Quantifying the loss of information in source attribution problems using the adjoint method in global models of atmospheric chemical transport.

MAURICIO SANTILLANA, *

Harvard University School of Engineering and Applied Sciences, Cambridge, Massachusetts

ABSTRACT

It is of crucial importance to be able to identify the location of atmospheric pollution sources in our planet. Global models of atmospheric transport in combination with diverse Earth observing systems are a natural choice to achieve this goal. It is shown that the ability to successfully reconstruct the location and magnitude of an instantaneous source in global chemical transport models (CTMs) decreases rapidly as a function of the time interval between the pollution release and the observation time. A simple way to quantitatively characterize this phenomenon is proposed based on the effective –undesired– numerical diffusion present in current Eulerian CTMs and verified using idealized numerical experiments. The approach presented consists of using the adjoint-based optimization method in a state-of-the-art CTM, GEOS-Chem, to reconstruct the location and magnitude of a realistic pollution plume for multiple time scales. The findings obtained from these numerical experiments suggest a time scale of 2 days after which the accuracy of the adjoint-based optimization methodology is compromised considerably in current global CTMs. In conjunction with the mean atmospheric velocity, the aforementioned time scale leads to an estimate of a length scale of about 1700 km, downwind from the source, beyond which measurements, in conjunction with current global CTMs, may not be successfully utilized to reconstruct continuous-in-time sources. The approach presented here can be utilized to characterize the capabilities and limitations of adjoint-based optimization inversions in other regional and global Eulerian CTMs.

1. Introduction

Understanding and quantifying the fate of anthropogenic and natural emissions of chemically active gases in the atmosphere is an important endeavour in our changing climate. In particular, many efforts have been directed to the quantification of large-scale spatial and temporal variations of sources and sinks of gases such as carbon dioxide, methane and nitrous oxide, due to their direct role in changing the radiative properties of the atmosphere, see Kasibhatla et al. (2000); Enting (2002); Ciais et al. (2010) and the multiple references therein. Having reliable estimates of these sources and sinks is of crucial relevance in the evaluation of global policies such as the Kyoto protocol, designed to curb emissions of green-house gases.

Global chemical transport models (CTMs) have been designed to simulate the dynamics of the concentration fields of chemicals under the influence of atmospheric transport and chemical reactions. In conjunction with Earth observing systems such as satellite retrievals and monitoring stations, CTMs provide a natural modeling framework to estimate the strength and location of such chemical sources in a top-down fashion. In this context, CTMs are expected to have the capability of identifying pollution sources within the time scales of relevance to global atmospheric chemical transport events: weeks, for inter-continental transport of pollutants, and months to a year, for inter-hemispheric transport. It is shown in this study that modeling frameworks that use the adjoint-based optimization method to identify the strength and location of pollution sources, as implemented in current CTMs, may not achieve their goal in relevant time scales.

Our ability to reconstruct a pollution source using global CTMs depends mainly on three factors: (i) the quality of our measurement systems, (ii) the appropriateness of our mathematical model and our inversion approach, and (iii) our ability to numerically approximate the mathematical model and the inversion approach using computers (Enting 2002). In this study, it is assumed that a perfect observation system is in place (no noise) at every location in the atmosphere, thus ignoring errors coming from (i), and it is also assumed that the mathematical model (equations [1]) approximates well the dynamics of the atmosphere, thus ignoring the issue of the appropriateness of the mathematical model in (ii). Specifically, the efficacy of the adjoint-based optimization methodology is studied in a practical and realistic computational framework (iii).
Eulerian global CTMs simulate the chemical composition of the atmosphere by numerically solving a set of coupled partial differential equations of the form,

$$\frac{\partial C_i}{\partial t} + \mathbf{u} \cdot \nabla C_i = R_i + s_i(x, t)$$  \hspace{1cm} (1)

where $C_i$ is the concentration of chemical species $i$, $\mathbf{u}$ is the wind velocity field (obtained from a global atmospheric circulation model), $R_i$ is the effective chemical production rate (typically a function of the mass fractions of other chemicals), and $s_i$ describes local emissions and non-chemical sinks.

