JOURNAL OF GEOPHYSICAL RESEARCH, VOL. ???, XXXX, DOI:10.1029/,

Inverse Modeling of Aerosol Dynamics: Condensational Growth

D. K. Henze, J. H. Seinfeld

Department of Chemical Engineering, California Institute of Technology,

Pasadena, CA, USA

W. Liao, A. Sandu,

Department of Computer Science, Virginia Tech, Blacksburg, VA, USA

G. R. Carmichael

Center for Global and Regional Environmental Research, University of Iowa,

Iowa City, IA, USA

J. H. Seinfeld, Department of Chemical Engineering, Caltech, Pasadena, CA 91125, USA. (seinfeld@caltech.edu)

Abstract. The feasibility of inverse modeling a multicomponent, size resolved aerosol evolving by condensation / evaporation is investigated. The adjoint method is applied to the multicomponent aerosol dynamic equation in a box model (zero-dimensional) framework. Both continuous and discrete formulations of the model (the forward equation) and the adjoint are considered. A test example is studied in which the initial aerosol size-composition distribution and the pure component vapor concentrations (i.e. vapor pressures) are estimated based upon measurements of all species, or a subset of the species, and the entire size distribution, or a portion of the size distribution. It is found that the adjoint method can successfully retrieve the initial size distribution and the pure component vapor concentrations even when only a subset of the species or a portion of the size distribution is observed. The results presented here provide a basis for the inverse modeling of aerosols in three-dimensional atmospheric chemical transport models.

1. Introduction

In recent years, data assimilation techniques have been used to increase one's ability to predict and characterize atmospheric chemical phenomena by providing valuable estimates of surface emissions, improved model sensitivities, and optimized measurement strategies. By enforcing closure between model predictions and experimental observations, these methods constrain the variance of chemical transport models (CTMs) to produce optimal representations of the state of the atmosphere. As the number of variables used to describe the state of the atmosphere increases, the process of integrating models and measurements becomes increasingly difficult. Fortunately, advances in algorithm efficiency, computational resources, and the theory of inverse modeling have facilitated extension of these techniques to systems of increasing complexity. Anticipating the point at which all main features of sophisticated atmospheric CTMs are endowed with an inverse, this work examines the possibilities of extending data assimilation studies to include explicit consideration of size and composition aerosol dynamics.

Although the actual implementation of data assimilation methods can be quite different, in general all techniques utilize some observational data set to provide an improved model representation of the system in question. Many previous studies on inverse modeling have utilized the Kalman filter, wherein propagation of the error covariance matrix is used to retain consistency between the model and the measurements. As the model integrates forwards in time, the entire model state is sequentially updated at each observation. Using data from satellites, this method has been employed to assimilate quantities such as aerosol optical depth [*Collins et al.*, 2001], column CO [*Lamarque et al.*, 1999], CO [*Clerbaux*

et al., 2001], CH_4 [*Lyster et al.*, 1997], total and partial ozone column [*Stajner et al.*, 2001], ozone measured from the Upper Atmosphere Research Satellite/Microwave Limb Sounder [*Levelt et al.*, 1998; *Khattatov et al.*, 2000], and tropospheric ozone [*Lamarque et al.*, 2002]. Recent studies by *Palmer et al.* [2003a, b] have used a similar theoretical approach in tandem with aircraft observations to provide estimates of Eastern Asian CO emissions and correlated anthropogenic halocarbon emissions.

At the fundamental level, for perfect linear models and a single observation, the basic Kalman filter technique is equivalent to another set of data assimilation algorithms categorized as variational calculations [*Lorenc*, 1986]. Despite their underlying similarities, when applied to nonlinear models, the Kalman filter and variational approaches become quite different. While using a Kalman filter has the distinct advantage that model error is explicitly included in the analysis, the large computational cost of this approach has historically been the prime motivation for development of alternative methods.

Variational data assimilation yields optimized estimates of model parameters (physical constants, inputs and initial / boundary conditions), χ, that minimize the discrepancy between model output and experimental observations as measured by a scalar valued cost function, $\mathcal{J}$. The entire evolution of the solution state is optimized to match the observations. Such calculations require knowledge of the gradient of the cost function with respect to the set of all variable parameters, $\nabla_\chi \mathcal{J}$. While the first applications of the variational method to simple meteorological cases appeared more than 30 years ago [*Sasaki*, 1970; *Thompson*, 1969], extension of this method to complex transport models, even in the absence of chemistry, was not feasible prior to development of efficient methods for calculation of $\nabla_\chi \mathcal{J}$.

Originating from the mathematics of systems optimization and control theory [*Cacuci*, 1981a, b], the adjoint method was first suggested for variational data assimilations in atmospheric transport models by *Marchuk* [1974] as an improved technique for calculating $\nabla_\chi \mathcal{J}$. One of the first applications in this field was the estimation of atmospheric diffusion coefficients [*Lamb et al.*, 1975]. The adjoint method uses a single backward integration of the model (with the state variable during the backward integration being the derivative of the cost function with respect to the original forward state variables) from the final time to the initial conditions in order to determine all elements of the gradient simultaneously. Compared to forward sensitivity analysis [*Hoffman*, 1986], in which the gradient is determined by consecutively propagating perturbations of each parameter individually through the model, the dependence of the calculation's complexity on the number of variable parameters is greatly reduced [*Talagrand and Courtier*, 1987]. Not only does this approach afford application to detailed models, it also facilitates the simultaneous estimation of large numbers of parameters. A drawback to the adjoint approach is that for nonlinear problems, trajectories from the forward integration must be available for the backward integration. This leads to large storage requirements; however, multiple level checkpointing schemes can be implemented to reduce this demand.

Applied to problems in fluid mechanics as early as 1974 [*Pironneau*, 1974], followed somewhat later by applications in meteorology [*Talagrand and Courtier*, 1987] and oceanography [*Tziperman and Thacker*, 1989], the adjoint method has only been applied to CTMs relatively recently. *Fisher and Lary* [1995] presented the first implementation of the adjoint method in a Four-Dimensional Variational (4-D Var) chemical data assimilation wherein they recovered the initial concentrations of stratospheric gases (O_3, NO_2)

over the assimilation period. Additional 4-D adjoint models for stratospheric chemical data assimilation have been developed by *Errera and Fonteyn* [2001], while tropospheric data assimilations have been performed by *Elbern et al.* [1997] and *Elbern and Schmidt* [1999, 2001]. *Sandu et al.* [submitted 2003] performed a regional sensitivity study using the adjoint method. Progress has also been made in developing software that automatically generates efficient adjoint models of atmospheric chemistry modules [*Sandu et al.*, 2003; *Daescu et al.*, 2003]. A nice review of the fundamentals of developing adjoints of 4-D adjoint models for chemical data assimilation routines is given by *Wang et al.* [2001].

The inclusion of detailed aerosol chemistry and physics has become requisite in atmospheric CTMs. Future implementation of 4D-Var assimilation techniques will likewise require the inclusion of aerosols in the adjoint models. To lay the groundwork for this endeavor, the fundamental capabilities (and limitations) of applying such techniques to aerosols need to be investigated. In this paper, we apply the first inverse models of multicomponent aerosol dynamics, and evaluate their performance under conditions designed to facilitate incorporation of these routines into existing adjoint CTMs. A paper presenting derivations of the necessary equations for several other forms of inverse aerosol models, and evaluation of these for a simple, single component aerosol has also been submitted [*Sandu et al.*, submitted 2004]. These works differ substantially from the only previous data assimilation study involving aerosols [*Collins et al.*, 2001] in that the aerosol distribution is allowed to evolve according to the aerosol dynamic equation [*Pilinis*, 1990] and that the inversion is performed using the adjoint technique. In the study by *Collins et al*, the aerosols were represented as growing via empirical correlations and growth rates, and the total aerosol optical depth was assimilated sequentially using a Kalman filter.

With the above goal in mind, adjoint aerosol models are developed and are tested using simulated observations (commonly known as an identical twin experiment). The (forward) aerosol model used is a simplified, yet numerically and physically consistent, version of the aerosol submodel currently employed in several 4-D CTMs [*Meng et al.*, 1998; *Song and Carmichael*, 2001]. As operator splitting is used in such models to isolate all aerosol processes into a single 0-D (box) routine which is called within each cell of the discretized 3-D spatial field, it is sufficient to use a forward box model that does not include gas-phase chemistry or spatial advection. Within this forward box model, emphasis is placed on gas-to-particle conversion, wherein gas-phase transport is the rate-limiting step for particle growth. The details of the forward model are given in Section 2.

An immediate application of an inverse aerosol model is to infer the size distributions of aerosol sources using surface, airborne, or possibly even satellite measurements. This involves reconstructing back trajectories of the distribution by repetitive calls to the adjoint box model from within the overall adjoint 4-D CTM, asking each time to recover the shape of the distribution at a previous time step. Therefore, an important capability of the aerosol adjoint routine is to recover an initial size distribution based upon knowledge of the distribution at some later time(s). The length of the assimilation period will depend upon the temporal resolution of the forward model and the frequency of the observations; herein we consider periods ranging from several minutes to a few hours.

