matrix

A correct characterization of the background errors is necessary for obtaining a meaningful analysis, i.e., for the success of the data assimilation procedure. Under the usual assumption, the background errors are normally distributed where their probability density is described by the background state $\mathbf{x}^b$ and the background error covariance matrix $\mathbb{B}$. In variational data assimilation both $\mathbf{x}^b$ and $\mathbb{B}$ enter directly into the formulation of the cost function; errors in their specification directly impact the analysis result (Daescu, 2008).

A non-diagonal background error covariance matrix allows the information from local observations to spread out in space to contribute to corrections of state variables in neighboring locations; similarly, it allows observations of certain components of the state vector to contribute to corrections of other components. These contributions are more prominent in case of 3D-Var since state variables are corrected every subsequent observation interval. The spread of information results in a smooth analysis state, and allows different sets of observations to complement each other. In case of 4D-Var, the state variables are corrected only at the start of the assimilation window using information from all observations combined. The spreading of information is achieved through background error covariance matrix and transport of adjoint sensitivities.

Despite these advantages, most chemical data assimilation studies to date have employed diagonal background covariances. Little work has been devoted to date to modeling off-diagonal terms (Chai etal., 2006; Constantinescu et al., 2007a). This is due to a number of practical difficulties that arise in the construction of background covariance matrices. The "true" state is, fundamentally, unknown, and so are the "true" errors; surrogate states have to be used to mimic forecast errors. Ensembles can be employed to estimate error correlations; however, the number of ensemble members is necessarily very small and only low rank approximations of the covariance matrix can be obtained. Localization is often employed to remove spurious correlations and to improve the rank of the resulting matrix (Gaspari and Cohn, 1999). The large number of state variables make the construction and storage of full covariance matrices impractical.

We next discuss the proposed approach to constructing a background error covariance matrix $\mathbb{B}$ that

accounts for both vertical and horizontal correlations without explicitly constructing the full covariance matrix. We explain the construction of the matrix in the two-dimensional case, i.e., for capturing horizontal correlations; the extensions to correlations in three dimensions and to correlations among multiple state variables are immediate. Our target application is global chemical data assimilation using GEOS-Chem.

Consider a uniform latitude-longitude grid and denote by x the longitude and by y the latitude level. A grid point (x_i,y_j) has longitude coordinate x_i, $i=1,\ldots,n_x$, and latitude coordinate y_j, $i=1,\ldots,n_y$. The state vector contains the state values at all gridpoints ordered latitude-first:

$$\Big[(x_1,y_1),\cdots,(x_1,y_{n_y}),(x_2,y_1),\cdots,(x_2,y_{n_y}),(x_{n_x},y_1),\cdots,(x_{n_x},y_{n_y})\Big] \tag{12}$$

4.1 Directional error correlation matrices

The one-dimensional correlation between errors at two locations (x_i,y_k) and (x_j,y_k) situated at the same latitude y_k is modeled as

$$\left(\widetilde{\mathbb{C}}_x^k\right)_{i,j} = \mathrm{corr}\left((x_i,y_k),(x_j,y_k)\right) = e^{-\frac{\mathrm{dist}\left((x_i,y_k),(x_j,y_k)\right)^2}{\ell_x^2}}; \quad i,j=1,\cdots,n_x; \quad k=1,\cdots,n_y; \tag{13}$$

where ℓ_x is the correlation distance in the longitude direction. For a uniform lat-lon grid the distance between x_i and x_j depends only on $\min(|i-j|,n_x-|i-j|)$. This distance also depends on the y_k; for this reason equation (13) defines a different longitudinal correlation matrix $\widetilde{\mathbb{C}}_x^k \in \mathbb{R}^{n_x\times n_x}$ for each latitude y_k. Due to the periodicity along each latitude circle the point x_1 is strongly correlated with both x_2 and x_{n_x}, etc. The periodicity is captured by the distance function in (13).

Similarly, the one-dimensional correlation between errors at two locations (x_k,y_i) and (x_k,y_j) situated at the same longitude x_k is modeled as

$$\left(\widetilde{\mathbb{C}}_y^k\right)_{i,j} = \mathrm{corr}\left((x_k,y_i),(x_k,y_j)\right) = e^{-\frac{\mathrm{dist}\left((x_k,y_i),(x_k,y_j)\right)^2}{\ell_y^2}}, \quad i,j=1,\cdots,n_y; \quad k=1,\cdots,n_y; \tag{14}$$

where ℓ_y is the correlation distance in the latitude direction. Equation (14) defines a single latitudinal correlation matrix $\widetilde{\mathbb{C}}_y \in \mathbb{R}^{n_y\times n_y}$. For a uniform lat-lon grid this correlation matrix is the same for each longitude x_k; consequently the superscript k is dropped. To simplify the construction the correlations due to the periodicity along a meridional circle are ignored. Otherwise, error correlations across the poles would lead to correlations between errors at all longitudes; such cross-correlations are not captured by (14).

The cost function and gradient calculations described in equations (5), (7), (9), and (11) require the inverse of the background error covariance matrix; this involves the inverses of the correlation matrices in longitudinal and latitudinal directions. The construction of the directional correlation matrices $\widetilde{\mathbb{C}}_x$ and $\widetilde{\mathbb{C}}_y$ does not guarantee that they are non-singular. To avoid a possible singularity we take a convex

combination between the identity matrix and tensor product correlations as follows:

$$\mathbb{C}_x^k = \theta_x \, \mathbb{I}_{n_x \times n_x} + (1 - \theta_x) \, \widetilde{\mathbb{C}}_x^k \, , \tag{15}$$

and

$$\mathbb{C}_y^k = \theta_y \, \mathbb{I}_{n_y \times n_y} + \left(1 - \theta_y\right) \, \widetilde{\mathbb{C}}_y^k \, . \tag{16}$$

The above procedure brings a shift in the spectrum and ensures the positive definiteness of $\mathbb{C}_x$ and $\mathbb{C}_y$. In all our experiments both θ_x and θ_y are chosen to be equal to 0.2.

The longitudinal correlation matrix between all the points on the two-dimensional grid (12) can be constructed from the one-dimensional longitudinal correlation matrices as follows

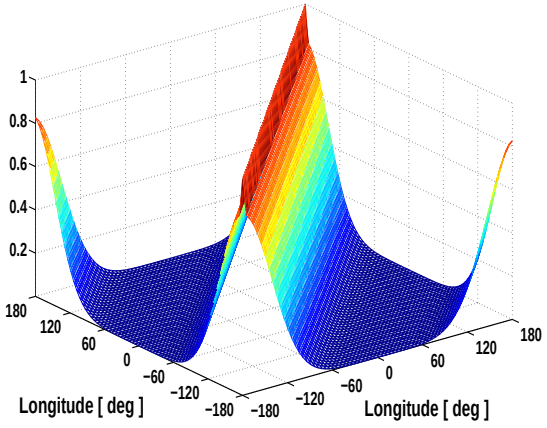
$$\mathbb{C}_x^{1:n_y} \otimes \mathbb{I}_{n_y \times n_y} = \begin{pmatrix} \begin{pmatrix} \left(\mathbb{C}_x^1\right)_{1,1} & 0 & \cdots & 0 \\ 0 & \left(\mathbb{C}_x^2\right)_{1,1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\mathbb{C}_x^{n_y}\right)_{1,1} \end{pmatrix} & \cdots & \begin{pmatrix} \left(\mathbb{C}_x^1\right)_{1,n_x} & 0 & \cdots & 0 \\ 0 & \left(\mathbb{C}_x^2\right)_{1,n_x} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\mathbb{C}_x^{n_y}\right)_{1,n_x} \end{pmatrix} \\ \vdots & \ddots & \vdots \\ \begin{pmatrix} \left(\mathbb{C}_x^1\right)_{n_x,1} & 0 & \cdots & 0 \\ 0 & \left(\mathbb{C}_x^2\right)_{n_x,1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\mathbb{C}_x^{n_y}\right)_{n_x,1} \end{pmatrix} & \cdots & \begin{pmatrix} \left(\mathbb{C}_x^1\right)_{n_x,n_x} & 0 & \cdots & 0 \\ 0 & \left(\mathbb{C}_x^2\right)_{n_x,n_x} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\mathbb{C}_x^{n_y}\right)_{n_x,n_x} \end{pmatrix} \end{pmatrix}$$

With some abuse of notation we extend the use of the Kronecker product symbol $\otimes$ in the above equation in order to highlight the structure of the two-dimensional longitudinal correlation matrix.

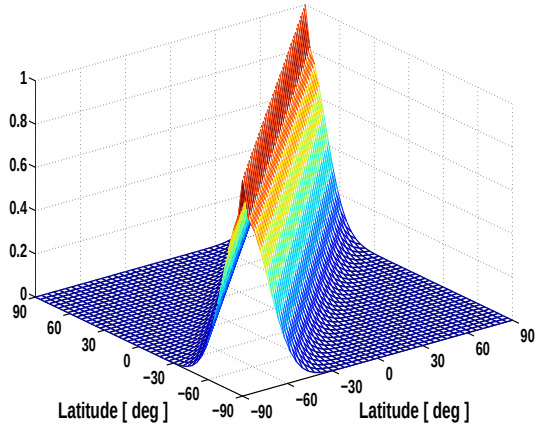
Similarly, the latitudinal correlation matrix between all the points on the two-dimensional grid (12) can be constructed from the one-dimensional latitudinal correlation matrices as follows

$$\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y = \begin{pmatrix} \begin{pmatrix} \left(\mathbb{C}_y\right)_{1,1} & \left(\mathbb{C}_y\right)_{1,2} & \cdots & \left(\mathbb{C}_y\right)_{1,n_y} \\ \left(\mathbb{C}_y\right)_{2,1} & \left(\mathbb{C}_y\right)_{2,2} & \cdots & \left(\mathbb{C}_y\right)_{2,n_y} \\ \vdots & \vdots & \ddots & \vdots \\ \left(\mathbb{C}_y\right)_{n_y,1} & \left(\mathbb{C}_y\right)_{n_y,2} & \cdots & \left(\mathbb{C}_y\right)_{n_y,n_y} \end{pmatrix} & \cdots & \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \\ \vdots & \ddots & \vdots \\ \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} & \cdots & \begin{pmatrix} \left(\mathbb{C}_y\right)_{1,1} & \left(\mathbb{C}_y\right)_{1,2} & \cdots & \left(\mathbb{C}_y\right)_{1,n_y} \\ \left(\mathbb{C}_y\right)_{2,1} & \left(\mathbb{C}_y\right)_{2,2} & \cdots & \left(\mathbb{C}_y\right)_{2,n_y} \\ \vdots & \vdots & \ddots & \vdots \\ \left(\mathbb{C}_y\right)_{n_y,1} & \left(\mathbb{C}_y\right)_{n_y,2} & \cdots & \left(\mathbb{C}_y\right)_{n_y,n_y} \end{pmatrix} \end{pmatrix}$$

The structure of the one-dimensional correlation matrices is represented in Figure 1. The longitudinal correlation $\mathbb{C}_x^k$ is represented at latitude $y_k = 20^\circ N$; note that due to the periodicity along each latitude



(a) Longitudinal correlation matrix $\mathbb{C}_x^k$ at latitude $y_k = 20^\circ N$



(b) Latitudinal correlation matrix $\mathbb{C}_y$

Fig. 1. Mesh representation of the one-dimensional longitudinal and latitudinal correlation matrices. The latitude-longitude model grid resolution is $4^\circ \times 5^\circ$ (about 400Km $\times$ 500Km near the equator) and the correlation lengths are ℓ_x = 1500Km and ℓ_y = 1200Km.

circle not only the elements near the diagonal, but also the elements in the corners of the matrix have non-zero values. The latitudinal correlation $\mathbb{C}_y$ does not account for periodicity (along each meridian
205 the grids 1 and n_y correspond to the South and to the North pole, respectively).