A source attribution problem in a global CTM can be stated as follows, given a set of observations of the state of the atmosphere (typically concentration fields) and atmospheric wind fields for a time period $T = t_f - t_0 > 0$, find the best emission and deposition rates, $s_i(x, t)$, for the time period $T$, such that when they are utilized as parameters in the CTM, the simulated chemical concentrations $C_i(t)$ are consistent with observations. Finding the solution (the best emission and deposition rates) to this inverse problem is achieved by minimizing the misfit between observations and model results, frequently using gradient-based iterative optimization methods. The adjoint-based optimization method studied here, is a widely accepted methodology to calculate the gradients needed to identify the direction of steepest descent in the minimization of a cost function that typically quantifies the misfit between model simulations and measurements. This methodology has been widely utilized in the reconstruction of emission sources using global CTMs, see for example (Henze et al. 2007; Zhang et al. 2009; Kopacz et al. 2009, 2010; Zhang et al. 2011; Wecht et al. 2011).

This paper is organized as follows, the idealized source attribution problem is introduced in section 2, the adjoint-based optimization method, as well as the technical details of its numerical implementation are presented in section 3. The numerical results of our computational experiments along with their implications in real-world applications, including data assimilation methodologies, are discussed in section 4.

2. A source attribution problem

An idealized accident-type release source attribution problem for an inert gas is studied here in order to evaluate the efficacy of the adjoint-based methodology to reconstruct local-in-time and space perturbations. Thus, from here on $R_i = 0$ in equations (1). An observation system capable of providing exact (synthetic) observations of the atmosphere everywhere at time $t_f$ (obtained directly from a forward simulations using the CTM) is assumed. The problem consists of finding the best field of instantaneous emissions at time $t = t_0$, $s_i(x, t_0)$, such that the misfit between observations and the simulation results, at time $t_f$, is minimized. The problem is simplified by placing the instantaneous release at a vertical height of approximately 4 km above the ground at time $t = 0$. This choice minimizes the effects of planetary boundary layer mixing processes, as represented (by sub-grid parametrizations) in global CTM (Lin et al. 2008). By having the same number of observations, $p$ (all grid boxes in the domain at $t = t_f$), as the number of discrete instantaneous sources to be de-
terminated, $N$ (all grid boxes in the domain at $t = 0$), the usual ill-posed nature of the inversion problem caused by the under-determination of parameters when $p \ll N$ in real-life Earth observations systems, is removed (Bocquet 2005b[a]). This fact, in combination with the numerical stability provided by the use of the numerical approximation of the continuous adjoint (see section 3) as a means to calculate the gradients in the minimization process, makes it possible for the idealized experiments under examination to be solved without the need of prior information about the plume or a regularization technique (Bocquet 2005b[a]).

The aforementioned problem is solved numerically, using a three dimensional state-of-the-art CTM, GEOS-Chem (Bey et al. 2001). GEOS-Chem is a state-of-the-art CTM, driven by GEOS-5 analyzed meteorological data from the NASA Global Modeling and Assimilation Office (GMAO). The GEOS-Chem adjoint-based optimization was developed by (Henze et al. 2007). In order to produce synthetic observations, an inert tracer plume is propagated for different simulation times in GEOS-Chem, with $4^\circ \times 5^\circ$ and $2^\circ \times 2.5^\circ$ horizontal resolutions, using the high-order advection scheme native of GEOS-Chem (Lin and Rood 1996). The true initial plume concentration was set to be twice that of the constant global background in all other grid cells, with a horizontal extent of approximately $1330 \text{ km} \times 1680 \text{ km}$ ($3 \times 3$ grid cells), and a vertical extent of one vertical pressure level at 4km (20th pressure level) above the ground. The horizontal location of the initial plume is shown in Figures 1 and 2 (Top left plot). Through successive iterations, the adjoint-based optimization methodology generates an “optimized” initial condition estimate. The maximum number of iterations was set to 99 to keep our experiments within the practical limit set by our computational resources. We performed six numerical experiments for each of the two spatial resolutions, assimilating data after 3, 12, 24, 48, 96, and 168 hours.

Since any continuous emission field in time can be thought of as a sum of instantaneous releases, the structural results of our numerical investigation are valid in constant emission flux reconstructions and more generally, in data assimilation frameworks that use Eulerian CTMs and the adjoint-based optimization method to estimate source locations and/or magnitudes.