In addition to recovering initial distributions, an inverse aerosol model can be used to estimate physical properties key to the dynamic evolution of the distribution by treating these quantities as variable parameters. The growth of aerosol particles due to condensation / evaporation is heavily influenced by the thermodynamic properties of the transfer-

ring species. A significant fraction of organic aerosol particles are comprised of chemical compounds whose thermodynamic properties in the particulate phase are not well characterized. Better estimates of such properties would not only increase the accuracy of CTMs, but would also aid in interpretation of laboratory studies of aerosol dynamics. Hence another desired capability of an adjoint aerosol model is to provide estimates of the thermodynamic properties of the aerosol species.

The aerosol adjoint models can also help refine experimental measurement strategies. Conditions can be simulated in which either individual species are not measured, or the size distribution is only partially sampled. Comparison of the assimilations between these scenarios leads to sampling schemes that provide an optimum balance between data recoverability and observational burden.

One of the primary reasons for choosing the adjoint method to construct an inverse aerosol model is the computational efficiency of this approach. As variations in the actual implementation of this methodology affect the overall computational requirements, it is beneficial to consider different approaches to constructing the adjoint models, of which there are two generally recognized types—continuous and discrete [*Giles and Pierce*, 2000; *Tziperman and Thacker*, 1989]. The first method is to derive the continuous adjoint equations from the governing partial differential equations, and then solve these numerically. The second approach is to cast the forward equations into a numerical discretized form, and then take the adjoint of this discretized formula. Numerical discretization and adjoint operations do not commute in general, therefore the continuous and discrete approaches lead to final gradients that differ in accuracy and computational expense; hence, it is

advantageous to assess both tactics when introducing the adjoint method to a new field [*Sandu et al.*, submitted 2004].

2. Multicomponent Gas-to-Particle Conversion (the Forward Model)

We consider a multicomponent aerosol that is growing/evaporating as a result of gas-to-particle conversion. The continuous governing equation for a 0-D, multicomponent, internally mixed aerosol distribution is then [*Pilinis*, 1990; *Meng et al.*, 1998]

$$\frac{\partial p_i(\mu,t)}{\partial t} = H_i(\mu, p_1, \ldots, p_n, t)p(\mu,t) - \frac{1}{3}\frac{\partial (Hp_i)}{\partial \mu}. \tag{1}$$

The boundary conditions are

$$p_i(\mu = \mu_{min}, t) = 0, \quad p_i(\mu = \mu_{max}, t) = 0, \quad p_i(\mu, t = t^0) = p_i^0(\mu),$$

and the terms are

$$p(\mu,t) = \sum_{i=1}^{n} p_i(\mu,t), \quad H(\mu, p_1, p_2, \cdots, p_n, t) = \sum_{i=1}^{n} H_i(\mu, p_1, p_2, \cdots, p_n, t),$$

where p is the total mass distribution, p_i is the mass distribution of the ith species, n is the number of species, μ is the log of the particle diameter over a reference diameter, H_i is the condensation/evaporation rate of a single species, and H is the total condensation/evaporation rate. H_i is given by the expression [*Wexler and Seinfeld*, 1990]

$$H_i = \frac{1}{m}\frac{dm_i}{dt} = \frac{2\pi D_p D_i}{m(1 + \frac{2\ell}{\alpha D_p})}(g_i - c_i), \tag{2}$$

where D_p is the diameter of the aerosol particle, D_i is the molecular diffusivity of species i in air, m_i is the mass of species i in a particle of diameter D_p, m is the total mass of the particle, ℓ is the mean free path, α is the sticking coefficient, g_i is the concentration of species i in the gas phase and c_i is the surface concentration of species i.

To solve (1), the aerosol distribution is discretized using a sectional approach [*Gelbard and Seinfeld*, 1980; *Gelbard et al.*, 1980]. The discrete form of the equation is solved using operator splitting techniques [*Yanenko*, 1971] and a modified Bott advection scheme [*Bott*, 1989; *Dhaniyala and Wexler*, 1996] in which the growth term is calculated before the advection term in order to avoid particles being left behind in the lower bins [*Dabdub and Seinfeld*, 1994; *Zhang et al.*, 1999].

3. The Inverse Problem

The goal of inverse modeling is to estimate model parameters which, when implemented in the forward model, yield solutions that are in optimal agreement with a set of observational data. The first step is to calculate a trial solution of the forward model (1) using a background (first guess) value for the model parameters, χ. The discrepancy between the trial solution and what is known from observations is measured by the cost function, which can be represented in general form as

$$\mathcal{J}(p_i, \chi) = \int_{t^0}^{T} \int_{\mu_{min}}^{\infty} J_0(p_i(\mu, t)) d\mu dt. \tag{3}$$

More specifically, for data assimilation problems, the cost function $\mathcal{J}$ is given as

$$\mathcal{J}(p_i, \chi) = \frac{1}{2}(\chi - \chi_b)B^{-1}(\chi - \chi_b) + \frac{1}{2}\sum_{i}^{n}\sum_{k\in\Omega}(y^k - h(p_i^k))R_k^{-1}(y^k - h(p_i^k)). \tag{4}$$

where Ω is the set of discrete time points t^k for which data are known, y^k are the observations at time t^k, h maps the solution from the model space to the observational space, χ_b is the apriori (background) estimate of χ, the matrix B is the error covariance associated with the background term, and the R_k are error covariances of the observations. The optimal model solution and parameters are found by solving the minimization problem

$$\min_{\chi} \mathcal{J}(p_i, \chi),$$

where $\mathcal{J}_{min}$ is found using the gradient resulting from taking the derivative of (3) with respect to χ. The difficulty lies in the fact that there is typically no single equation relating the model parameters to the model solution. $\mathcal{J}$ depends on χ implicitly through the dependency of p_i on χ given by the forward model. The inverse model provides a means of calculating the derivative of the cost function with respect to the model parameters, $\nabla_\chi \mathcal{J}$.

3.1. The Adjoint Method

In the following subsections, we give the equations for $\nabla_\chi \mathcal{J}$ derived using both the continuous and discrete adjoint methods. While there is no formal advantage of one method over another in any general sense, one approach may be better suited to a given application. Typically, the discrete approach yields analytical gradients by implementing in reverse order the exact numerical code used to calculate the forward model, thereby capturing whatever variable dependencies and nonlinearities are included in the discetized forward model. Alternatively, to derive the continuous adjoint equations by hand, one must linearize the equations first, leading to gradients that can be highly approximate. Furthermore, if the governing equation is solved using an explicit numerical algorithm, it can be possible to generate the discrete adjoint codes easily and quickly using automatic differentiation software. On the other hand, deriving the continuous adjoint equations often provides insight into the physical meanings of the adjoint variables and boundary conditions, and the solution to these equations can usually be implemented more efficiently than automatically generated adjoints of the discretized model.

We present the continuous adjoint equation first. Then we consider the adjoint of the discretized governing equation as is generated by the Tangent Adjoint Model Compiler

(TAMC) [*Giering and Kaminski*, 1998]. In Section 4 we compare the results of each approach using a sample system representative of atmospheric aerosols.

3.2. Continuous Adjoint Equations

For the continuous adjoint equations, we consider the case where the model parameters are simply the initial distributions of each species,

$$\chi = p_i(\mu, 0) = p_i^0.$$

The equation adjoint to (1) is

$$\frac{\partial \lambda_i}{\partial t} = -\sum_{j=1}^{n} \lambda_j H_j - p \sum_{j=1}^{n} \lambda_j \frac{\partial H_j}{\partial p_i} - \frac{1}{3} \sum_{j=1}^{n} p_j \frac{\partial \lambda_j}{\partial \mu} \frac{\partial H}{\partial p_i} - \frac{H}{3} \frac{\partial \lambda_i}{\partial \mu} - \frac{\partial J_0}{\partial p_i}, \tag{5}$$

the derivation of which is given in Appendix A. The adjoint equation is integrated backward in time from the "initial conditions"

$$\lambda(\mu, T) = 0$$

to the "final conditions"

$$\lambda_i(\mu, t^0) = \nabla_{p_i^0} \mathcal{J} \tag{6}$$

to solve for the adjoint variable $\lambda(\mu, t)$ at $t = 0$, which we see from (6) is the gradient of the cost function with respect to the initial distribution.

Although we have derived the adjoint equation (5) in continuous form, the continuous method is, in practice, still a hybrid of continuous and discrete calculations. The nonlinear dependence of H upon $p_i(\mu, t)$ for growth laws such as that given by (2) makes the $\frac{\partial H}{\partial p_i}$ term of the adjoint equation (5) difficult to evaluate using continuous equations; therefore, automatic differentiation is used to calculate this term. (This nonlinearity also makes it difficult to distinguish between those variations in H caused by variations of parameters

within the growth law, and those caused by variations in $p_i(\mu, t)$, which is why we have limited the scope of the continuous analysis to $\chi = p_i^0$.) In addition, both continuous forward and adjoint equations are eventually integrated numerically, further blurring the distinction between the continuous and discrete approaches.

3.3. Discrete Adjoint Equations

In this section we explicitly derive the discrete adjoint formulas to illustrate the differences between the continuous and discrete approaches. The actual formulas used were created automatically using TAMC. A complete explanation of the theory and algorithms used in TAMC is given by *Giering and Kaminski* [1998].