Figure 2 represents contour lines of the longitudinal correlation $\mathbb{C}_x$ for points at different latitudes. The correlation length ℓ_x is short (top panel), medium (middle panel), and large (bottom panel). Note that the same correlation length ℓ_x translates into a larger number of correlated grid points at higher latitudes.

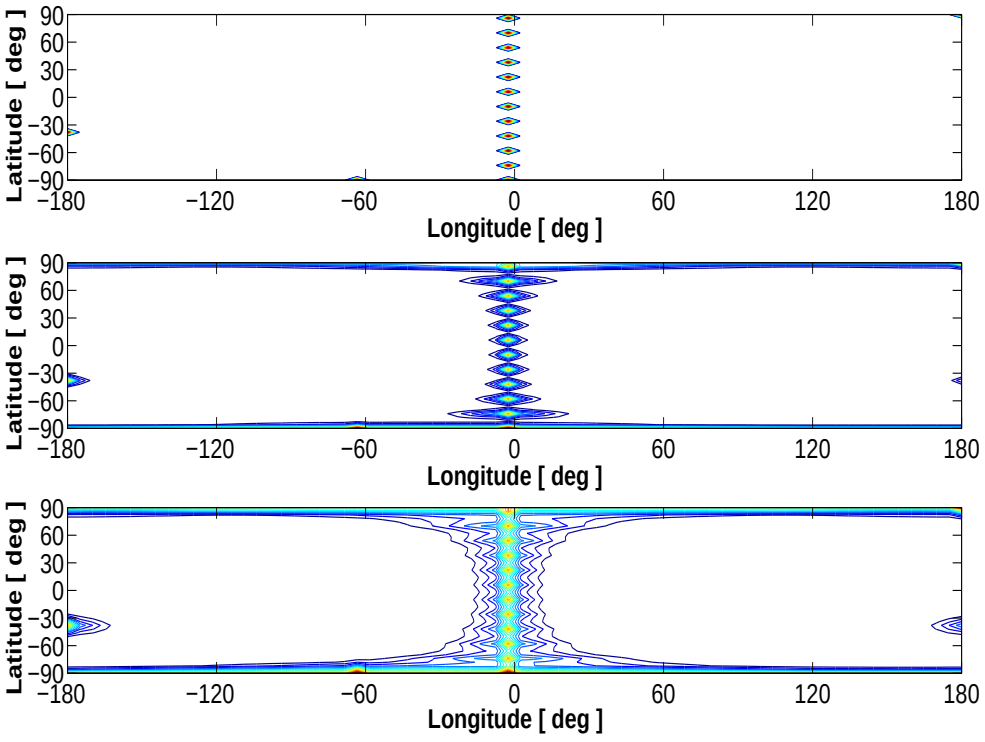


Fig. 2. Contour lines of the longitudinal correlation $\mathbb{C}_x$ for points at different latitudes. The correlation length ℓ_x is short (top panel), medium (middle panel), and large (bottom panel). Note that the same correlation length ℓ_x translates into a different number of correlated grid points depending on the latitude.

4.2 Two-dimensional covariance matrices

Formally the full background error correlation matrix $\mathbb{C} \in \mathbb{R}^{n_x n_y \times n_x n_y}$ (which accounts for both latitudinal and longitudinal correlations) is constructed via the following relation

$$\mathbb{C} = \left(\mathbb{C}_x^{1:n_y} \otimes \mathbb{I}_{n_y \times n_y} \right) \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y \right). \tag{17}$$

Note that this (huge) matrix is never explicitly formed. One needs to form and store only n_y one-dimensional longitudinal correlation matrices (13) and a single one-dimensional latitudinal correlation matrix (14). Note that the diagonal entries of the tensor product matrix (17) are all equal to one. The tensor product matrix (17) is not symmetric.

A symmetric version of the two-dimensional correlation matrix can be constructed as follows. Any symmetric positive definite matrix $\mathbb{C}$ has a matrix square root $\mathbb{C}^{1/2}$ such that

$$\mathbb{C} = \mathbb{C}^{1/2} \mathbb{C}^{T/2}.$$

The matrix square root is not uniquely defined; in particular it can be symmetric or not depending on the decomposition method used as described in equations (21) and (22). Let $\left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y \right)^{1/2}$ be a

square root of the longitudinal correlation matrix. The symmetric two-dimensional correlation matrix can be constructed as:

$$\mathbb{C}^{\text{sym}} = \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y\right)^{1/2} \cdot \left(\mathbb{C}_x^{1:n_y} \otimes \mathbb{I}_{n_y \times n_y}\right) \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y\right)^{T/2}. \tag{18}$$

Let $\sigma_{i,j}$ be the standard deviation of the error at location (x_i, y_j) and

$$\Sigma = \text{diag}_{1 \le i \le n_x, 1 \le j \le n_y} \sigma_{i,j}$$

the diagonal matrix with all standard deviations at all grid points ordered according to (12). The two-dimensional covariance matrix is constructed from the correlation matrix (17) by scaling it from left and right with the diagonal matrix of standard deviations Σ

$$\mathbb{B} = \Sigma \cdot \mathbb{C} \cdot \Sigma = \Sigma \cdot \left(\mathbb{C}_x^{1:n_y} \otimes \mathbb{I}_{n_y \times n_y}\right) \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y\right) \cdot \Sigma. \tag{19}$$

Similarly, a symmetric version of the covariance matrix can be constructed from the symmetric correlation (18) as follows:

$$\mathbb{B}^{\text{sym}} = \Sigma \cdot \mathbb{C}^{\text{sym}} \cdot \Sigma = \Sigma \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y\right)^{1/2} \left(\mathbb{C}_x^{1:n_y} \otimes \mathbb{I}_{n_y \times n_y}\right) \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y\right)^{T/2} \cdot \Sigma. \tag{20}$$

4.3 Efficient covariance matrix function calculations

The symmetric positive definite one-directional longitudinal correlation matrix has a matrix square root $\mathbb{C}_y^{1/2}$. A symmetric square root, the inverse of the symmetric square root, and the matrix inverse can be obtained via the singular value decomposition (SVD)

$$\mathbb{C}_y = \mathbb{U}\Theta\mathbb{U}^T, \quad \left(\mathbb{C}_y\right)^r = \mathbb{U}\Theta^r\mathbb{U}^T, \quad \text{for } r \in \left\{\frac{1}{2}, -\frac{1}{2}, -1\right\}. \tag{21}$$

while a nonsymmetric square root can be obtained via a Cholesky decomposition

$$\mathbb{C}_y = \mathbb{L}_y\mathbb{L}_y^T, \quad \mathbb{C}_y^{1/2} = \mathbb{L}_y. \tag{22}$$

By the properties of the Kronecker product we have that the square root, the inverse square root, and the inverse of the two-dimensional longitudinal correlation matrix can be constructed in terms of the same matrix functions applied to the one dimensional longitudinal correlations:

$$\left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y\right)^r = \mathbb{I}_{n_x \times n_x} \otimes \left(\mathbb{C}_y\right)^r \quad \text{for } r \in \left\{\frac{1}{2}, -\frac{1}{2}, -1\right\}. \tag{23}$$

Consequently, the symmetric covariance (20) can be implemented as

$$\mathbb{B}^{\text{sym}} = \Sigma \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \left(\mathbb{C}_y\right)^{1/2}\right) \cdot \left(\mathbb{C}_x^{1:n_y} \otimes \mathbb{I}_{n_y \times n_y}\right) \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \left(\mathbb{C}_y\right)^{T/2}\right) \cdot \Sigma$$

using either of the one-dimensional square roots.