The idealized numerical problem studied here is called an “identical twin” experiment in the inverse modeling community (Kaipio and Somersalo 2004), since synthetic observations are produced using the same forward model utilized in the inversion approach. It is well known that assessing the performance of inverse modeling strategies using an identical twin experiment leads to optimistic results and constitutes what the inverse modeling community calls an “inverse crime” (Kaipio and Somersalo 2004). Indeed, in identical twin experiments, the exact solution of the minimization problem lives in the parameter space (the space of initial conditions in our specific problem) of the modeling framework, and thus, with an appropriate methodology one should be able to recover the optimal solution. From this perspective, the idealized numerical experiments studied here assess only the algorithmical capabilities and limitations of the adjoint-based inversion methodology, as implemented in CTMs, and can be interpreted as optimistic since the further complications of real-life source reconstructions are ignored.

3. The adjoint-based optimization method

Inverse modeling and data assimilation approaches involve finding an optimal set of model parameters that best match observations (Enting 2002; Bennet 2002; Wunnsch 2006). In order to find this optimal set of parameters (in our source attribution problem we find the best initial condition), one minimizes the distance or misfit between the model results and observations. This minimization can be performed using a gradient based approach. As shown in (Talagrand and Courtier 1987) the “adjoint equations” of the model can be used to compute explicitly the functional derivative (or the gradient) of the distance function, between observations and model results, with respect to the initial conditions. The computation of one gradient requires one forward-in-time integration of the full model equations over the time interval on which the observations are available, followed by one backward-in-time integration of the adjoint equations. Subsequent gradients required for a steepest descent algorithm in a minimization routine are computed in a similar fashion.

When $R_i = 0$ in equations 1, the forward model becomes a linear transport problem with sources and sinks. The adjoint operator of the transport operator is again a linear transport operator with reversed winds and reversed time integration (Bocquet 2005[a]; Hourdin et al. 2006). A subtle point arises when implementing the adjoint-based optimization method numerically, since one faces the dilemma of whether using the adjoint of the discrete approximation to the transport operator or the discretization of the continuous adjoint (Vukicevic et al. 2001; Bocquet 2005[a]; Hourdin et al. 2006; Liu and Sandu 2008). This issue has been previously studied not only in the context of atmospheric chemistry modeling (Gou and Sandu 2011), but in the context of oceanography and weather modeling (Sirkes and Tziperman 1997). While the adjoint of the discrete approximation of the transport may produce better point-wise results in sensitivity analyses, it can be very unstable numerically and potentially lead to non-physical sensitivity results, specially when using stabilized high-order advect-
Fig. 1. Area of influence at \( \sim 4 \) km of altitude (20th pressure level) as a function of simulation time, as reconstructed by the first iteration of the adjoint-based optimization method for \( 4^\circ \times 5^\circ \) numerical simulations. Latitude vs Longitude.

Fig. 2. Area of influence at \( \sim 4 \) km of altitude (20th pressure level) as a function of simulation time, as reconstructed by the first iteration of the adjoint-based optimization method for \( 2^\circ \times 2.5^\circ \) numerical simulations. Latitude vs Longitude.

Numerical errors in transport routines of Eulerian global CTMs have been shown to be significant, compromising the quality of simulation results, at spatial resolutions larger than \( 1^\circ \times 1^\circ \) even with high-order advection schemes (Wild and Prather, 2006). Moreover, the decay and spatial broadening observed in Eulerian global CTM simulations of realistic pollution plumes signal the presence of excessive numerical diffusion. This undesired numerical diffusion is shown to be much higher than expected by the order of the advection scheme, and shown to be nearly governed by the theoretical upper limit, set by the value of the local finite-time Lyapunov exponent, almost independently of the grid-box size, in Rastigeyev et al. (2010) for currently practical spatial resolutions. As a consequence, high values of numerical diffusion are observed in regions with high values of finite-time Lyapunov exponents, i.e. in regions where wind shear is high and where chaotic mixing (and thus where loss of information) is pronounced. This excessive numerical diffusion will eventually compromise our ability to recover the plume information (location and magnitude) in the multiple iterations of the adjoint-based optimization process.