We begin with a discretized form of the governing equation, which we shall represent below as

$$[p_i]_j^k = F_j(p_i^{k-1}, g_i^{k-1}), \quad k = 1, \ldots, N, \quad i = 1, \ldots, n, \quad j = 1, \ldots, s. \tag{7}$$

where $[p_i]_j^k$ is the concentration of species i in the jth bin at time step k, p_i^k is the vector of all particulate concentrations, g_i^k is the vector of all gas concentrations, and F_j represents the numerical operator describing gas / particle transport and advection in diameter space. An informative example to consider is when the observations are simply the concentrations at the final time step, and the only recoverable parameters are the initial conditions. In this case, $\Omega = N$, h is simply an identity, and, ignoring background terms, the cost function can be written as

$$\mathcal{J}(p_i^0) = \frac{1}{2} \sum_{j=1}^{s} \sum_{i=1}^{n} ([y_i]_j^N - [p_i]_j^N) R_k^{-1} ([y_i]_j^N - [p_i]_j^N). \tag{8}$$

The desired quantity to be computed is the derivative of the cost function with respect to changes in the vector of initial conditions,

$$\nabla_{p_i^0}\mathcal{J} = \frac{\partial \mathcal{J}(p_i^N)}{\partial p_i^0}. \tag{9}$$

Using the chain rule, one can expand the RHS of (9)

$$\nabla_{p_i^0}\mathcal{J} = \left[\frac{\partial p_i^1}{\partial p_i^0}\right]^T \left[\frac{\partial p_i^2}{\partial p_i^1}\right]^T \cdots \left[\frac{\partial p_i^N}{\partial p_i^{N-1}}\right]^T \left[\frac{\partial \mathcal{J}(p_i^N)}{\partial p_i^N}\right] \tag{10}$$

Evaluation of the RHS of (10) from right to the left corresponds to calculating $\nabla_{p_i^0}\mathcal{J}$ via the adjoint method, while calculating this series of matrix products from left to right constitutes a forward sensitivity calculation. Careful consideration of the number of required scalar multiplications shows that the computational demands of the adjoint method are significantly less than those of the forward method when the dimension of $\mathcal{J}$ is smaller than the dimension of p [*Kaminski et al.*, 1999; *Sandu et al.*, 2003]. Since in this case $\mathcal{J}$ is a scalar and p has $n \times s$ elements, calculating this series of matrix products in reverse is preferable.

Defining the discrete adjoint variable as

$$\lambda^k = \left[\frac{\partial p_i^N}{\partial p_i^k}\right]^T \left[\frac{\partial \mathcal{J}(p_i^N)}{\partial p_i^N}\right] = \nabla_{p_i^k}\mathcal{J} \tag{11}$$

and initializing λ^k as $\lambda^N = \nabla_{p_i^N}\mathcal{J}$, $\lambda^0 = \nabla_{p_i^0}\mathcal{J}$ can be found iteratively (beginning with $k = N$ and ending with $k = 1$) using the following expression

$$\lambda^{k-1} = \left[\frac{\partial p_i^k}{\partial p_i^{k-1}}\right]^T \lambda^k. \tag{12}$$

In this manner, the adjoint method is reduced to calculating $\frac{\partial p_i^k}{\partial p_i^{k-1}} = \frac{\partial F_j(p_i^k, g_i^k)}{\partial p_i^{k-1}}$ at each step. $F_j(p_i^k, g_i^k)$ is implemented using standard FORTRAN constructs such as loops, conditionals, basic functions and algebraic manipulations, for which algorithms for calculating the

derivatives are known [*Giering and Kaminski*, 1998; *Giles et al.*, 2003], hence the adjoint code can be constructed automatically. One potentially problematic routine in $F_j(p_i^k, g_i^k)$ is the Bott-advection scheme: the positive-definite constraints contain many evaluations of min / max statements, whose derivatives are undefined if the arguments are equal. To avoid this problem, we use double precision floating point numbers and resign ourselves to arbitrarily choosing the path of dependence in the rare case that the arguments are exactly equal.

Due to the nonlinear nature of $F_j(p_i^k, g_i^k)$ introduced by the dynamic time step and nonlinearities in the growth law, $\frac{\partial F_j(p_i^k, g_i^k)}{\partial p_i^{k-1}}$ will depend upon p_i^k and g_i^k, hence their values from the forward trajectories will be required at each step of the iteration. This can lead to significant storage requirements and read / write demands for full-scale models with many components in many cells. Similar situations have been handled gracefully by checkpointing schemes which minimize these types of computational demands (for example *Elbern and Schmidt* [1999], or the distributed scheme implemented for a parallel model of *Sandu et al.* [submitted 2003]); these techniques could be applied to the aerosol adjoint model as well.

4. Inverse Modeling of Aerosol Size-Composition Dynamics

In order to assess the various adjoint models, we perform multiple twin experiments on a test system that consists of three species whose properties are designed to be representative of conditions commonly encountered in atmospheric aerosols. Observations are sampled from the reference, or true, solution generated using the forward model. The simulation is repeated with perturbed values of the parameters, and the reference values are recovered through inverse modeling. The adjoint method is used to calculate the gradient of the cost

function with respect to the initial distributions and/or pure species vapor concentrations. For this type of test, we can consider the observations to be exact, and the cost function reduces to

$$\mathcal{J}(p_i^0) = \frac{1}{2}\sum_{j=1}^{8}\sum_{i=1}^{3}\sum_{k\in\Omega}([y_i]_j^k - [p_i]_j^k)^2.$$

This cost function does not penalize departure from the background estimates since we know that the observations are correct while the initial guesses are wrong. The cost function is then minimized using a limited memory BFGS algorithm [*Byrd et al.*, 1995; *Zhu et al.*, 1994], providing optimized estimates of the desired quantities. To simplify the calculations, the components of the test system are assumed to have ideal thermodynamic properties. Ignoring surface tension and non-ideal effects, Raoults law and the ideal gas equation can be used to express the surface vapor concentration as a function of the particle composition,

$$c_i = x_i c_i^\circ$$

where x_i is the aerosol phase mole fraction and c_i° is the pure component vapor concentration of species i. If we assume, for simplicity, that each species has equal molecular mass, then the fractions are equivalent to the mass fractions, and the growth rate can be written as

$$H_i = \frac{2\pi D_p D_i}{m(1+\frac{2\ell}{\alpha D_p})}(g_i - \frac{p_i}{\sum_{i=1}^{s} p_i}c_i^\circ) \tag{13}$$

The initial conditions for the reference (true) solution used throughout this study are given in Table 1, and the physical properties of the aerosols are $\alpha = 0.1$, $\ell = 65$ nm, and $D_i = 1\times10^{-5}$ m^2/s. In the aerosol phase, each species is initially log-normally distributed: species 1 is located in the smaller bins, species 2 in the larger bins, and species 3 across all bins. The gas-phase concentrations and pure component vapor concentrations are

selected such that species 1 condenses and species 2 evaporates, while the third species is nonvolatile. Figures 1, 2 and 3 show the reference run at $t = 0$, 15 min, and 2.5 h, respectively. Most of the progress towards an equilibrium distribution is made during the first 15 min. Figure 4 shows the time evolution of the gas phase concentrations. Species 1 condenses before species 2 evaporates because gas / particle transport takes longer for the larger particles. The initial decrease in the vapor concentration of species 2 occurs because its mole fraction is very low in the smaller particles, causing the effective surface vapor concentration for these particles to be lower than the surrounding gas concentration.

For use with the discrete adjoint model, the time step for the forward numerical simulation is adjusted dynamically to be as long as possible while still meeting the following criteria: it always satisfies the Courant stability condition, and it is sufficiently small to justify operator splitting. After an initial brief period during which most of species 1 condenses, the time step levels off to a value of ~ 18 s, leading to a simulation in which 50 steps spans ~ 15 min.

The continuous adjoint equation (5) for the forward model is solved using finite differences. Due to the nonlinearity of (1), solving the adjoint equation requires values from the forward solution. Rather than allow each integration to have a different time stepping scheme, and then attempt to match the trajectories by interpolating, it is preferable to use a static time step for both forward and backward runs. In order to avoid the possibility of either solution becoming unstable, the time step is fixed at 5.0 s. Consequently, the number of time steps required to run the continuous model is almost four times greater than that required to run the discrete model.

Multiple assimilation studies were performed using the test system described above. The studies were grouped into four scenarios according to how much information was initially known and how observations were used to recover the unknown data. Discrete adjoint codes were generated using TAMC for each scenario. Reconstructing the adjoint model for each set of dependent and independent variables did not present a major challenge, as the calculation of an adjoint model of this system using TAMC takes less than a few minutes.

Table 3 summarizes the conditions and results of each of the cases considered. The RECOVER column lists which parameters were being assimilated; the numbers refer to species whose initial distribution (p_i^0) or pure surface concentrations (c_i°) were unknown. The initial guesses for these unknown parameters are given in the GUESS column. The notation $\times(a, b, c)$ indicates that the initial guess was equal to the true value multiplied by a factor of a, b, c for the 1^{st}, 2^{nd} and 3^{rd} species, respectively, while $+(a, b, c)$ implies that the true values were amended by these amounts. The extent to which details of the reference solution were included as observations is summarized by the three columns under the OBSERVE heading. The numbers in the bin column indicate which of the bins were observed (terms like $\overline{12}$ indicate that only the total concentration in bins 1 through 2 was known), and the numbers in the species column indicate which species were measured. The ratio in the time column is the time between observations over the total simulation time (both in minutes). The R column gives the results of each test. A measure of the relative success of the data assimilation is the percent of the error in the initial guess that is still present after optimization,

$$R(z) = \left[\frac{\sum_{bins/species} (z_{optimized} - z_{true})^2}{\sum_{bins/species} (z_{guess} - z_{true})^2} \right]^{1/2} \tag{14}$$

where z is either p_i^0 or c_i°. Low values of R imply that either the initial guess was extremely bad or the assimilation converged to the true value.