Similarly, different powers of each one-dimensional latitudinal correlation matrix can be obtained via a singular value decomposition:

$$\mathbb{C}_x^{1:n_y} = \mathbb{V}\Gamma\mathbb{V}^T, \quad \left(\mathbb{C}_x^{1:n_y}\right)^r = \mathbb{V}\Gamma^r\mathbb{V}^T, \quad \text{for } r \in \left\{\frac{1}{2}, -\frac{1}{2}, -1\right\}.$$

By the properties of the extended Kronecker product we have that the square root, the inverse square root, and the inverse of the two-dimensional latitudinal correlation matrix can be constructed in terms of the same matrix functions applied to the one dimensional latitudinal correlations:

$$\left(\mathbb{C}_x^{1:n_y} \otimes \mathbb{I}_{n_y \times n_y}\right)^r = \left(\mathbb{C}_x^{1:n_y}\right)^r \otimes \mathbb{I}_{n_y \times n_y} \quad \text{for } r \in \left\{\frac{1}{2}, -\frac{1}{2}, -1\right\}. \tag{24}$$

We now use these relations to build functions of the covariance matrices. The inverse of the background covariance is needed in the formulation of the variational cost function. The inverse of the non-symmetric covariance (19) is

$$\mathbb{B}^{-1} = \Sigma^{-1} \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \left(\mathbb{C}_y\right)^{-1}\right) \cdot \left(\left(\mathbb{C}_x^{1:n_y}\right)^{-1} \otimes \mathbb{I}_{n_y \times n_y}\right) \cdot \Sigma^{-1}. \tag{25}$$

The inverse of the symmetric covariance (20) matrix is

$$(\mathbb{B}^{\text{sym}})^{-1} = \Sigma^{-1} \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \left(\mathbb{C}_y\right)^{-T/2}\right) \left(\left(\mathbb{C}_x^{1:n_y}\right)^{-1} \otimes \mathbb{I}_{n_y \times n_y}\right) \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \left(\mathbb{C}_y\right)^{-1/2}\right) \cdot \Sigma^{-1}. \tag{26}$$

Finally, the symmetric covariance (20) has a (non-symmetric) matrix square root

$$(\mathbb{B}^{\text{sym}})^{1/2} = \Sigma \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \left(\mathbb{C}_y\right)^{1/2}\right) \cdot \left(\left(\mathbb{C}_x^{1:n_y}\right)^{1/2} \otimes \mathbb{I}_{n_y \times n_y}\right). \tag{27}$$

This is built out of tensor products involving the square roots of the one-dimensional correlation matrices. The inverse of the square root matrix (27) is

$$(\mathbb{B}^{\text{sym}})^{-1/2} = \left(\left(\mathbb{C}_x^{1:n_y}\right)^{-1/2} \otimes \mathbb{I}_{n_y \times n_y}\right) \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \left(\mathbb{C}_y\right)^{-1/2}\right) \cdot \Sigma^{-1}. \tag{28}$$

4.4 Efficient linear algebra operations involving the covariance matrix

Matrix vector operations involving $\mathbb{B}$ can be performed effectively by exploiting its structure. Consider a vector of concentrations (or concentration errors) $u_{i,j}$ - indexed by latitude and longitude, but stored as a state vector with the convention (12).

Consider the non-symmetric covariance matrix. The covariance matrix-vector product $v = \mathbb{B} \cdot u$ can be computed in stages. Each stage produces a temporary result which is a two-dimensional vector.

Expression	Computation	
$\alpha = \Sigma \cdot u$	$\alpha_{i,j} = \sigma_{i,j} u_{i,j}$,	for $i = 1,\cdots,n_x, j = 1,\cdots,n_y$.
$\beta = \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y\right) \cdot \alpha$	$\beta_{i,1:n_y} = \mathbb{C}_y \cdot \alpha_{i,1:n_y}$,	for $i = 1,\cdots,n_x$.
$\gamma = \left(\mathbb{C}_x^{1:n_y} \otimes \mathbb{I}_{n_y \times n_y}\right) \cdot \beta$	$\gamma_{1:n_x,j} = \mathbb{C}_x^j \cdot \beta_{1:n_x,j}$,	for $j = 1,\cdots,n_y$.
$v = \Sigma \cdot \gamma$	$v_{i,j} = \sigma_{i,j} \gamma_{i,j}$,	for $i = 1,\cdots,n_x, j = 1,\cdots,n_y$.

(29)

Similarly the inverse covariance matrix-vector product $v = \mathbb{B}^{-1} \cdot u$ can be computed as follows

$$
\begin{array}{lll}
\text{Expression} & \text{Computation} & \\
\alpha = \Sigma^{-1} \cdot u & \alpha_{i,j} = u_{i,j} / \sigma_{i,j}, & \text{for } i = 1, \cdots, n_x, j = 1, \cdots, n_y. \\
\beta = \left(\mathbb{C}_x^{1:n_y} \otimes \mathbb{I}_{n_y \times n_y} \right)^{-1} \cdot \alpha & \text{Solve} \mathbb{C}_x^j \cdot \beta_{1:n_x,j} = \alpha_{1:n_x,j}, & \text{for } j = 1, \cdots, n_y. \\
\gamma = \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbb{C}_y \right)^{-1} \cdot \beta & \text{Solve} \mathbb{C}_y \cdot \gamma_{i,1:n_y} = \beta_{i,1:n_y}, & \text{for } i = 1, \cdots, n_x. \\
v = \Sigma^{-1} \cdot \gamma & v_{i,j} = \gamma_{i,j} / \sigma_{i,j}, & \text{for } i = 1, \cdots, n_x, j = 1, \cdots, n_y.
\end{array}
\tag{30}
$$

Similar procedures can be developed for the symmetric covariance time vector products.

The square root (27) times vector product $v = (\mathbb{B}^{\text{sym}})^{1/2} u$ is computed as:

$$
\begin{array}{lll}
\text{Expression} & \text{Computation} & \\
\alpha = \left(\left(\mathbb{C}_x^{1:n_y} \right)^{1/2} \otimes \mathbb{I}_{n_y \times n_y} \right) \cdot u & \gamma_{1:n_x,j} = \mathbb{C}_x^j \cdot \beta_{1:n_x,j}, & \text{for } j = 1, \cdots, n_y \\
\beta = \left(\mathbb{I}_{n_x \times n_x} \otimes \left(\mathbb{C}_y \right)^{1/2} \right) \cdot \alpha & \beta_{i,1:n_y} = \mathbb{C}_y^{1/2} \cdot \alpha_{i,1:n_y}, & \text{for } i = 1, \cdots, n_x \\
v = \Sigma \cdot \beta & v_{i,j} = \sigma_{i,j} \gamma_{i,j}, & \text{for } i = 1, \cdots, n_x, j = 1, \cdots, n_y.
\end{array}
\tag{31}
$$

All the above implementations are based on repeated operations involving the one-dimensional covariance matrices and their square roots. These operations are very efficient since all the linear algebra operations (matrix-vector multiplication, SVD, Choesky factorization, the solution of linear systems) are performed on small dimensional matrices ($n_x \times n_x$ or $n_y \times n_y$).

5 Numerical experiments

For numerical experiments, we employ GEOS-Chem v7-04-10 adjoint code (Singh et al., 2009b), capable of performing both 3D-Var and 4D-Var data assimilations with real data. We assimilate Tropospheric Emission Spectormeter (TES) satellite ozone profile retrievals into the model and validate the generated analyses through an independent observation dataset provided by direct ozone profile measurements from Ozonesondes. The numerical optimization method used in all experiments is the limited memory bound-constrained BFGS (Zhu et al., 1997). This quasi-Newton approach has become the "gold standard" in solving large scale chemical data assimilation problems (Sandu et al., 2005).

5.1 Experimental setting

Simulations with GEOS-Chem v7 adjoint can be carried out at $4^\circ \times 5^\circ$ and $2^\circ \times 2.5^\circ$ resolutions. We have used $4^\circ \times 5^\circ$ resolution in all our experiments. There are 46 x 72 latitude-longitude grid boxes at this resolution, and 55 vertical levels; near the equator and at ground level each grid box covers an area of about 400 Km $\times$ 500 Km. The current GEOS-Chem model does not capture well the dynamics of the upper troposphere and of the stratosphere. Therefore, we performed data assimilation for only the first 23 model levels (for up to about 50 hPa). The model has been modified to use the linearized

ozone (linoz) scheme for a better estimation of ozone exchanges at troposphere-stratosphere boundary. This scheme is going to be used in the next release of the standard GEOS-Chem model.

The 3D-Var data assimilation experiments were performed over the months of July and August 2006, starting at 00:00(GMT) on July 1st. The TES satellite data was read once every 4 simulation hours; the observation operator called at model time t (hours) reads in all the measurements collected within the interval $t-2$ (hours) to $t+2$ (hours). 3D-Var data assimilation treats all observations in this interval as instantaneous, and assimilates them in the same optimization run. In all our 3D-Var experiments, we performed 8 iterations per analysis since the cost function decreased significantly within the first few iterations. It is important to note that 3D-Var does not involve any model adjoint calculations; gradients require only the adjoint of the observation operator. The optimization adjusts ozone concentrations.

The 4D-Var data assimilation experiments were performed over a 5 day assimilation window starting at 00:00(GMT) on August 1st, 2006 and ending at 00:00(GMT) on August 6th of the same year. The background initial conditions were generated through a free GEOS-Chem model run. There were 12 optimization iterations performed in order to improve the ozone initial condition. Each iteration during 4D-Var assimilation includes a forward model and a backward model adjoint run. TES satellite profile retrievals are read every 4 hours during the model adjoint run, and the cost function and adjoint gradients accumulate the impact of all 4 hour data sets throughout the assimilation window.

NOTE: The TES data can be biased by as much as 10% (Nassar et al., 2008). We removed this bias as estimated by (Nassar et al., 2008) before assimilating the data.

5.1.1 Computational costs

As described in Section 1, the construction of the background error covariance matrix $\mathbb{B}$ impacts the result of the data assimilation. If one considers no correlation among different model grid points, or among different chemical species, $\mathbb{B}$ turns out to be diagonal. However, this approximation is inaccurate as the ozone errors are highly correlated spatially (Constantinescu et al., 2007a,c,d) and correlated to errors in other chemical species; this correlation is not discussed in this paper. In Section 4, we have introduced an efficient methodology to construct a non-diagonal background error covariance matrix, $\mathbb{B}$. Its inverse, $\mathbb{B}^{-1}$, needed in 3D-Var (5) and in 4D-Var (9) cost function formulations, can be obtained either via a Cholesky decomposition or via a singular value decomposition. (Note that by the "computation of the inverse" we mean the solution of a linear system).

Table 1 illustrates the computational cost of data assimilation compared to the cost of free running model for a 24 hour simulation. All the simulations are performed on a Dell Precision T5400 workstation with 2 quadcore Intel(R) Xeon(R) processors with clock speed 2.33GHz and a RAM of 16GB shared between the two processors. Performing 3D-Var with a non-diagonal error covariance matrix whose inverse is computed by Cholesky decomposition is only about 1.4% more expensive than the 3D-Var with a diagonal covariance matrix. The calculation of the inverse via the Cholesky decomposition is

285 considerably more efficient than the calculation via a singular value decomposition, as expected.

3D-Var framework is built on top of GEOS-Chem v7 package which uses Sparse Matrix Vectorized GEAR (SMVGEAR) solver for chemistry. However, to construct the adjoint of chemistry required by the 4D-Var, we implemented KPP solver (Damian et al., 2002) into GEOS-Chem which not only provides a suite of high performance chemical solvers to choose from but also generates automatically
290 the continuous and discrete adjoint codes (Daescu, 2000, 2003; Sandu et al., 2003). A detailed discussion on interfacing KPP with GEOS-Chem and comparison with native SMVGEAR solver for accuracy and computational performance is presented in (Eller et al., 2009). As pointed out in (Henze et al., 2007), the computational cost of Rosenbrock solver increases significantly with the tolerance levels; higher tolerances use smaller internal time steps requiring more computation. In our experiments, we have
295 set RTOL=10^{-3} and ATOL=10^{-2} to achieve moderate to high accuracy.