Some amount of diffusion is to be expected in any atmospheric transport simulation. Indeed, turbulent mixing, consequence of the multi-scale chaotic velocity field in the atmosphere, can be thought of and is often parametrized as an eddy diffusion term in meso-scale Eulerian atmospheric simulations and represented by stochastic perturbations on the velocity field in Lagrangian global CTMs (Gifford 1982; Sillman et al. 1990; Mauzerall et al. 1998; Pisso et al. 2009). Studies combining Lagrangian trajectories and aircraft measurements of actual passive tracer plumes have placed the values of this eddy diffusivity around \( 10^4 \) m\(^2\) s\(^{-1}\) in the horizontal, and \( 0.3 - 1 \) m\(^2\) s\(^{-1}\) in the vertical (Pisso et al. 2009). In Eulerian global CTMs, however, the values of the “effective” horizontal numerical diffusion are estimated to be approximately \( 10^5 \) m\(^2\) s\(^{-1}\) (Pisso et al., 2009), an order of magnitude larger than those estimated by aircraft studies. Specifically, this horizontal numerical
The effective (numerical) equations governing the dynamics of a pollution plume in global Eulerian CTMs can be incorporated in the mathematical model\(^1\) as:

\[
\frac{\partial C_i}{\partial t} + \mathbf{u} \cdot \nabla (C_i) = \nabla \cdot (\mathbf{D}_h \nabla (C_i)) + R_i + s_i(x, t) \tag{2}
\]

where the “effective” numerical diffusivity matrix \(\mathbf{D}_h\) is, in fact, a function of the grid size, \(\Delta x\), the time step, \(\Delta t\), the characteristics of the flow field given by \(\mathbf{u}(x, t)\), and the spatial extent of the pollution plume with respect to the grid size (LeVeque, 2002; Majda and P. 1999; Rastigeyev et al., 2010).

4. Numerical Results

A gradual process to investigate the impact of the effective numerical diffusion in the solution of the aforementioned idealized source attribution problems was designed. First, one iteration of the adjoint-based optimization was performed in order to identify the geographic area of influence, \(i.e.\) where information came from, as calculated by the numerical advection scheme in our numerical experiments for multiple time-scales. Since no prior knowledge of the pollution plume was assumed, the first meaningful iteration of the adjoint-based optimization consists of utilising the resulting concentration field of the forwardly advected idealized plume as an initial condition, and integrate it back-in-time using the advection scheme, with reversed winds, for exactly the same time period \(t_f - t_0 = t_f\). In the absence of numerical errors, and due to the linearity of the transport operator in equations\(^1\), this procedure should reconstruct the location and magnitude of the original plume at time \(t = 0\).

a. Geographic region of influence and loss of information

The obtained results from this single iteration of the adjoint-based optimization, displayed in Figures\(^1\) and\(^2\) for spatial resolutions \(4^\circ \times 5^\circ\) and \(2^\circ \times 2.5^\circ\) respectively, show the numerical reconstruction of the geographic region of influence, \(A_h(t_0)\), also called the numerical sensitivity region, at \(\sim 4\) km altitude (at time \(t = 0\)). Note that for a simulation time \(t_f = 3h\), the numerical region of influence coincides with the original region of the plume in both spatial resolutions. As the simulation time \(t_f\) (time between release and observations) increases, the numerical sensitivity region broadens more and more (the boundary of this geographic region was drawn where the reconstructed source was 100 times smaller than the original excess concentration). The broadening –a direct consequence of the numerical diffusion– is slightly anisotropic reflecting the underlying structure of the atmospheric winds characterized by the jet stream in mid-latitudes. Using estimates of the local finite-time Lyapunov exponent for mid-latitudes in GEOS-Chem from Rastigeyev et al. (2010), for \(4^\circ \times 5^\circ\) and \(2^\circ \times 2.5^\circ\) spatial resolutions, we can infer approximate values of \(D_h \sim 3.5 \times 10^6\) m\(^2\) s\(^{-1}\) and \(D_h \sim 1.4 \times 10^6\) m\(^2\) s\(^{-1}\) for the numerical horizontal diffusion for each resolution respectively, in the chosen geographic location of the idealized plume.