4.1. Case 1: Recovery of Initial Distributions

The most important aspect of the data assimilation is the ability to recover the initial distribution, as determination of other parameters is dependent upon the adjoint of the concentration variable. Case 1a is the easiest test, with all 3 species being measured in all 8 bins and all the surface concentrations considered known. Cases 1a-c.*i* used the discrete adjoint model while cases 1a-c.*ii* used the continuous adjoint model. The reference, guessed, and optimized initial distributions for cases 1a.*i* and 1a.*ii* are shown in Figures 5 and 6, respectively. Both adjoint models recover the true distribution very well, and the continuous model converges more completely than the discrete model in this case. Considering a longer assimilation period (40 min), yet still only making an observation at the final time, the results of Case 1b.*i* and 1b.*ii* show that in this situation the discrete model optimizes to a more accurate set of initial distributions. In Case 1c, the simulation time is 2.5 h, but observations are still taken every $\sim$ 15 min. Figures 7 and 8 show that the optimized p_i^0 are greatly improved over the initial guess, yet still noticeably far from the true distribution. Overall, when the interval between consecutive observations is relatively short ($\sim$15 min), the continuous method provides better estimates than the discrete method; however, the opposite becomes true as the distribution of observations becomes increasingly sparse. Given only a single observation over a period of 2.5 h, the discrete model performs much better than the continuous model (Case 1d).

In addition to comparing the ability of the two types of adjoint models to recover the initial distributions, it is important to compare the computational expense of each

approach. Table 2 summarizes the timing statistics for the first scenario, where t_f is the time for the forward calculation, t_b is the time for the backward calculation, and N_{cf} is the number of cost function evaluations during minimization. These data can be used to define a total expense ratio, η_{tot}, where

$$\eta_{tot} = \frac{\text{Total computational time (discrete)}}{\text{Total computational time (continuous)}} \tag{15}$$

Since the total computational time for each test is approximately equal to $N_{cf} * (t_f + t_b)$, this ratio can be further broken down into a product of ratios which are fairly consistent in magnitude throughout each test,

$$\eta_{tot} = \eta_{nf}\eta_{fwd}\eta_{adj} \tag{16}$$

where

$$\begin{aligned} \eta_{nf} &= \frac{N_{cf}(d)}{N_{cf}(c)} \\ \eta_{fwd} &= \frac{t_f(d)}{t_f(c)} \\ \eta_{adj} &= \frac{\left(1 + \frac{t_b(c)}{t_f(c)}\right)}{\left(1 + \frac{t_b(d)}{t_f(d)}\right)} \end{aligned}$$

The ratio of the backward to forward calculation times, η_{adj}, is much lower for the continuous model than the discrete model. However, as indicated by the overall number of function calls required during the optimization, the gradients from the continuous model are not as accurate as those from the discrete model. Both these results are consistent with what one would expect from these two types of models. Simplifications made to derive the adjoint equations in continuous form lead to faster calculations that are more approximate in nature.

In addition to analyzing the fundamental capabilities of the adjoint method in this test system, we would like to make recommendations for the direction of future work involving

more sophisticated aerosol models. As the complexity of the model increases, a continuous derivation will require an increasingly large number of approximations, leading to adjoint times that are faster, yet gradients that are not as accurate; hence, we speculate that η_{nf} will decrease and η_{adj} will increase. If, to a first order, these effects cancel each other out, the overall efficiency of a more complex aerosol model will depend upon η_{fwd}. In this simple model, η_{fwd} is $\sim 1/4$ because the average time step taken in the discrete model is about four times as long as the static time step set in the continuous model. For detailed aerosol models, the range of the dynamic time step can span several orders of magnitude. Using a static time step will force the forward calculation for the continuous model to be much slower than the forward calculation for the discrete model, causing η_{fwd}, and likely η_{tot}, to be less than unity by several orders of magnitude. To avoid this, one could use dynamic time steps for both forward and backward runs of the continuous model; however, the interpolation process required to utililize data from the forward trajectory when solving the adjoint equation may increase the error in the resulting gradient. While there are no inherent restrictions on the types of time steps that can be used to solve the continuous equations, these issues can complicate their implementation. In short, the discrete adjoint formulation appears to be the more viable method.

4.2. Case 2: Recovery of Pure Species Vapor Concentrations

The next set of tests examines the situation in which the initial distributions of all the components are known, but the pure component surface vapor concentrations are not. The value of $R(c_i^\circ)$ for Case 2a is 0.00 because the true values of c_i° are recovered to at least six significant digits. For example, the optimized value of c_1° is 1.0000028. Case 2b considers the situation in which the initial guesses for c_i° are such that the overall transport

of each species is in the opposite direction than in the true solution. For example, with c_1° = 20 μg/m^3, species 1 evaporates instead of condensing. Again, the optimized c_i° matches the true value to at least six significant digits, indicating that c_i° can be recovered even when the overall direction of the mass transport is not known before the initial analysis.

4.3. Case 3: Recovery of Initial Distribution and Vapor Concentrations

The third scenario addresses a common question encountered in aerosol measurement — based upon accurate information of a subset of the aerosol components, what can be inferred about an unmeasured species? In this set, no information about species 1 is used in performing the assimilation, and the cost function is

$$\mathcal{J}(p_i^0) = \frac{1}{2}\sum_{j=1}^{8}\sum_{i=2}^{3}([y_i]_j^N - [p_i]_j^N)^2.$$

Results for Case 3a indicate that both p_1^0 and c_1° can be recovered simultaneously. While these results look promising, to say that "nothing" was known about species 1 is perhaps misleading in that the initial guess for p_1^0 had the same shape as the true solution, greatly facilitating the assimilation.

To determine how much the success of the assimilation depends upon the shape of the initial guess, Case 3b starts with p_1^0 being a constant value of 5 μg/m^3 throughout the size distribution. Not surprisingly, with such a poor initial guess, the performance is drastically decreased, as indicated by $R(p_1^0) = 0.49$. However, a plot of the initial distribution shows that the assimilation is very successful for all parameters except the concentrations in the two largest size bins (Figure 9). To understand why this would be the case, it is useful to recall that the driving term for the discrete adjoint model is $\frac{\partial \mathcal{J}(p_i^N)}{\partial p_i^N}$. In other words, the adjoint model is forced by the difference in the concentration of the observed species

between the guessed and the reference solutions at the time when the observations were made. For Case 3b, the simulation results at $t = 15$ min are shown in Figure 10, and we see optimization of p_1^0 in bins 7 and 8 was stopped prematurely because there was no longer any driving force for the adjoint model; the optimized solution had already converged to the true value. Since the characteristic time for condensation / evaporation in bin 7 is several hours, the concentrations in the larger bins had yet to change significantly after only 15 min. In this situation, as confirmed by the results of Case 3c, it is advantageous to run the simulation longer before taking an observation in order to provide ample forcing for the adjoint model. On the other hand, if the observation time is delayed too long, the assimilation would become impossible (imagine trying to determine the initial condition for an aerosol that has equilibrated to an evenly distributed profile).

4.4. Case 4: Recovery from Partial Distributions

In addition to considering variations in the observation frequency and species detection, it is of interest to examine the performance of the data assimilation when only portions of the size distribution are measured. Scenario 4a addresses the situation in which observations are made only in the smaller four size bins,

$$\mathcal{J}(p_i^0) = \frac{1}{2}\sum_{j=1}^{4}\sum_{i=1}^{3}([y_i]_j^N - [p_i]_j^N)^2$$

Based upon this information, the initial concentrations in the larger bins were determined and are shown in Figure 11. At first glance, the results appear to be fairly poor; however, one must take into account the direction that each species is advecting. Considering the initial guess as a perturbation of the reference solution, the effect that this perturbation has on the concentrations in the smaller four bins is the driving force for the adjoint model.

For species 1, the lower half of the distribution is largely invariant to perturbations in the upper four bins because this component is growing. However, for species 2, particles are evaporating and advection is bringing information about the contents of bins 5-8 to bins 1-4, hence we would expect the assimilation to have performed better for species 2 than for species 1. Indeed this is the case. Providing further forcing by running the simulation longer also leads to better results (Case 4b), and not surprisingly, if distributions 1 and 3 are considered known, then the assimilation of species 2 is even more improved (Case 4c, Figure 12).

Tests 4d-f address cases in which the observed concentrations are actually sums over two or more adjacent size bins. Since the observations are no longer exactly equivalent to the state variables, this averaging is represented by the function h in the cost function,

$$\mathcal{J}(p_i^0) = \frac{1}{2}\sum_{\bar{j}=1}^{2}([y]_{\bar{j}}^N - h_{\bar{j}}([p]_j^N))^2,$$

where $\bar{j}$ is the index of the lumped bins. In Case 4d, each pair of adjacent bins is averaged, while in 4e the observed distribution is of only two bins—one that containes particles whose diameter is smaller than 2.76 μm, and one that containes particles that are larger. The adjoint method is only able to resolve the initial distributions to a level consistent with the resolution of the initial guess. Given an initial guess that is resolved on the scale of an 8-bin distribution, the assimilations are fairly successful. However, the optimized distributions become increasingly featureless as the resolution of the initial guess is decreased, see Figure 13.