The 4D-Var assimilation is considerably more expensive than the 3D-Var. The use of the non-diagonal $\mathbb{B}$ (with Cholesky decomposition) in 4D-Var causes a minimal to zero increase in the computational time when compared to the diagonal $\mathbb{B}$ case.

Table 1. Timing results for GEOS-Chem free model run, 3D-Var and 4D-Var data assimilations with diagonal and non-diagonal $\mathbb{B}$ for a 24 hour simulation starting July 1st, 2006.

Experiment Description	CPU Time
Free model run, SMVGEAR chemistry solver	2 min 50 sec
Free model run, KPP chemistry solver	3 min 18 sec
3D-Var with diagonal $\mathbb{B}$	3 min 57 sec
3D-Var with non-diagonal $\mathbb{B}$, Cholesky	4 min 00 sec
3D-Var with non-diagonal $\mathbb{B}$, SVD	9 min 38 sec
4D-Var with diagonal $\mathbb{B}$ (per model run)	16 min 51 sec
4D-Var with non-diagonal $\mathbb{B}$, Cholesky (per model run)	16 min 51 sec

5.1.2 Tropospheric Emission Spectrometer (TES) observations

300 TES (Beer et al., 2001), one of four science instruments aboard NASA's Aura satellite, measures the infrared-light energy (radiance) emitted by Earth's surface, and by the chemical tracers in the atmosphere (http://tes.jpl.nasa.gov). Vertical profiles of chemical concentrations are retrieved from the radiance measurements using an off-line inversion process. In this work we assimilate the retrieved ozone vertical profiles. Figure 3 shows the location of TES profiles for two days.

305 A-priori information about the vertical concentration profile of the species of interest is needed to solve the retrieval inverse problem (the prior information does not come from the measurement). Let $\mathbf{x}^{\text{prior}}$ be the prior vertical ozone concentration profile (in volume mixing ratio units), and let

$\mathbf{z}^{\text{prior}} = \log \mathbf{x}^{\text{prior}}$. Let $\mathbf{z}^{\text{radiance}}$ (=$\log \mathbf{x}^{\text{true}}$) be the atmospheric profile as resulting directly from the radiances.

The vertical ozone profile retrieval can be expressed according to the formula

$$\widehat{\mathbf{z}} = \mathbf{z}^{\text{prior}} + A\left(\mathbf{z}^{\text{radiance}} - \mathbf{z}^{\text{prior}}\right) + G\eta, \quad \widehat{\mathbf{x}} = \exp(\widehat{\mathbf{z}}). \tag{32}$$

Here A is the averaging kernel matrix, G is the gain matrix, and η is the spectral measurement error (assumed to have mean zero and covariance S_η). More details can be found in (Bowman et al., 2002; Jones et al., 2003; Worden et al., 2004).

The corresponding TES observation operator 3 is linear with respect to the logarithm of the concentrations, but nonlinear with respect to the concentration profile:

$$\mathcal{H}(\mathbf{x}) = \mathbf{z}^{\text{prior}} + A\left(\log(L(\mathbf{x})) - \mathbf{z}^{\text{prior}}\right)$$

where L is an interpolation operator that transforms $\mathbf{x}$ from the GEOS-Chem N-level vertical grid to the TES profile retrieval P-level grid.

For this reason several chemical data assimilation studies based on TES retrieved profiles (Jones et al., 2003; Bowmann et al., 2006; Parrington et al., 2009) have opted to perform the suboptimal Kalman filtering step in the logarithm of the concentrations:

$$\log \mathbf{x}^a = \log \mathbf{x}^f + K\left(\widehat{\mathbf{z}} - \mathcal{H}\left(\mathbf{x}^f\right)\right).$$

For variational data assimilation the forcing calculation is carried out in concentrations. For this reason, an adjoint of the observation operator needs to be derived to update the gradients as described in equations (7) and (11)

$$(\mathcal{H}'(\mathbf{x}))^T \cdot v = \left(\frac{\partial}{\partial \mathbf{x}}(A \log(L(\mathbf{x})))\right)^T \cdot v = \left(\frac{\partial L}{\partial \mathbf{x}}\right)^T \cdot \begin{pmatrix} (L\mathbf{x})_0^{-1} & 0 & \cdots & 0 \\ 0 & (L\mathbf{x})_1^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (L\mathbf{x})_P^{-1} \end{pmatrix} \cdot A^T \cdot v$$

Here, $(\mathcal{H}'(\mathbf{x}))^T$ is a matrix and $v = \mathbb{R}^{-1}\left(\mathcal{H}(\mathbf{x}) - \mathbf{x}^{obs}\right)$. The TES averaging kernel A is usually a non-symmetric matrix, and the result of $A^T \cdot v$ is fed to the interpolation operator to construct the diagonal matrix with the i-th element being $1/(L\mathbf{x})_i$. The term $(\partial L/\partial \mathbf{x})^T$ is the adjoint of the interpolation operator and brings entities from the TES profile retrieval domain back to the GEOS-Chem model domain.

5.1.3 Ozonesonde observations

For validation, we use the ozonesonde profiles measured by the INTEX Ozonesonde Network Study 2006 (IONS-6) (http://croc.gsfc.nasa.gov/intexb/ions06.html (Thompson et al., 2007a, 2007b)) for the

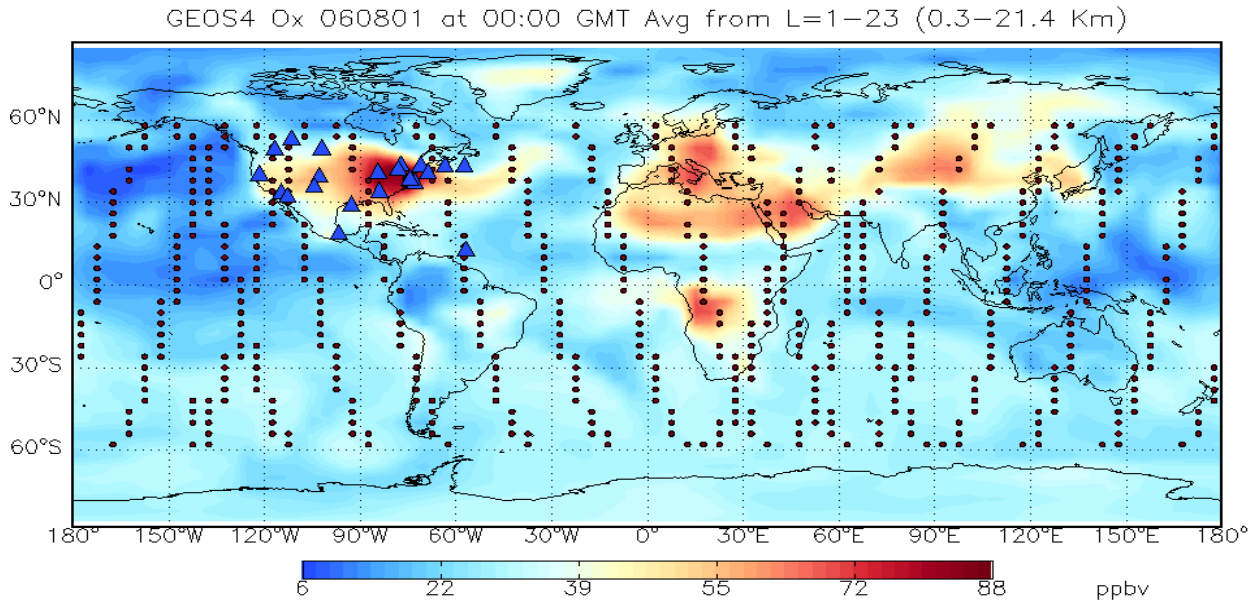


Fig. 3. Ozonesonde sounding stations (triangles) used during IONS06 campaign and AURA/TES satellite trajectory snapshots (dots) plotted over the global ozone distribution on August 1st, 2006.

month of August, assuming that these measurements provide values close to the true state of the atmosphere. The ozonesonde observations are not used in data assimilation, and therefore provide an independent data set against which the analysis results are validated. There are 418 ozonesondes launched from 22 stations across North America as shown in the Figure 3. A detailed description of the number of ozonesondes launched per station with longitude and latitude information can be found in (Parrington et al., 2008).

5.2 Impact of non-diagonal background error covariance in 3D-Var assimilation

In order to demonstrate the benefits of including spatial correlations in the background error covariance matrix, we first compare the tropospheric ozone concentrations generated through 3D-Var analysis against the ozonesonde observations. The left panel of Figure 4 shows the forecast, the analysis and the ozonesonde ozone concentrations averaged over all ozonesonde launches in July and August 2006. The model ozone fields are interpolated to the space-time location of each ozonesonde launch for comparison. The center panel of Figure 4 shows the mean relative errors in model predicted ozone concentrations (the relative differences between the forecast/analysis profiles and the ozonesonde profiles), averaged over all ozonesonde launches. The rightmost panel provides an estimate of the variability of ozonesonde against the variability of ozone concentration predicted through different assimilation