Assuming horizontal isotropic diffusion only, an estimate of the “diffusive deformation” of the numerical region of influence can be calculated by adding the length scale \(\sqrt{D_h t}\) to each the boundary lengths of the plume at time \(t\). We can thus expect that the area of the numerical region of influence will have increased its size nearly four-fold in \(4^\circ \times 5^\circ\) simulations and three-fold in \(2^\circ \times 2.5^\circ\), after a day of forward simulation (and a day of back integration), since the tracer concentration will have spread \(\sim 1000\) km on each of the four boundaries of the plume. As shown in the Sensitivity Area (1 day) panel of Figures\(^1\) and\(^2\) this estimate captures well the behaviour of the numerical results. For two-days and four-days forward simulation times, this estimate suggests a 4.5-fold and 6-fold (for \(4^\circ \times 5^\circ\) simulations) and 4-fold and 5.4-fold (for \(2^\circ \times 2.5^\circ\) simulations) broadening of the numerically reconstructed plume, respectively. In fact, the introduced deformation estimates of the area of the numerical region of influence appear to be a lower limit of the actual one shown in our numerical experiments in the panels of Figure\(^1\). The disagreement, however, is not significant. Moreover, it is shown in the next section that this estimates of the –undesired– numerical spatial deformation of the original plume can be used to estimate the loss of information of the full adjoint-based optimization.

While this estimate of the deformation of the numerical sensitivity region, using only (spatially and time averaged) horizontal isotropic broadening, may appear incomplete (local finite-time Lyapunov exponent values, from which we estimated the spatially and time averaged diffusivity constant in the diagonal of \(\mathbf{D}_h\), change as a function of time and region of space), it seems appropriate when cal-

\(^1\)For this estimate we assumed that the decay rate for a spatial resolution of \(4^\circ \times 5^\circ\) is governed by the Lyapunov exponent \(\lambda\), thus \(\alpha \approx 9 \times 10^{-5}\) sec\(^{-1}\), the plume width \(W \sim 10^6\) m, and the characteristic length scale where the concentration of the plume decays to zero \(r_b \approx \Delta x = 4 \times 10^2\) m.
c. Estimation of the loss of information

A simple calculation to estimate of the relative error observed in the adjoint-based source reconstructions is proposed here based on the horizontal deformation of the numerical region of influence, due to undesired numerical diffusive processes, as a function of time. Estimates of the
Fig. 5. Normalized, zonally-integrated reconstructed plumes as a function of simulation time, as reconstructed by the adjoint-based optimization method for 4° × 5° numerical simulations. Latitude vs Pressure level (atm)

The quality of the gradient information provided by the adjoint operator is more accurate, as discussed before, for shorter simulations. As a consequence, the BFGS gradient-based minimization routine utilized in GEOS-Chem (Zhu et al., 1994), is capable of producing excellent results by decreasing the normalized value of the cost function by 15 orders of magnitude for the 4° × 5°, 3 hr simulations, and nearly by 14 orders of magnitude for the 2° × 2.5°, 3 hr simulations. This is shown in Figures 10 and 11. As the simulation time increases though, the gradient information loses quality rapidly and leads to significantly poorer performances in the minimization processes. Indeed, for 7-day simulations, the minimization only achieves an 8-order of magnitude reduction of the normalized cost function, leading to very poor reconstructed plumes.

The center of mass and total mass relative errors were calculated, between the reconstructed plume and the true plume, for each spatial resolution and simulation period, in order to provide a more complete characterization of the numerical errors. The results are shown in Figures 8 and 9. For these calculations, an error in the center of mass equal to the maximal radius of the plume (1680 km) was considered a 100% relative error. Errors above 100% are observed in simulations longer that 2 days. The total mass relative error does not reach very high values, for 7-day simulations these are of the order of 20 – 30%, as shown in
Fig. 7. Relative $L^2$-error (%) between the reconstructed initial condition and the true solution as a function of the simulation time for both $4^\circ \times 5^\circ$ and $2^\circ \times 2.5^\circ$ spatial resolutions. The estimate on the loss of information is obtained assuming that information is lost proportionally to the excess area of the numerical area of influence when compared to the true area of influence.