5. Conclusions

As part of a broad effort to better the understanding of the state of the atmosphere using inverse modeling techniques, this paper focused on the specific goal of incorporating multicomponent, size resolved aerosols in data assimilation studies. The adjoint method has been explored as a means of recovering parameters of an aerosol distribution evolving by condensation / evaporation. Within the field of adjoint modeling, we have explored two general tactics for creating the inverse model—discrete and continuous. Evaluating these methods with a simplified, yet representative, model of an atmospheric aerosol, we have attempted to recover parameters of the distribution by assimilating observations that are sparse in time, size and / or chemical resolution.

Intricacies of what was still a simple test model (compared to the aerosol routines implemented in detailed CTMs) limited the feasibility of formulating the adjoint equations in an entirely continuous fashion. In particular, nonlinearities introduced by the particle growth rate limits the extent to which the continuous equations can be derived in full. Nonetheless, the results of problems that have been addressed using the continuous approach are comparable to those found the using discrete approach. However, the flexibility of discrete adjoint models, combined with the ease of creating them automatically using programs such as TAMC, makes them the more viable method for solving inverse problems involving increasingly complex aerosol systems.

In the test problem considered, we attempted to recover parameters such as the initial distribution and the species' pure surface concentrations. Either of these were easily recovered for all three species when at least one observation of the entire distribution was known sufficiently prior to equilibration. Additionally, if both of these properties for a

single species were unknown, and this species was never even observed, the adjoint calculations allowed us to adequately infer this information from measurements of the dynamic evolution of the other two species. The most difficult task attempted was the recovery of initial distributions based upon observations in only a subset of the size range. For understandable reasons, this type of assimilation required the most observational information in order to yield decent estimates of the aerosol parameters. Overall, we demonstrated that the adjoint method can be used to recover information about a dynamic, size and chemically resolved aerosol distribution under a variety of conditions.

Appendix A: Derivation of Continuous Adjoint Equation

We will use the Lagrangian multiplier method to derive the continuous adjoint derivations. The cost function is defined as

$$\begin{aligned}\mathcal{J} \;=\; & \int_{t^0}^{T}\int_{0}^{\infty} J_0(p_1(\mu,t),p_2(\mu,t),\ldots,p_n(\mu,t))\,d\mu\,dt \\ & -\sum_{i=1}^{n}\int_{t_0}^{T}\int_{0}^{\infty}\lambda_i(\mu,t)(LHS_{p_i}-RHS_{p_i})\,d\mu\,dt, \end{aligned} \tag{A1}$$

Here LHS_{p_i} and RHS_{p_i} refer to the left side and right side of (1), respectively. J_0 is the local cost function component,

$$J_0 = \frac{1}{2}\sum_{i=1}^{n}(y_i - h(p_i))R_k^{-1}(y_i - h(p_i))\delta(t-t^k), \tag{A2}$$

where $t^k \in \Omega$. Taking the variation of (A1), we get

$$\begin{aligned}\delta\mathcal{J} \;=\; & \int_{t^0}^{T}\int_{0}^{\infty}\sum_{i=1}^{n}\frac{\partial J_0}{\partial p_i}\delta p_i(\mu,t)\,d\mu\,dt \\ & -\int_{t^0}^{T}\int_{0}^{\infty}\sum_{i=1}^{n}\delta\lambda_i(\mu,t)(LSH_{p_i}-RHS_{p_i})\,d\mu\,dt \\ & -\int_{t^0}^{T}\int_{0}^{\infty}\sum_{i=1}^{n}\lambda_i(\mu,t)\delta(LSH_{p_i}-RHS_{p_i})\,d\mu\,dt \end{aligned} \tag{A3}$$

Inserting the expressions of LSH_{p_i} and RHS_{p_i}, (A3) can be written as

$$\begin{aligned}\delta\mathcal{J} \;=\; & \int_{t^0}^{T}\int_0^{\infty}\sum_{i=1}^{n}\frac{\partial J_0}{\partial p_i}\delta p_i(\mu,t)\,d\mu\,dt \\ & -\int_{t^0}^{T}\int_0^{\infty}\sum_{i=1}^{n}\delta\lambda_i(\mu,t)(LSH_{p_i}-RHS_{p_i})\,d\mu\,dt \\ & -\int_{t^0}^{T}\int_0^{\infty}\sum_{i=1}^{n}\lambda_i(\mu,t)\delta(\frac{\partial p_i(\mu,t)}{\partial t}-H_i(\mu,p_1,p_2,\cdots,p_n,t)p(\mu,t)+\frac{1}{3}\frac{\partial}{\partial\mu}\left(Hp_i\right))\,d\mu\,dt\end{aligned} \tag{A4}$$

Then we can re-write (A4) as

$$\delta\mathcal{J} \;=\; \sum_{i=1}^{n}\int_{t^0}^{T}\int_0^{\infty}\delta p_i(\mu,t)\,\frac{\partial J_0(\mu,p,t)}{\partial p_i}\,d\mu\,dt \tag{A5a}$$

$$-\sum_{i=1}^{n}\int_{t^0}^{T}\int_0^{\infty}\delta\lambda_i(\mu,t)(LHS_{p_i}-RHS_{p_i})\,d\mu\,dt \tag{A5b}$$

$$-\sum_{i=1}^{n}\int_{t^0}^{T}\int_0^{\infty}\lambda_i(\mu,t)\frac{\partial(\delta p_i(\mu,t))}{\partial t}\,d\mu\,dt \tag{A5c}$$

$$+\sum_{i=1}^{n}\int_{t^0}^{T}\int_0^{\infty}\lambda_i(\mu,t)H_i(\mu,p,t)\sum_{j=1}^{n}\delta p_j(\mu,t)\,d\mu\,dt \tag{A5d}$$

$$+\sum_{i=1}^{n}\int_{t^0}^{T}\int_0^{\infty}\lambda_i(\mu,t)\sum_{j=1}^{n}\frac{\partial H_i(\mu,p,t)}{\partial p_j}\delta p_j p(\mu,t)\,d\mu\,dt \tag{A5e}$$

$$-\frac{1}{3}\sum_{i=1}^{n}\int_{t_0}^{T}\int_0^{\infty}\lambda_i(\mu,t)\frac{\partial}{\partial\mu}\left(\sum_{j=1}^{n}\frac{\partial H(\mu,p,t)}{\partial p_j}\delta p_j p_i+H(\mu,p,t)\delta p_i\right)\,d\mu\,dt \tag{A5f}$$

If we choose the final condition $\lambda(\mu,T)=0$, and integrate (A5c) by parts, (A5c) becomes

$$\sum_{i=1}^{n}\int_0^{\infty}\lambda_i(\mu,t_0)\delta p_i(\mu,t_0)\,d\mu+\sum_{i=1}^{n}\int_{t^0}^{T}\int_0^{\infty}\delta p_i(\mu,t)\frac{\partial(\lambda_i(\mu,t))}{\partial t}\,d\mu\,dt \tag{A6}$$

Likewise, letting $\lambda_i(0,t)=0$, $p_i(+\infty,t)=0$, (A5f) can be written as

$$\frac{1}{3}\sum_{i=1}^{n}\int_{t_0}^{T}\int_0^{\infty}\frac{\partial\lambda_i(\mu,t)}{\partial\mu}\left(p_i\sum_{j=1}^{n}\frac{\partial H(\mu,p,t)}{\partial p_j}\delta p_j+H(\mu,p,t)\delta p_i\right)\,d\mu\,dt \tag{A7}$$

If $p(\mu,t)$ is the solution of (1), $LHS_{p_i}-RHS_{p_i}=0$, then

$$\delta\mathcal{J} \;=\; \sum_{i=1}^{n}\int_{t^0}^{T}\int_0^{\infty}\delta p_i(\mu,t)\frac{\partial J_0(\mu,t)}{\partial p_i}\,d\mu\,dt \tag{A8a}$$

$$+\sum_{i=1}^{n}\int_{t^0}^{T}\int_0^{\infty}\frac{\lambda_i(\mu,t)}{\partial t}\delta p_i(\mu,t)\,d\mu\,dt \tag{A8b}$$

$$+\sum_{i=1}^{n}\int_{0}^{\infty}\lambda_i(\mu,t^0)\delta p_i(\mu,t^0)\,d\mu \tag{A8c}$$

$$+\sum_{i=1}^{n}\int_{t^0}^{T}\int_{0}^{\infty}\lambda_i(\mu,t)H_i(\mu,p,t)\sum_{j=1}^{n}\delta p_i(\mu,t)\,d\mu\,dt \tag{A8d}$$

$$+\sum_{i=1}^{n}\int_{t^0}^{T}\int_{0}^{\infty}\lambda_i(\mu,t)\sum_{j=1}^{n}\frac{\partial H_i(\mu,p,t)}{\partial p_j}p(\mu,t)\delta p_j\,d\mu\,dt \tag{A8e}$$

$$+\frac{1}{3}\sum_{i=1}^{n}\int_{t_0}^{T}\int_{0}^{\infty}\frac{\partial\lambda_i(\mu,t)}{\partial\mu}\sum_{j=1}^{n}\frac{\partial H(\mu,p,t)}{\partial p_j}\delta p_j p_i\,d\mu\,dt \tag{A8f}$$

$$+\frac{1}{3}\sum_{i=1}^{n}\int_{t_0}^{T}\int_{0}^{\infty}\frac{\partial\lambda_i(\mu,t)}{\partial\mu}H(\mu,p,t)\delta p_i\,d\mu\,dt \tag{A8g}$$

Assigning the coefficient in front of δp_i to 0 results in the adjoint equation,

$$\frac{\partial\lambda_i}{\partial t}=-\sum_{j=1}^{n}\lambda_j H_j-p\sum_{j=1}^{n}\lambda_j\frac{\partial H_j}{\partial p_i}-\frac{1}{3}\sum_{j=1}^{n}p_j\frac{\partial\lambda_j}{\partial\mu}\frac{\partial H}{\partial p_i}-\frac{H}{3}\frac{\partial\lambda_i}{\partial\mu}-\frac{\partial J_0}{\partial p_i} \tag{A9}$$

Acknowledgments. The authors thank the National Science Foundation for supporting this work through the award NSF ITR AP&IM 0205198. The work of A. Sandu was also partially supported by the award NSF CAREER ACI 0093139.