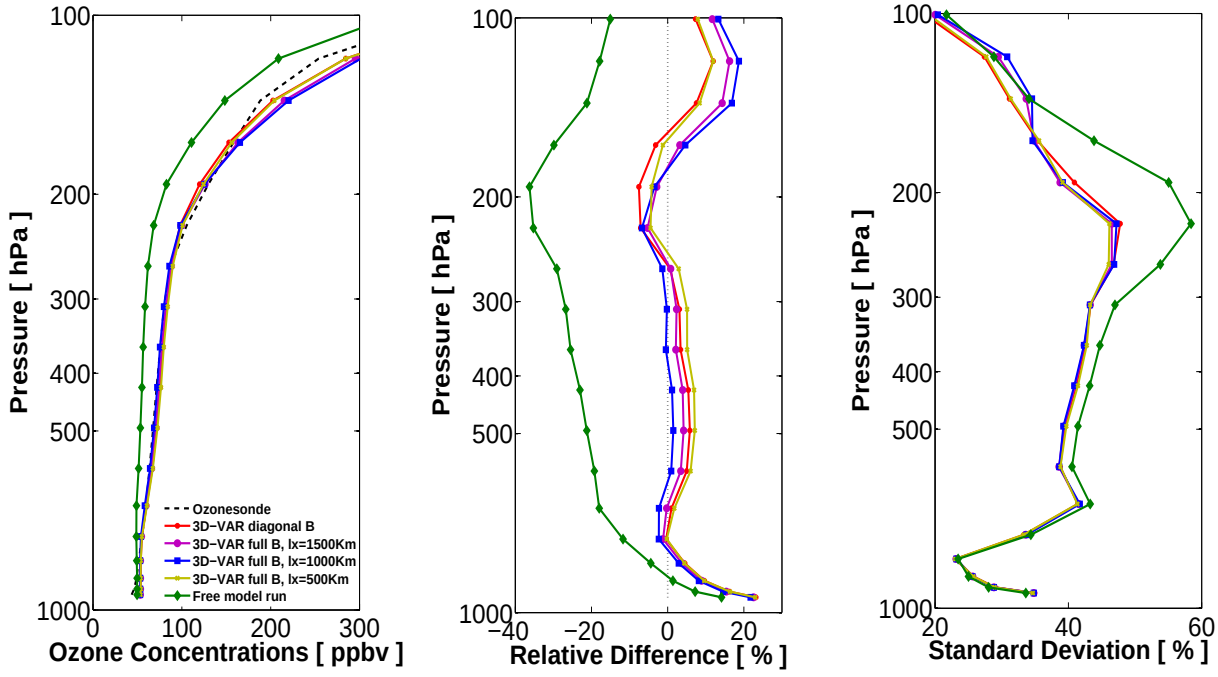
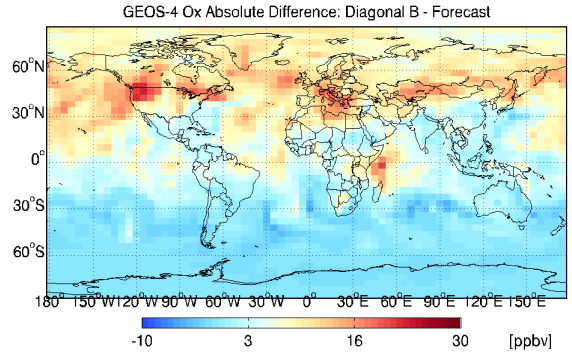


Fig. 4. The impact of non-diagonal background error covariances in 3D-Var data assimilation. Left panel: mean ozone concentrations at ozonesonde locations for 3D-Var analyses and free model trajectories. Center panel: relative mean errors of predicted ozone concentrations with respect to ozonesonde measurements. Right panel: standard deviation of absolute values of errors with respect to ozonesonde measurements. The data is averaged over all ozonesonde launches. These plots were generated from 2 months simulation from 00:00 GMT July 1st to 23:00 GMT August, 2006 and compared against ozonesonde data available for the month of August.

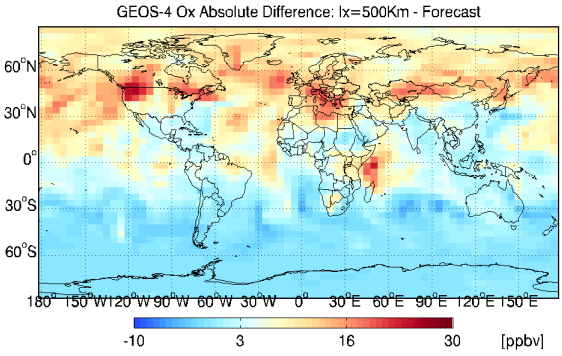
techniques.

In all our experiments, correlation lengths in latitudinal direction varied in proportion with correlation lengths in longitudinal direction. A value of 500 for ℓ_x implicitly indicates ℓ_y is 400, and refers to correlation between two neighboring grid boxes both in East/West and North/South directions.

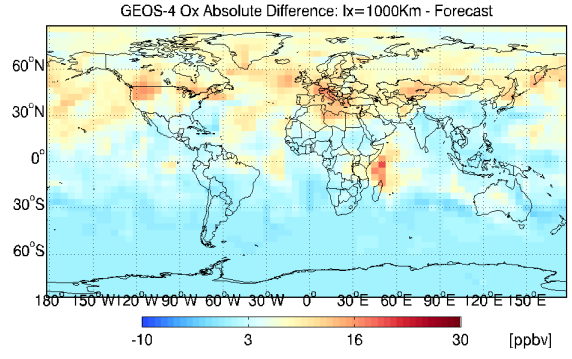
The results indicate that 3D-Var is sensitive to the correlation length used in the construction of the background error covariance matrix (a zero correlation length corresponds to a diagonal matrix). Note that the assimilation results using a non-diagonal $\mathbb{B}$ with higher correlation length are superior to those using a diagonal $\mathbb{B}$ in the lower and mid troposphere. Above 180 hPa, however, the errors in the assimilated ozone fields are larger for the non-diagonal case. This could be attributed to the fact that a uniform correlation length across all vertical levels is only a very coarse approximation of the real error correlations. Higher correlation lengths might be smearing off the ozone in the upper troposphere leading to an overestimate.



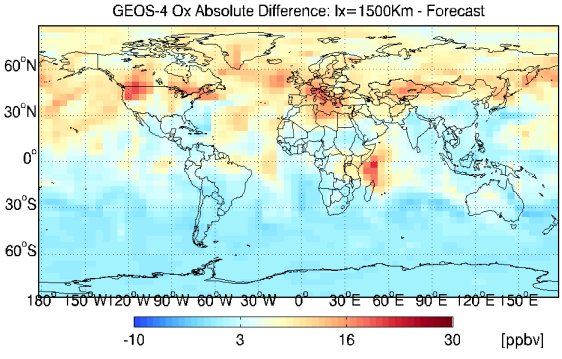
(a) Absolute difference between the 3D-Var analysis using diagonal $\mathbb{B}$ ($\ell_x = 0\,Km$) and the free model run



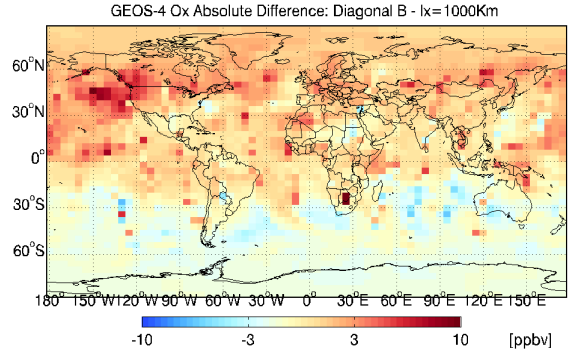
(b) Absolute difference between the 3D-Var analysis using non-diagonal $\mathbb{B}$ ($\ell_x = 500\,Km$) and the free model run



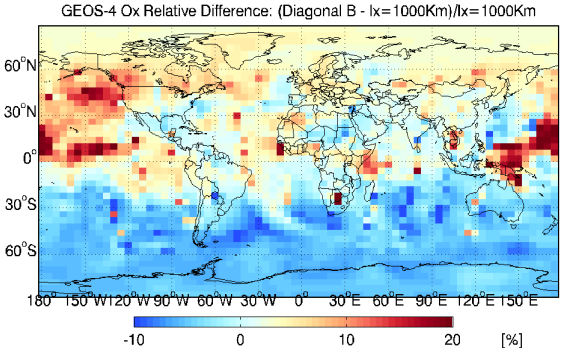
(c) Absolute difference between the 3D-Var analysis using non-diagonal $\mathbb{B}$ ($\ell_x = 1000\,Km$) and the free model run



(d) Absolute difference between the 3D-Var analysis using non-diagonal $\mathbb{B}$ ($\ell_x = 1500\,Km$) and the free model run



(e) Absolute difference between the 3D-Var analyses using diagonal $\mathbb{B}$ and non-diagonal $\mathbb{B}$ ($\ell_x = 1000\,Km$)



(f) Relative difference between the 3D-Var analyses using diagonal $\mathbb{B}$ and non-diagonal $\mathbb{B}$ ($\ell_x = 1000\,Km$)

Fig. 5. Differences in global ozone concentrations at 23:00 GMT on August 31, 2006 averaged over the first 10 GEOS-Chem vertical levels. Panels (a)-(d): differences between the 3D-Var analysis fields and the model forecast (solution without data assimilation); the analyses use different correlation lengths between 0 Km and 1,500 Km. Panels (e)-(f): absolute and relative differences between 3D-Var analyses using diagonal and non-diagonal background covariance matrices.

To further understand the effect of using non-diagonal background error covariance matrices in 3D-Var we consider the corrections obtained with different correlation lengths (i.e., the differences between the assimilated ozone fields and forecast, or the non-assimilated ozone fields). Panels (a)-(d) of Figure 5 show the global spatial distribution plots of these differences. The assimilation with non-
355 diagonal covariance matrices generate much smoother analyses; note that the point-wise values of the increments is smaller, and that the corrections are distributed over larger areas. Panels (e)-(f) of Figure 5 compare directly the 3D-Var analyses obtained using a diagonal $\mathbb{B}$ and a non-diagonal $\mathbb{B}$ with a correlation length of 1000 Km. The corrections in the non-diagonal case are smoother.

5.3 Impact of non-diagonal background error covariance in 4D-Var assimilation

360 We now study the effects of using non-diagonal background error covariance matrix in 4D-Var data assimilation. We compare the analyses ozone concentrations generated by 4D-Var with different background error correlation lengths against the ozonesonde observations. The left panel of Figure 6 shows the forecast, analysis and ozonesonde measured ozone concentrations averaged over the two months assimilation window. The model ozone fields are interpolated to the space-time location of
365 each ozonesonde launch for comparison. The center panel shows the relative errors of model predictions with respect to ozonesonde data, averaged over all ozonesonde launches. The right panel provides the standard deviations of these errors.

The results indicate that 4D-Var is also sensitive to the structure of the background error covariance matrix. The use of non-diagonal correlations can lead to improved analyses. The best analysis is
370 obtained with a correlation length of 500 Km (about one grid cell near the equator). Note that 4D-Var accounts for all the data available within the assimilation window.

To better understand the impact of different background error correlation lengths on 4D-Var assimilation, we present in Figure 7 the differences in ozone concentrations generated by the free model run and by the 4D-Var assimilation using diagonal and non-diagonal $\mathbb{B}$. The use of a non-diagonal $\mathbb{B}$
375 with a properly-chosen correlation length not only provides a better estimate but also helps generate a smoother analysis. The panels (e)-(f) of Figure 7 compare directly the 4D-Var analyses obtained using a diagonal $\mathbb{B}$ and a non-diagonal $\mathbb{B}$ with a correlation length of 500 Km; the large localized corrections over North America provided by the diagonal $\mathbb{B}$ are smoothened out when the non-diagonal $\mathbb{B}$ is employed.