Figure 8 shows a source is about 1700 km, equivalent to 3 grid boxes in a $4^\circ \times 5^\circ$ simulation. This means that if a continuous source is to be successfully reconstructed using the adjoint-based optimization method in current global CTMs, then a continuous observation system, closer than 3 grid cells (in a $4^\circ \times 5^\circ$ simulation), should be in place down wind from the source. Moreover, the aforementioned length scale naturally suggests a spatial density over the globe below which sources may not be accurately reconstructed in global Eulerian CTMs in practical spatial resolutions. Depending on the tracer, this spatial density of observation sites may or may not be realistic with current Earth observing systems.

5. Conclusions

It is shown that the ability to reconstruct the location and magnitude of an accident-type, instantaneous inert pollution plume, using the adjoint-based optimization method in a global Eulerian CTM, decreases rapidly as the time between release and observations increases. The obtained numerical results suggest a time scale of 2 days after which significant numerical errors ($> 70\%$) compromise the ability of the adjoint-based optimization method to track information accurately back to the pollution source. This time scale is shorter than the time scales of inter-continental transport for example, for which it would be desirable to have a reliable methodology to reconstruct sources.

It is shown that the quality of the gradient utilized to
to minimize the misfit between observations and simulation results, used by the adjoint-based optimization method, decays fast as a function of simulation time. As a consequence of this fact, a simple way to quantitatively characterize the loss of information in the adjoint-based optimization is proposed based on the –undesired– deformation of the numerical area of influence caused by the effective diffusion in the CTM (which is much larger than expected by the order of the advection scheme due to the chaotic nature of atmospheric flow). This characterization of the loss of information is shown to successfully describe the behavior of the numerical relative error observed between the true solution and the numerical reconstructions, as a function of simulation time.

The approach presented in this study exhibits a structural (algorithmical) limitation also present in data assimilation systems utilizing the adjoint to obtain the gradient for successive optimization iterations in source attributions problems. Since identical twin experiments were utilized to evaluate the performance of the adjoint-based optimization methodology, and measurement errors and model errors were ignored, one should expect the results obtained here to be optimistic in the context of real-life source attribution problems.

The use of regularization techniques as well as a dense observation network near the source(s) to be reconstructed may help counter balance this structural numerical limitation of the adjoint method, as it is currently implemented in global CTMs. Inversion approaches not relying on gradient information should be more intensely studied in this context (Kaipio and Somersalo 2004).

Acknowledgments.

REFERENCES

Bennet, A., 2002: Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press.

Bey, I., et al., 2001: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. Journal of Geophysical Research, 106, 23,073–23,096.

Bocquet, M., 2005a: Grid resolution dependence in the reconstruction of an atmospheric tracer source. Nonlin. Processes Geophys., 12, 219–233.

Bocquet, M., 2005b: Reconstruction of an atmospheric tracer source using the principle of maximum entropy. i:theory. Q. J. R. Meteorol. Soc., 131, 2191–2208.

Ciais, P., P. Rayner, F. Chevallier, P. Bousquet, M. Logan, P. Peylin, and M. Ramonet, 2010: Atmospheric inversions for estimating CO2 fluxes: methods and perspectives. Climatic Change, 103, 69–92.

Enting, I., 2002: Inverse Problems in Atmospheric Constituent Transport. Cambridge University Press.
Fig. 11. Behavior of the cost function as a function of the number of iterations for multiple simulations times for $2^2 \times 2.5^2$ spatial resolution. Logarithm of the normalized cost function vs number of iterations.

Gifford, F., 1982: Horizontal diffusion in the atmosphere: A lagrangian dynamical theory. *Atmos. Environ.*, 16, 502–512.

Gou, T. and A. Sandu, 2011: Continuous versus discrete advection adjoints in chemical data assimilation with cmaq. *Atmos. Environ*.

Henze, D., A. Hakami, and J. H. Seinfeld, 2007: Development of the adjoint of geochem. *Atmos. Chem. Phys.*, 7, 2413–2433.

Henze, D., J. H. Seinfeld, and D. T. Shindell, 2009: Inverse modeling and mapping u.s. air quality influences of inorganic pm2.5 precursor emissions with the adjoint of geochem. *Atmos. Chem. Phys.*, 9, 5877–5903.