References

Bott, A., A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes, *Mon. Weather Rev.*, *117*, 1006, 1989.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, A limited memory algorithm for bound constrained optimization, *Scientific Computing*, *16*(5), 1190–1208, 1995.

Cacuci, D. G., Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach, *J. Math. Phys.*, *22*(12), 2794, 1981a.

Cacuci, D. G., Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses, *J. Math. Phys.*, *22*(12), 2803, 1981b.

Clerbaux, C., J. Hadji-Lazaro, D. Hauglustaine, G. Megie, B. Khattatov, and J.-F. Lamarque, Assimilation of carbon monoxide measured from satellite in a three-dimensional chemistry-transport model, *J. Geophys. Res.*, *106*(D14), 15,385–15,394, 2001.

Collins, W., P. Rasch, B. Eaton, B. Khattatov, J.-F. Lamarque, and C. Zender, Simulating aerosols using a chemical transport model with assimilation of satellite aerosol retrievals: Methodology for INDOEX, *J. Geophys. Res.*, *106*(D7), 7313–7336, 2001.

Dabdub, D., and J. H. Seinfeld, Numerical advective schemes used in air quality models–sequential and parallel implementation, *Atmos. Environ.*, *28*(20), 3369–3385, 1994.

Daescu, D., A. Sandu, and G. R. Carmichael, Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: II — Numerical validation and applications, *Atmos. Environ.*, *37*, 5097–5114, 2003.

Dhaniyala, S., and A. S. Wexler, Numerical schemes to model condensation and evaporation of aerosols, *Atmos. Environ.*, *30*(6), 919–928, 1996.

Elbern, H., and H. Schmidt, A four-dimensional variational chemistry data assimilation scheme for eulerian chemistry transport modeling, *J. Geophys. Res.*, *104*(D15), 18,583–18,598, 1999.

Elbern, H., and H. Schmidt, Ozone episode analysis for four-dimensional variational chemistry data assimilation, *J. Geophys. Res.*, *106*, 3569–3590, 2001.

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

Errera, Q., and D. Fonteyn, Four-dimensional variational chemical assimilation of crista stratospheric measurements, *J. Geophys. Res.*, *106*(D11), 12,253–12,265, 2001.

Fisher, M., and D. J. Lary, Lagrangian four-dimensional variational data assimilation of chemical species, *Q. J. R. Meteorol. Soc.*, *121*, 1681–1704, 1995.

Gelbard, F., and J. H. Seinfeld, Simulation of multicomponent aerosol dynamics, *J. Colloid Interface Sci.*, *78*(2), 485, 1980.

Gelbard, F., Y. Tambour, and J. H. Seinfeld, Sectional representation for simulating aerosol dynamics, *J. Colloid Interface Sci.*, *76*(2), 541, 1980.

Giering, R., and T. Kaminski, Recipes for adjoint code construction, *ACM Trans. Math. Software*, *24*(4), 437, 1998.

Giles, M. B., and N. A. Pierce, An introduction to the adjoint approach to design, *Flow Turbul. Combust.*, *65*, 393–415, 2000.

Giles, M. B., M. C. Duta, J.-D. Müller, and N. A. Pierce, Algorithm developments for discrete adjoint methods, *AIAA Journal*, *41*(2), 198, 2003.

Hoffman, R. N., A four-dimensional analysis exactly satisfying equations of motion, *Mon. Weather Rev.*, *114*, 338–397, 1986.

Kaminski, T., M. Heimann, and R. Giering, A coarse grid three-dimensional global inverse model of the atmospheric transport, *J. Geophys. Res.*, *104*(D15), 18,535–18,553, 1999.

Khattatov, B., J.-F. Lamarque, L. v. Lyjak, R. Menard, P. F. Levelt, X. X. Tie, G. P. Brasseur, and J. C. Gille, Assimilation of satellite observations of long-lived chemical species in global chemistry-transport models, *J. Geophys. Res.*, *105*(D23), 29,135–29,144, 2000.

Lamarque, J.-F., B. Khattatov, J. C. Gille, and G. P. Brasseur, Assimilation of measurement of air pollution from space (MAPS) CO in a global three-dimensional model, *J. Geophys. Res.*, *104*(D21), 26,209–26,218, 1999.

Lamarque, J.-F., B. Khattatov, and J. C. Gille, Constraining tropospheric ozone column through data assimilations, *J. Geophys. Res.*, *107*(D22), 4651–4662, 2002.

Lamb, R. G., W. H. Chen, and J. H. Seinfeld, Numerico-empirical analysis of atmospheric diffusion theories, *J. Atmos. Sci.*, *32*, 1794–1807, 1975.

Levelt, P. F., B. Khattatov, J. C. Gille, G. P. Brasseur, X. X. Tie, and J. Waters, Assimilation of MLS ozone measurements in a global three-dimensional chemistry-transport model, *Geophys. Res. Lett.*, *25*, 4493–4496, 1998.

Lorenc, A. C., Analysis methods for numerical weather prediction, *Q. J. R. Meteorol. Soc.*, *112*, 1177–1194, 1986.

Lyster, P. M., S. E. Cohn, R. Menard, L.-P. Chang, S.-J. Lin, and R. Olsen, An implementation of a two-dimensional filter for atmospheric chemical constituent assimilation on massively parallel computers, *Mon. Weather Rev.*, *125*, 1674–1686, 1997.

Marchuk, G., Numerical solution of the problems of the dynamics of the atmosphere and the ocean (in Russian), *Gidrometeoizdat*, 1974.

Meng, Z., D. Dabdub, and J. H. Seinfeld, Size-resolved and chemically resolved model of atmospheric aerosol dynamics, *J. Geophys. Res.*, *102*(D3), 3419–3435, 1998.

Palmer, P. I., D. J. Jacob, D. B. A. Jones, C. L. Heald, R. M. Yantosca, J. A. Logan, G. W. Sachse, and D. G. Streets, Inverting for emissions of carbon monoxide from Asia using aircraft observations over the western Pacific, *J. Geophys. Res.*, *108*(D21), 2003a.

Palmer, P. I., D. J. Jacob, L. J. Mickley, D. R. Blake, G. W. Sachse, H. E. Fuelberg, and C. M. Kiley, Eastern Asian emissions of anthropogenic halocarbons deduced from aircraft concentration data, *J. Geophys. Res.*, *108*, 4753, 2003b.

Pilinis, C., Derivation and numerical solution of the species mass distribution equations for multicomponent particulate systems, *Atmos. Environ.*, *24A*(7), 1923–1928, 1990.

Pironneau, O., On optimum design in fluid mechanics, *J. Fluid Mech*, *64*, 97–110, 1974.

Sandu, A., D. N. Daescu, and G. R. Carmichael, Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: I—Theory and software tools, *Atmos. Environ.*, *37*, 5083–5096, 2003.

Sandu, A., D. N. Daescu, G. R. Carmichael, and T. Chai, Adjoint sensitivity analysis of regional air quality models, submitted 2003.

Sandu, A., D. Daescu, and G. R. Carmichael, Inverse modeling of particle dynamics, submitted 2004.

Sasaki, Y., Some basic formalisms in numerical variational analysis, *Mon. Weather Rev.*, *98*, 875–883, 1970.

Song, C. H., and G. R. Carmichael, A three-dimensional modeling investigation of the evolution processes of dust and sea-salt in East Asia, *J. Geophys. Res.*, *106*, 18,131–18,154, 2001.

Stajner, I., L. Riishojgaard, and R. Rood, The GEOS ozone data assimilation system: specification of error statistics, *Q. J. R. Meteorol. Soc.*, *127*(573), 1069–1094, 2001.

Talagrand, O., and P. Courtier, Variational assimilation of meteorological observations with the adjoint of the vorticity equations. Part I. Theory, *Q. J. R. Meteorol. Soc.*, *113*, 1311–1328, 1987.

Thompson, P. D., Reduction of analysis error through constraints of dynamical consistency, *J. Appl. Meteorol.*, *8*, 738–742, 1969.

Tziperman, E., and W. C. Thacker, An optimal-control / adjoint-equations approach to studying the oceanic general circulation, *J. Phys. Oceanogr.*, *19*, 1471, 1989.