380 5.4 Determining the correlation length through experiments

The correlation length is a very important parameter that impacts the quality of the assimilation when using non-diagonal error covariance matrix. The value of the correlation length depends on various factors such as the lifetime of the tracer under consideration, the grid resolution, the pressure level, and the wind velocity. We propose a method to determine experimentally a value of the correlation

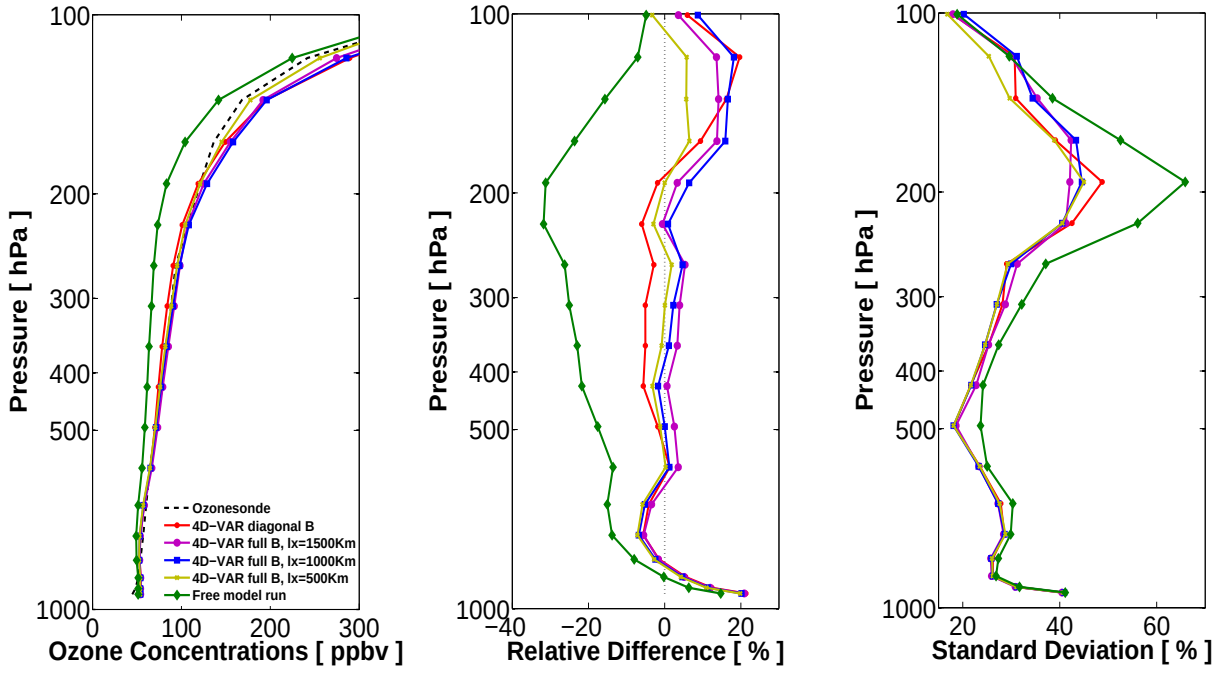
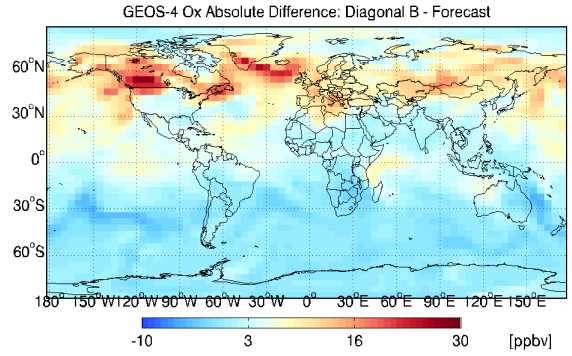


Fig. 6. The impact of non-diagonal background error covariances on 4D-Var data assimilation. The results shown are for a single 5-day assimilation window from 00:00 GMT August 1st to 00:00 GMT August 6th, 2006. Left panel: mean ozone concentrations at ozonesonde locations for 4D-Var analyses and free model trajectories. Center panel: relative mean errors of predicted ozone concentrations with respect to ozonesonde measurements. Right panel: standard deviation of absolute values of errors with respect to ozonesonde measurements.

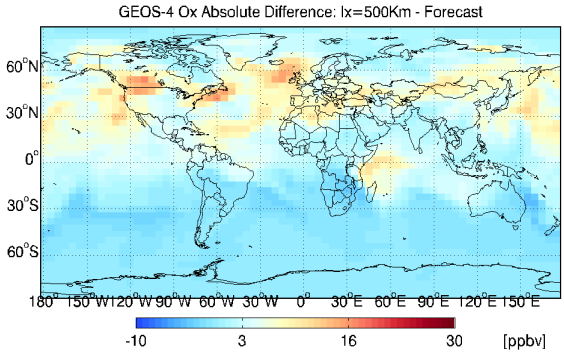
length that is appropriate for the model and data at hand.

Recall the construction of one dimensional correlation matrices in equations (13) and (14). Our aim is to determine the number of grid cells (in each direction) where the errors are correlated. For this we use adjoint sensitivity analysis. Specifically, we initialize the adjoint variable to 1 in a specific cell at the end of a given window (and to zero everywhere else), perform a backwards adjoint simulation, and analyze the adjoint fields at the beginning of the window. The error in the specific cell is correlated with errors in those grid cells where the adjoint values are above $1/e$. The length of the time window depends on the time scale of the model under consideration.

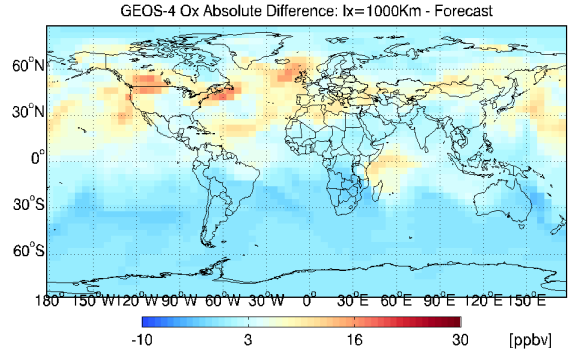
Here we consider a time window of 8 hours. We run the forward GEOS-Chem model starting at July 1st, 00:00 GMT for 20 hours. The adjoint variable for ozone at 20:00 GMT are initialized to 1 in a subset of the grid points $((i,j)$ chosen such that $i \mod 10 = 1$ and $j \mod 10 = 1)$. Adjoint variables for all other grid points and species are initialized to zero. The gap in the initialization helps avoid the interactions between adjoint "plumes" initialized at different locations. The ozone adjoint variable



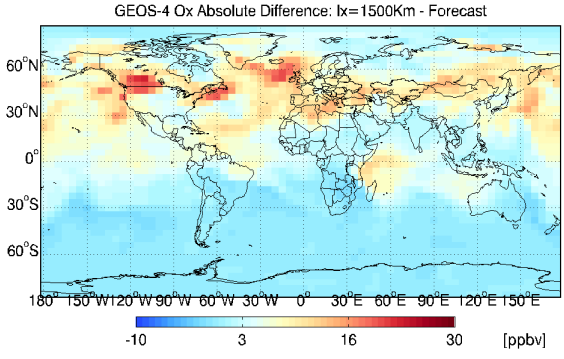
(a) Absolute difference between the 4D-Var analysis using diagonal $\mathbb{B}$ ($\ell_x = 0\,Km$) and the free model run



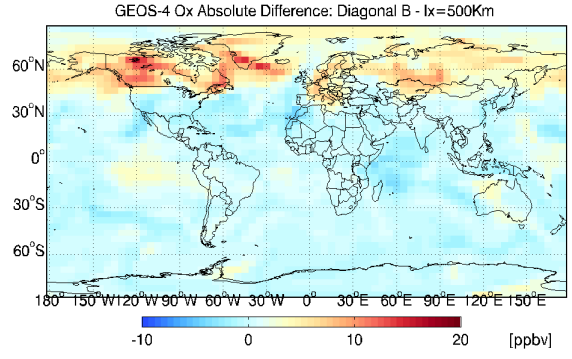
(b) Absolute difference between the 4D-Var analysis using non-diagonal $\mathbb{B}$ ($\ell_x = 500\,Km$) and the free model run



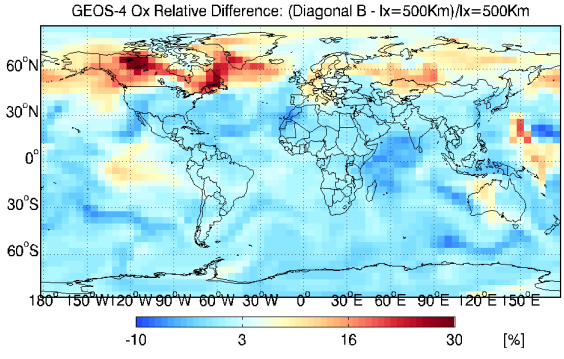
(c) Absolute difference between the 4D-Var analysis using non-diagonal $\mathbb{B}$ ($\ell_x = 1000\,Km$) and the free model run



(d) Absolute difference between the 4D-Var analysis using non-diagonal $\mathbb{B}$ ($\ell_x = 1500\,Km$) and the free model run

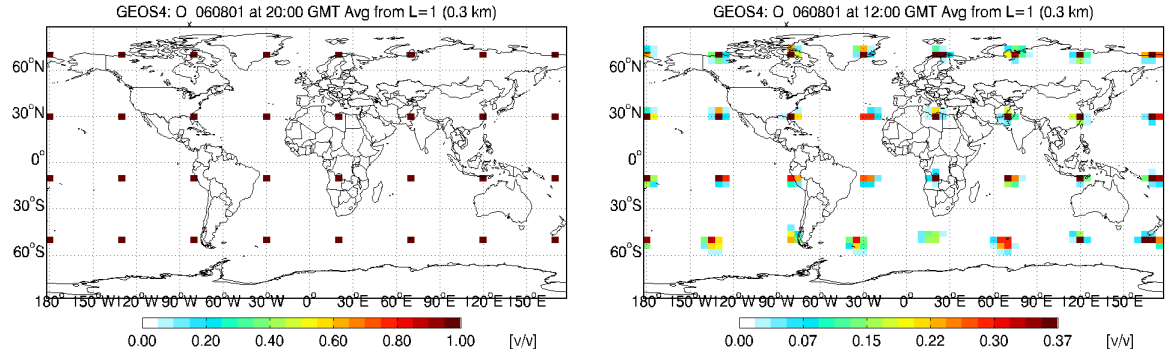


(e) Absolute difference between the 4D-Var analyses using diagonal $\mathbb{B}$ and non-diagonal $\mathbb{B}$ ($\ell_x = 500\,Km$)



(f) Relative difference between the 4D-Var analyses using diagonal $\mathbb{B}$ and non-diagonal $\mathbb{B}$ ($\ell_x = 500\,Km$)

Fig. 7. Differences in global ozone concentrations at 00:00 GMT on August 06, 2006 (end of assimilation window) averaged over the first 10 GEOS-Chem vertical levels. Panels (a)-(d): differences between the 4D-Var analysis fields and the model forecast (solution without data assimilation); the analyses use different correlation lengths between $0\,Km$ and $1{,}500\,Km$. Panels (e)-(f): absolute and relative differences between 4D-Var analyses using diagonal and non-diagonal background covariance matrices.