Hourdin, F., O. Talagrand, and A. Idelkadi, 2006: Eulerian backtracking of atmospheric tracers. ii: Numerical aspects. *Quarterly Journal of the Royal Meteorological Society*, 132, 585–603.

Kaipio, J. and E. Somersalo, 2004: *Statistical and Computational Inverse Problems*. Springer.

Kasibhatla, P., M. Heimann, P. Rayner, N. Mahowald, R. G. Prinn, and D. E. H. (Eds), 2000: Inverse methods in global biogeochemical cycles. *Geophys. Monogr. Series*, 114, 324.

Kopacz, M., D. J. Jacob, D. K. Henze, C. L. Heald, D. G. Streets, and Q. Zhang, 2009: 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.*, 114.

Kopacz, M., et al., 2010: Global estimates of co sources with high resolution by adjoint inversion of multiple satellite datasets (mopitt, airs, sciamachy, tes). *Atmos. Chem. Phys.*, 10, 855–876.

LeVeque, R., 2002: *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics.

Lin, J., D. Youn, X. Z. Liang, and D. J. Wuebbles, 2008: Global model simulation of summertime u.s. ozone diurnal cycle and its sensitivity to pbl mixing, spatial resolution, and emissions. *Atmos. Environ*.

Lin, S. and R. Rood, 1996: Multidimensional flux-form semi-lagrangian transport schemes. *Monthly Weather Review*, 124, 20462070.

Liu, Z. and A. Sandu, 2008: Analysis of discrete adjoints of numerical methods for the advection equation. *International Journal for Numerical Methods in Fluids*, 56, 769–803.

Majda, A. and K. P., 1999: Simplified models for turbulent diffusion: Theory, numerical modelling, and physical phenomena. *Physics Reports*, 314, 237–574.

Mauzerall, D. L., et al., 1998: Photochemistry in biomass burning plumes and implications for tropospheric ozone over the tropical south atlantic. *Journal of Geophysical Research: Atmospheres*, 103, 8401–8423.

Pisso, I., E. Real, K. S. Law, B. Legras, N. Bousserez, J. L. Attié, and H. Schlager, 2009: Estimation of mixing in the troposphere from lagrangian trace gas reconstructions during long-range pollution plume transport. *J. Geophys. Res.*, 114.

Rastigeyev, Y., R. Park, M. Brenner, and D. Jacob, 2010: Resolving intercontinental pollution plumes in global models of atmospheric transport. *J. Geophys. Res.*, 115.

Sillman, S., L. Logan, and S. Wofsy, 1990: A regional scale model for ozone in the united states with subgrid representation of urban and power plant plumes. *J. Geophys. Res.*, 95, 57315748.

Sirkes, Z. and E. Tziperman, 1997: Finite difference of adjoint or adjoint of finite difference? *Mon. Wea. Rev.*, 125, 33733378.

Talagrand, O. and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. i: Theory. *Q. J. R. Meteorol. Soc.*, 133, 1311–1328.
Vukicevic, T., M. Steyskal, and M. Hecht, 2001: Properties of advection algorithms in the context of variational data assimilation. *Mon. Weather Rev.*

Wecht, K. J., et al., 2011: Validation of tes methane with hippo aircraft observations: implications for inverse modeling of methane sources. *Atmos. Chem. Phys. Discuss.*, **11**, 27887–27911.

Wunsch, C., 2006: *Discrete Inverse and State Estimation Problems*. Cambridge University Press.

Zhang, L., D. J. Jacob, M. Kopacz, D. K. Henze, K. Singh, and D. A. Jaffe, 2009: Intercontinental source attribution of ozone pollution at western u.s. sites using an adjoint method. *Geophys. Res. Lett.*, **36**.

Zhang, L., et al., 2011: Improved estimate of the policy-relevant background ozone in the united states using the geos-chem global model with 1/2 degrees x 2/3 degrees horizontal resolution over north america. *Atmospheric Environment*, **45**, 6769–6776.

Zhu, C., R. Byrd, P. Lu, and J. Nocedal, 1994: Lbfgs-b: a limited memory fortran code for *Tech. rep. Northwestern University*. 