Wang, K.-Y., D. J. Lary, D. E. Shallcross, S. M. Hall, and J. A. Pyle, A review on the use of the adjoint method in four-dimensional atmospheric-chemistry data assimilation, *Q. J. R. Meteorol. Soc.*, *127*, 2181–2204, 2001.

Wexler, A. S., and J. H. Seinfeld, The distribution of ammonium salts among a size and composition dispersed aerosol, *Atmos. Environ.*, *24A*(5), 1231–1246, 1990.

Yanenko, N. N., *The Method of Fractional Steps*, Springer, New York, 1971.

Zhang, Y., C. Seigneur, J. H. Seinfeld, M. Z. Jacobson, and F. S. Binkowski, Simulation of aerosol dynamics: A comparative review of algorithms used in air quality models, *Aerosol Sci. Tech.*, *31*, 487–514, 1999.

Zhu, C., R. H. Byrd, P. Lu, and J. Nocedal, L-BFGS-B: a limited memory FORTRAN code for solving bound constrained optimization problems, *Tech. rep.*, Northwestern University, 1994.

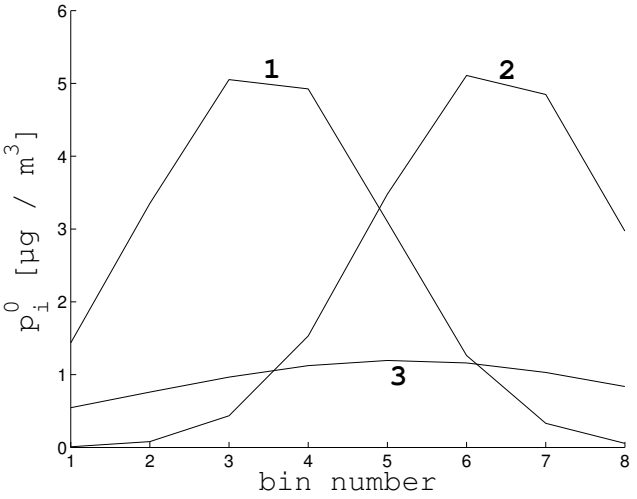

Figure 1. Initial distribution of reference solution.

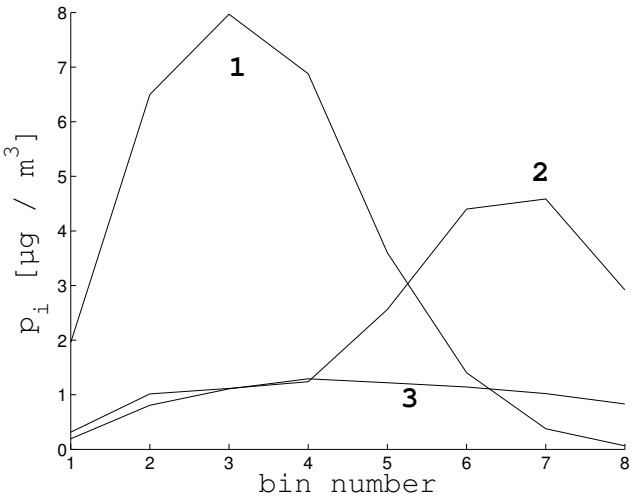

Figure 2. Reference solution, $t = 15$ min.

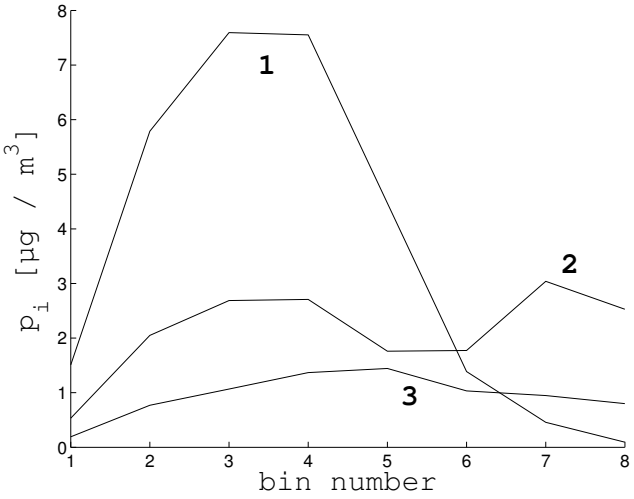

Figure 3. Reference solution, $t = 2.5$ h.

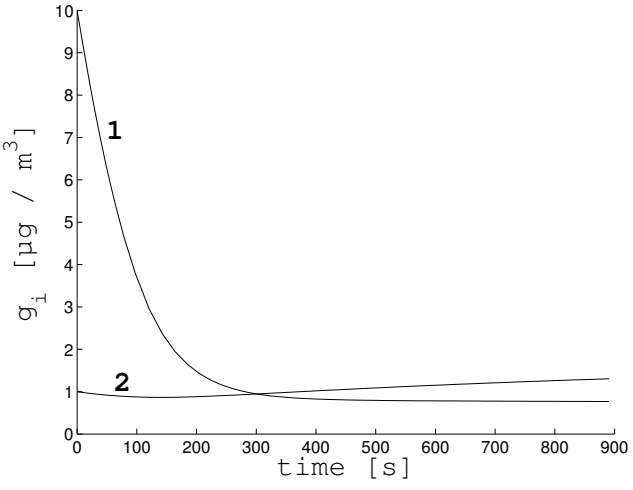

Figure 4. Time evolution of ambient gas concentrations

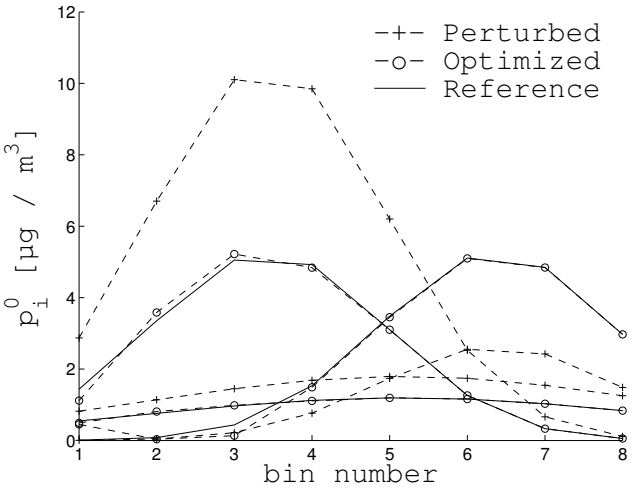

Figure 5. Case 1a.*i*: Simultaneous recovery of the initial distribution of all three species from an observation at the final time (15 min) using the discrete adjoint model.

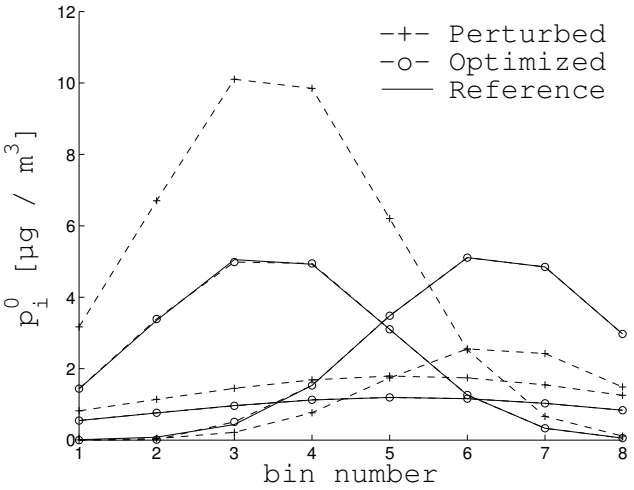

Figure 6. Case 1a.*ii*: Simultaneous recovery of the initial distributions of all three species from an observation at the final time (15 min) using the continuous adjoint model.

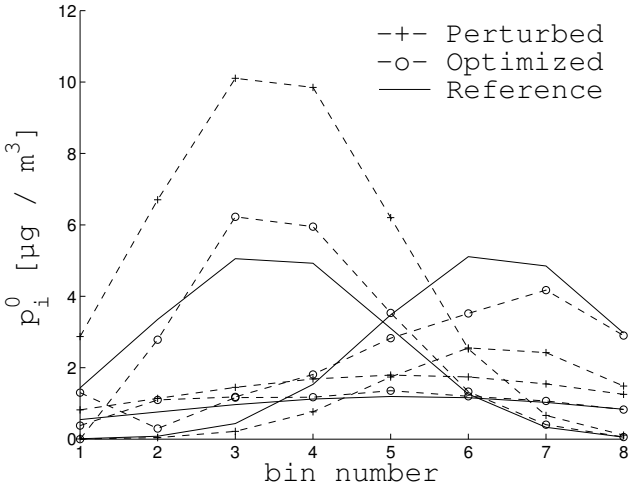

Figure 7. Case 1c.*i*: Simultaneous recovery of the initial distributions of all three species from 10 observations taken every 15 min using the discrete adjoint model.