(a) Adjoint ozone variables are initialized to one in selected grid cells

(b) Adjoint ozone variables spread after 8 hours of backward sensitivity run

Fig. 8. Ground level ozone adjoint variable values are initialized to one on July 1st 2006, 20:00 GMT, every tenth grid point in longitudinal and latitudinal directions. An 8 hour backward adjoint integration spreads the adjoint fields, and helps identify grid cells where ozone errors are correlated.

field is analyzed at 12:00 GMT to find out the number of grid cells where the value is greater than or equal to $1/e$. In our current setup, we use the same correlation length for all pressure levels and thus consider the spread only at ground level.

The procedure can be easily extended to considering different correlation lengths for different vertical levels and for different geographic areas.

Figure 8 shows that the ozone adjoint variables have spread (on average) over one to two grid cells in both longitudinal and latitudinal direction. For the $4^\circ \times 5^\circ$ model resolution each grid box is of size $400Km \times 500Km$ near the equator. The adjoint sensitivity analysis indicates that the correlation lengths should be chosen in the ranges of $\ell_x \in [500Km, 1000Km]$ and $\ell_y \in [400Km, 800Km]$ respectively. This confirms the best correlation lengths empirically observed in the data assimilation results reported in Figures 4 and 6.

6 Conclusions

This paper presents an efficient methodology to construct non-diagonal background error covariance matrices for data assimilation. The two- or three- dimensional covariance matrices are not formed explicitly. Rather, multi-dimensional correlations are represented by tensor products of one dimensional correlation matrices along longitudinal and latitudinal directions. The technique can be easily extended to include correlations in the vertical direction as well. Highly efficient linear algebra operations are obtained by performing successive matrix-vector products, Cholesky decompositions, etc. with one-dimensional correlation matrices. The correlation lengths are important parameters that need to be specified for each directional correlation. The performance of the data assimilation system depends on

a correct specification of the correlation lengths. We propose an adjoint sensitivity analysis approach to guide the choice of proper correlation lengths; the approach implicitly accounts for factors such as chemical activity, grid resolution, etc.

The proposed methodology can model non-isotropic error correlations, since different correlation lengths can be used for different directions. Moreover, different correlation lengths can be used in different parts of the computational domain. The covariances can be recomputed periodically (e.g., at the beginning of each assimilation window) after re-evaluating the correlation lengths estimates. The model does not account for correlations across different variables, e.g., for correlations among error in different chemical species.

The approach to construct non-diagonal covariance matrices has been tested using the 3D-Var and 4D-Var data assimilation frameworks developed for GEOS-Chem. The experiments assimilate observations from TES satellite ozone profile retrievals, and validate the results against an independent data set provided by IONS ozonesondes. The change of the covariance matrix formulation in data assimilation from diagonal to non-diagonal adds only a negligible computational overhead. At the same time, the inclusion of spatial correlations with appropriately chosen correlation lengths leads consistently to improved analyses in both the 3D-Var and the 4D-Var settings.

Future work will compare the proposed background covariances against other modeling methods available in the literature. We will also consider the vertical error correlations which can help spread the information from higher altitudes, where the TES sensitivity is greater, to lower altitudes, where accurate ozone predictions are of interest from an air quality perspective.

Acknowledgements. This work has been supported by NASA through the ROSES-2005 AIST project. The authors would like to thank Mark Parrington and Paul Hamer for providing processed TES data and for their help with the data visualization script.

References

Akella, S., and Navon, I. M.: Different approaches to model error formulation in 4D-Var: a study with high resolution advection schemes, Tellus A , Vol 61A, 112Ű128, 2009.

Beer, R., Glavich, T. A., and Rider, D. M.: Tropospheric emission spectrometer for the Earth Observing System's Aura satellite, Appl. Opt., 40(15), 2356-2367, 2001.

Bey, I., Jacob, D. J. , Yantosca, R. M., Logan, J. A., Field, B., Fiore, A. M., Li, Q., Liu, H., Mickley, L. J. and Schultz, M.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, J. Geophys. Res., 106, 23, 073-23,096, 2001.

Bowman, K. W., Worden, J., Steck, T., Worden, H. M., Clough, S. and Rodgers, C.: Capturing time and vertical variability of tropospheric ozone: A study using TES nadir retrievals, J. Geophys. Res.,107, (D23), 2002.

Bowman, K. W., Rodgers, C. D., et al: Tropospheric Emission Spectrometer: Retrieval method and error analysis, IEEE Transactions on Geoscience and Remote Sensing, vol. 44, no. 5, May 2006.

Bowman, K. W., Jones, D. B. A., Logan, J. A., Worden, H., Boersma, F., Kulawik, S., Osterman, G., Worden, J. and Chang, R.: Impact of surface emissions to the zonal variability of tropical ozone and carbon monoxide for November 2004, Atmos. Chem. Phys. Disc., 8, 1505–1548, 2008.

Carmichael, G. R., Sandu, A., Chai, T., Daescu, D., Constantinescu, E. M. and Tang, Y.: Predicting air quality: Improvements through advanced methods to integrate models and measurements, Journal of Computational Physics, Vol. 227, Issue 7, p. 3540–3571, 2008.

Chai, T., Carmichael, G. R., Sandu, A., Tang, Y. and Daescu, D. N.: Chemical data assimilation of transport and chemical evolution over the Pacific (TRACE-P) aircraft measurements, Journal of Geophysical Research, 111, D02301, doi:10.1029/2005JD005883, 2006.

Chai, T., Carmichael, G. R., Tang, Y., Sandu, A., Hardesty, M., Pilewskie, P., Whitlow, S., Browell, E. V., Avery, M. A., Thouret, V., Nedelec, P., Merrill, J. T. and Thomson, A. M.: Four dimensional data assimilation experiments with ICARTT (International Consortium for Atmospheric Transport and Transformation) ozone measurements, Journal of Geophysical Research, Vol. 112, D12S15, doi:10.1029/2006JD007763, 2007.

Chai, T., Carmichael, G. R., Tang, Y. and Sandu, A.: Regional NO_X emission inversion through a four-dimensional variational approach using SCIAMACHY tropospheric NO_2 column observations, Atmospheric Environment, doi:10.1016/j.atmosenv.2009.06.052, in print, 2009.

Clark, H. L., Cathala, M. -L., Teysse'dre, H., Cammas, J. -P. and Peuch.,V. -H.: Cross-tropopause fluxes of ozone using assimilation of MOZAIC observations in a global CTM, Tellus, Ser. A and Ser. B, 59B, 39-49, 2006.

Constantinescu, E. M., Chai, T., Sandu, A. and Carmichael, G. R.: Autoregressive models of background errors for chemical data assimilation, Journal of Geophysical Research, Vol. 112, D12309, doi:10.1029/2006JD008103, 2007.

Constantinescu, E. M., Sandu, A., Chai, T. and Carmichael, G. R.: Investigation of ensemble-based chemical data assimilation in an idealized setting, Atmospheric Environment, Vol. 41, Issue 1, p. 18–36, 2007.

Constantinescu, E. M., Sandu, A., Chai, T. and Carmichael, G. R.: Ensemble-based chemical data assimilation. I: general approach, Quarterly Journal of the Royal Meteorological Society, Volume 133, Issue 626, p. 1229–1243, Online ISSN: 1477-870X, Print ISSN: 0035-9009, July 2007 Part A.

Constantinescu, E. M., Sandu, A., Chai, T. and Carmichael, G. R.: Ensemble-based chemical data assimilation. II: covariance localization, Quarterly Journal of the Royal Meteorological Society, Volume 133, Issue 626, p.

1245–1256, Online ISSN: 1477-870X, Print ISSN: 0035-9009, July 2007 Part A.

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

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M.: The ECMWF implementation of three-dimensional variational assimilation (3D-Var) I: Formulation, Quarterly Journal of the Royal Meteorological Society, 124(550):1783, 1998.

Daescu, D., Carmichael, G.R., and Sandu, A.: Adjoint Implementation of Rosenbrock Methods Applied to Variational Data Assimilation Problems, Comp. Phys, 165, 496-510, 2000.

Daescu, D., Sandu, A., and Carmichael, G.R.: Direct and Adjoint Sensitivity Analysis of Chemical Kinetic Systems with KPP: II - Validation and Numerical Experiments, Atmos. Environ., 37, 5097-5114, 2003.

Daescu, D.N.: On the sensitivity equations of four-dimensional variational (4D-Var) data assimilation, Monthly Weather Review, 136 (8), 3050-3065, 2008.

Daley, R.: Atmospheric Data Analysis, Cambridge University Press, p. 457pp, 1991.

Damian, V., Sandu, A., Damian, M., Potra, F., and Carmichael,G.R.: The Kinetic PreProcessor KPP - A Software Environment for Solving Chemical Kinetics, Comp. and Chem. Eng., 26, 11, 1567-1579, 2002.

Derber, J. C. and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. Tellus, 51 A, 195Ű221.

Elbern, H., and Schmidt, H.: A four-dimensional variational chemistry data assimilation scheme for eulerian chemistry transport modeling. J. Geophys. Res., 104, 18583-18598, 1999.

Elbern, H., and Schmidt, H.: Ozone episode analysis by four dimensional variational chemistry data assimilation, J. Geophys. Res., 106(D4), 3569–3590, 2001.

Eller, P., Singh, K., Sandu, A., Bowman, K. W., Henze, D. K. and Lee, M.: Implementation and evaluation of an array of chemical solvers in a global chemical transport model, Geophysical Model Development, Vol. 2, p. 1–7, 2009.