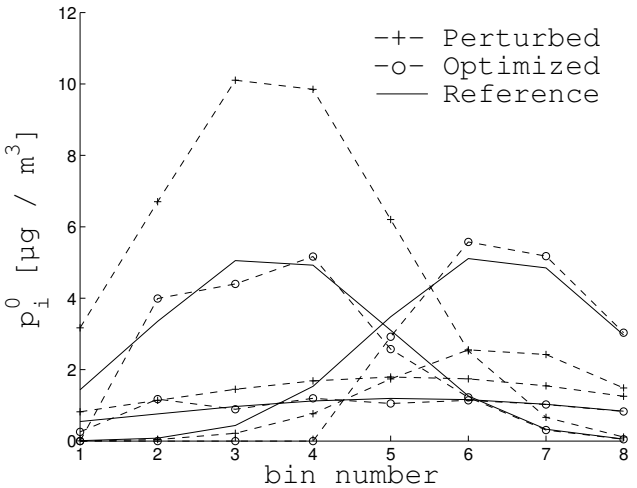

Figure 8. Case 1c.*ii*: Simultaneous recovery of the initial distributions of all three species simultaneously from 10 observations taken every 15 min using the continuous adjoint model.

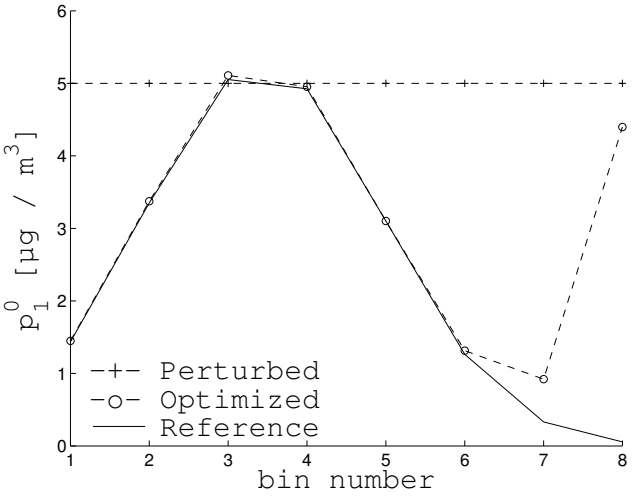

Figure 9. Case 3b: Recovery of the initial distribution of species 1 from an observation of species 2 and 3 at the final time (15 min).

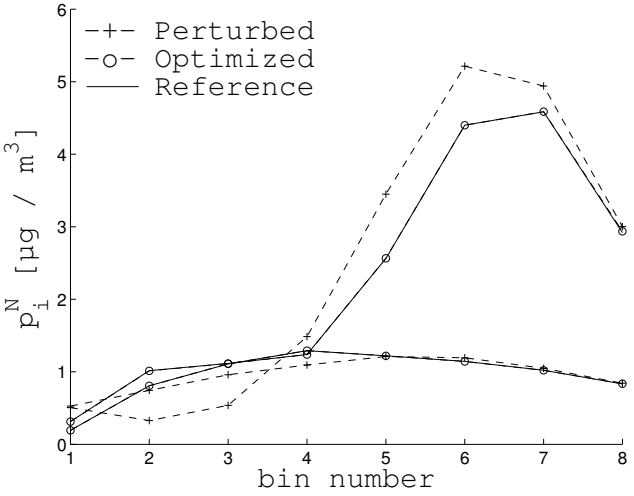

Figure 10. Case 3b: The converged profiles of species 2 and 3 after 15 min.

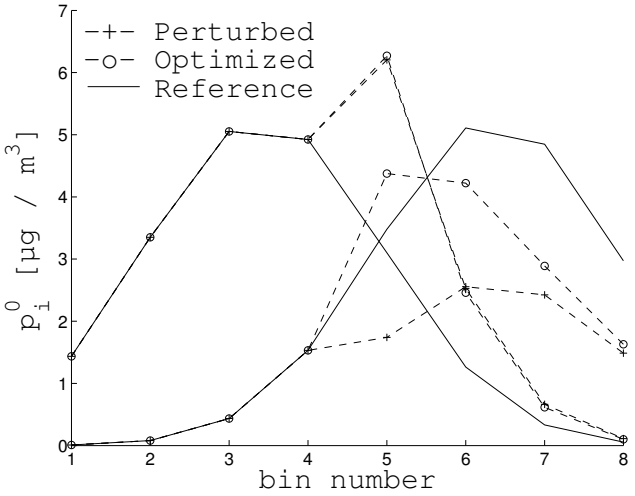

Figure 11. Case 4a: Simultaneous recovery of the initial distribution of all three species from an observations at the final time (15 min) of only bins 1-4. Species 3 omitted for clarity.

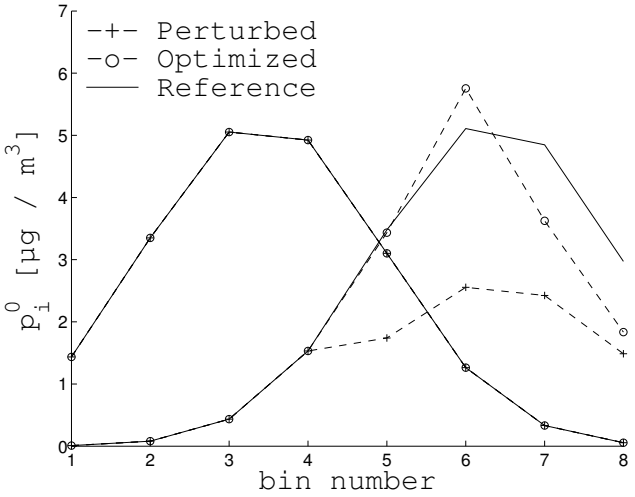

Figure 12. Case 4c: Recovery of the initial distribution of only species 2 from an observations at the final time (46) of only bins 1-4. Species 3 omitted for clarity.

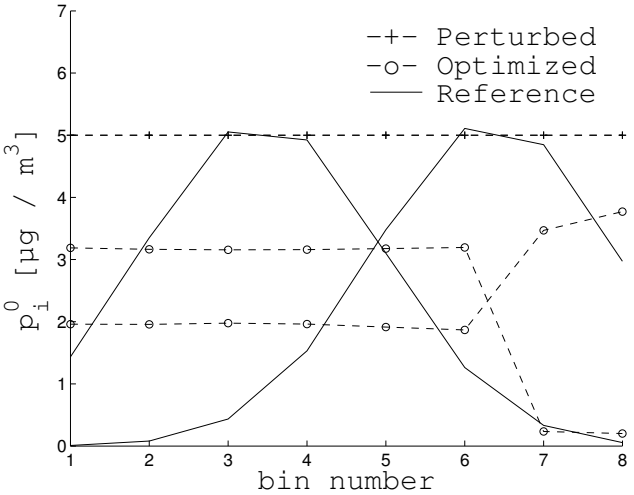

Figure 13. Case 4f: Simultaneous recovery of the initial distribution of all three species from observations of the total particulate concentration in bins 1 - 6 and bins 7 - 8 using a flat initial guess.

Table 1. Test problem specifications

Species	g_i [μg/m^3]	c_i° [μg/m^3]	p_i [μg/m^3]	$\bar{D}_p$ [μm][a]	σ[b]
1	10.0	1.0	20.0	0.3	2.8
2	1.0	10.0	20.0	2.3	2.8
3	0.0	0.0	10.0	1.0	10.0

[a] Mean particle diameter of a log-normal distribution

[b] Standard deviation

Table 2. Timing Statistics

Method	Total Time[a][s]	t_f [s]	t_b [s]	N_{cf}
	Case 1a			
Discrete	21	0.02	0.17	113
Continuous	10	0.05	0.02	148
	Case 1b			
Discrete	81	0.04	0.47	156
Continuous	58	0.24	0.10	172
	Case 1c			
Discrete	47	0.16	1.73	25
Continuous	30	0.48	0.20	45

[a] repressing output

Table 3. Conditions and Results of Assimilation Tests

CASE	RECOVER		OBSERVE			GUESS		R	
	p_i^0	c_i°	bins	species	time	p_i^0	c_i°	p_i^0	c_i°
1a.i	1-3	-	1-8	1-3	15/15	$\times$(2, 0.5, 1.5)	-	0.07	-
1a.ii	”	-	”	”	”	”	-	0.01	-
1b.i	”	-	”	”	40/40	”	-	0.19	-
1b.ii	”	-	”	”	”	”	-	0.26	-
1c.i	”	-	”	”	15/150	”	-	0.34	-
1c.ii	”	-	”	”	”	”	-	0.27	-
1d.i	”	-	”	”	150/150	”	-	0.21	-
1d.ii	”	-	”	”	”	”	-	0.68	-
2a	-	1-3	1-8	1-3	15/15	-	$\times 2, \times 5, +1$	-	0.00
2b	”	”	”	”	”	-	$+20, -10, +5$	-	0.00
3a	1	1	1-8	2-3	15/15	$\times$(2,-,-)	$\times$(10,-,-)	0.11	0.01
3b	”	”	”	”	”	5,-,-	”	0.49	0.02
3c	”	”	”	”	46/46	”	”	0.01	0.00
4a	1-3	-	1-4	1-3	15/15	$\times$(2, 0.5, 1.5)	-	0.84	-
4b	”	-	”	”	46/46	”	-	0.63	-
4c	2	-	”	”	”	$\times$(-,0.5,-)	-	0.46	-
4d	1-3	-	$\overline{12}, \overline{34}, \overline{56}, \overline{78}$	1-3	15/15	$\times$(2, 0.5, 1.5)	-	0.13	-
4e	”	-	$\overline{123456}, \overline{78}$	”	”	”	-	0.31	-
4f	”	-	”	”	”	5,5,5	-	0.31	-