Fisher, M. and Lary, D. J., Lagrangian four-dimensional variational data assimilation of chemical species, Q. Journal of the Royal Met. Society, 121, 1681-1704, 1995.

Fisher, M.: Background Error Covariance Modeling. Proc. ECMWF Seminar on recent developments in data assimilation for atmosphere and ocean, pp 45-64, 8-12 September, 2003.

Fisher, M.: ECMWF Data Assimilation Training Course: Background Error Covariance Modelling, 2006.

Gaspari, G., Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological Society, Vol 125 Issue 554,723-757, 1999.

Hakami, A., Henze, D. K., Seinfeld, J. H., Chai, T., Tang, Y., Carmichael, G. R. and Sandu, A.: Adjoint inverse modeling of black carbon during ACE-Asia, Journal of Geophysical Research, Vol. 110, D14301, doi:10.1029/2004JD005671, 25 pages, 2005.

Henze, D. K., Hakami, A. and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmos. Chem. Phys., 7, 2413-2433, 2007.

Henze, D. K., Seinfeld, J. H. and Shindell, D. T.: Inverse modeling and mapping U.S. air quality influences of inorganic PM2.5 precursor emissions with the adjoint of GEOS-Chem, Atmos. Chem. Phys., 9, 5877-5903, 2009.

Hollingsworth, A. and LŽnnberg, P.: The statistical structure of short-range forecast errors as determined from

radiosonde data. Part I: The wind field. Tellus, 38A, 111-136, 1986.

Ingleby, B.: The statistical structure of forecast errors and its representation in the Met Office Global Model, Q. Journal of Royal Meteorol. Soc., 124, 1783–1807, 2001.

Jones, D. B. A., Bowman, K. W., Logan, J. A., Heald, C. L., Liu, J., Luo, M., Worden, J. and Drummond, J.: Inversion analysis of carbon monoxide emissions using data from the TES and MOPITT satellite instruments, Atmospheric Chemistry and Physics Discussions,7,6, 17625–17662, 2007.

Jones, D. B. A., Bowman, K. W., Palmer, P. I., Worden, J. R., Jacob, D. J., Hoffman, R. N., Bey, I. and Yantosca, R. M.: Potential of observations from the Tropospheric Emission Spectrometer to constrain continental sources of carbon monoxide, J. Geophys. Res., 108, (D24), 2003.

Kalnay, E.: Atmospheric modeling, data assimilation and predictability, Cambridge University Press, 2002.

Khattatov, B. V., Gille, J. C., Lyjak, L. V., Brasseur, G. P., Dvortsov, V. L., Roche, A. E. and Walters, J.: Assimilation of photochemically active species and a case analysis of UARS data, Journal of Geophysical Research, 104:18715–18737, 1999.

Khattatov, B. V., Lamarque, J. -F., Lyjak, L. V., Menard, R., Levelt, P., Tie, X., Brasseur, G. P. and Gille, J. C.: Assimilation of satellite observations of long-lived chemical species in global chemistry transport models, J. Geophys. Res., 105(D23), 29–135, 2000.

Kopacz, M., Jacob, D. J., Henze, D. K., Heald, C. L., Streets, D. G. and Zhang, Q.: A comparison of analytical and adjoint Bayesian inversion methods for constraining Asian sources of CO using satellite (MOPITT) measurements of CO columns, J. Geophys. Res., doi:0.1029/2007JD009264, 2009.

Lahoz, W. A., et al.: The Assimilation of Envisat data (ASSET) project, Atmos. Chem. Phys., 7, 1773–1796, 2007.

Lamarque, J.-F., Khattatov, B. V. and Gille, J. C.: Constraining tropospheric ozone column through data assimilation, J. Geophys. Res., 107(D22), 4651, doi:10.1029/2001JD001249, 2002.

LeDimet, F.-X. and Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, Tellus 38A, 97–110, 1986.

Lions, J. L.: Optimal control of systems governed by partial differential equations, Springer-Verlag, 1971.

Menard, R., Cohn, S. E., Chang, L. -P. and Lyster, P. M.: Assimilation of stratospheric chemical tracer observations using a Kalman Filter I: Formulation, Mon. Weather Rev., 128, 2654-2671, 2000.

Nassar, R., Logan, J. A., Worden, H. M., et al.: Validation of Tropospheric Emission Spectrometer (TES) nadir ozone profiles using ozonesonde measurements,J. Geophys. Res., 113, D15S17, doi:10.1029/2007JD008819, 2008.

Navon, I. M.: Data assimilation for Numerical Weather Prediction: a review, in: Data Assimilation for Atmospheric, Oceanic, and Hydrologic Applications, XVIII, 475 p. 326 illus., Hardcover, ISBN: 978-3-540-71055-4, 2009.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A.V., Kostelich, E. J., Kostelich, M., Corazza, M., Sauer, T., Kalnay, E., Patil, D. J. and Yorke, J. A: A local ensemble Kalman Filter for Atmospheric Data Assimilation, Tellus, Vol. 56A, pp. 415-428, 2004.

Palmer, P. I., Jacob, D. J., Jones, D. B. A., Heald, C. L., Yantosca, R. M., Logan, J. A., Sachse, G. W. and Streets, D. G.: Observations over the western Pacific, 2003.

Parrington, M., Jones, D. B. A., Bowman, K. W., Horowitz, L. W., Thompson, A. M., Tarasick, D. W. and Witte, J. C.: Estimating the summertime tropospheric ozone distribution over North America through assimilation of

observations from the Tropospheric Emission Spectrometer, Journal of Geophysical Research, Vol 113, D18307, doi:10.1029/2007JD009341, 2008.

Parrington, M., Jones, D. B. A., Bowman, K. W., Thompson, A. M., Tarasick, D. W., Merrill, J., Oltmans, S. J., Leblanc, T., Witte, J. C. and Millet, D. B.: Impact of the assimilation of ozone from the tropospheric emission spectrometer on surface ozone across North America, Geophysical Research Letters 36 (4), 2009.

Parrish, D. F. and Derber, J. C.: The national meteorological center's spectral statistical-interpolation analysis system, Monthly Weather Review, (120), p. 1747–1763, 1992.

Pierce, R. B., et al.: Chemical data assimilation estimates of continental U. S. ozone and nitrogen budgets during the Intercontinental Chemical Transport Experiment-North America, J. Geophys. Res., 112, D12S21, doi:10.1029/2006JD007722, 2007.

Purser, R. J., Wu, W.-S., Parrish, D. F., and Roberts, N. M.: Numerical aspects of the application of recursive filters to variational statistical analysis. Advanced methods of 4-D Var Data Assimilation in NWP models Part II: Spatially inhomogeneous and anisotropic general covariances. Mon. Wea. Rev., 131, 1536-1548, 2003.

Rabier, F., Jarvinen, H., Klinker, E., Mahfouf, J. -F. and Simmons, A.: The ECMWF operational implementation of four-dimensional variational data assimilation I: Experimental results with simplified physics, Quarterly Journal of the Royal Meteorological Society, 126:1143–1170, 2000.

Sandu, A., Daescu, D., and Carmichael, G.R.: Direct and Adjoint Sensitivity Analysis of Chemical Kinetic Systems with KPP: I - Theory and Software Tools, Atmos. Environ., 37, 5083-5096, 2003.

Sandu, A., Daescu, D. N., Carmichael, G. R. and Chai, T.: Adjoint sensitivity analysis of regional air quality models, Journal of Computational Physics, Vol. 204, p. 222–252, 2005.

Sandu A. and Zhang L.: Discrete second order adjoints in atmospheric chemical transport modeling, Journal of Computational Physics, 227 (12), 5949–5983, 2008.

Sasaki, Y. K.: An objective analysis based on the variational method, J. Met. Soc. Jap. II(36), 77–88, 1958.

Segers, A. J., Eskes, H. J., van der A, R. J., van Oss, R. F. and van Velthoven, P. F. J.: Assimilation of GOME ozone profiles and a global chemistry-transport model, using a Kalman Filter with anisotropic covariance, Quarterly Journal of the Royal Meteorological Society, 131, 477-502, 2005.

Singh, K., Eller, P., Sandu, A., Bowman, K. W., Jones, D. B. A. and Lee, M.: Improving GEOS-Chem model forecasts through profile retrievals from Tropospheric Emission Spectrometer, in: Lecture Notes on Computational Science vol. 5545, p. 302–311, International Conference on Computational Science 2009, Baton Rouge, Louisiana, May 25–27, 2009.

Singh, K., Eller, P., Sandu, A., Henze, D., Bowman, K. W., Kopacz, M. and Lee, M.: Towards the construction of a standard adjoint GEOS-Chem model, High Performance Computing Symposium (HPC 2009) at Spring Simulation Multiconference (SpringSim'09), San Diego, California, March 22–27, 2009.

TES Science Team, TES L2 Data Users Guide, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California. (Available at http://tes.jpl.nasa.gov/docsLinks/DOCUMENTS/TESL2DataUsersGuidev2.0.pdf), 2006.

Thompson, A. M., et al.(2007a): Intercontinental chemical transport experiment ozonesonde network study (IONS) 2004: 1. Summertime upper troposphere/lower stratosphere ozone over northeastern North America, J. Geophys. Res., 112 D12S12, doi:10.1029/2006JD007441, 2007. Thompson, A. M., et al.(2007b): Intercontinental chemi-

cal transport experiment ozonesonde network study (IONS) 2004: 2. Tropospheric ozone budgets and variability over northeastern North America, J. Geophys. Res., 112 D12S13, doi:10.1029/2004JD005359, 2007.

Worden, J. R., Bowman, K. W. and Jones, D. B. A.: Characterization of atmospheric profile retrievals from Limb Sounding Observations of an inhomogeneous atmosphere, J. Quant. Spectrosc. Radiat. Transfer, 86, (03)00274-7, 2004.

Wu, L., Mallet, V., Bocquet, M. and Sportisse, B.: A comparison study of data assimilation algorithms for ozone forecasts, J. Geophys. Res., 113, D20310, 2008.

Zhang, L., Constantinescu, E. M., Sandu, A., Tang, Y., Chai, T., Carmichael, G. R., Byun, D., Olaguer, E.: An adjoint sensitivity analysis and 4D-Var data assimilation study of Texas air quality, Atmospheric Environment, Vol. 42, Issue 23, p. 5787–5804, 2008.

Zhu, C., Byrd, R. H. and Nocedal, J.: L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization, ACM Transactions on Mathematical Software, Vol 23, Num. 4, pp. 550 - 560, 1997.