Inverse modeling of aerosol dynamics: Condensational growth

D. K. Henze and J. H. Seinfeld

Department of Chemical Engineering, California Institute of Technology, Pasadena, California, USA

W. Liao and A. Sandu

Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

G. R. Carmichael

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

[1] 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 on the basis of 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, although this success is shown to depend upon the form of the initial estimates, the nature of the observations, and the length of the assimilation period. The results presented here provide a basis for the inverse modeling of aerosols in three-dimensional atmospheric chemical transport models. *INDEX TERMS:* 0305 Atmospheric Composition and Structure: Aerosols and particles (0345, 4801); 3260 Mathematical Geophysics: Inverse theory; 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; *KEYWORDS:* inverse modeling, data assimilation, aerosol dynamics

1. Introduction

[2] 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.

[3] 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 [*Lyster et al.*, 1997; *Khattatov et al.*, 2000; *Stajner et al.*, 2001; *Palmer et al.*, 2003a]. 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 other methods. As an alternative approach, the adjoint method was first suggested as an efficient technique for performing variational data assimilations in atmospheric transport models by *Marchuk* [1974]. Originating from the mathematics of systems optimization and control theory [*Cacuci*, 1981a, 1981b] and well established in the fields of fluid mechanics [*Pironneau*, 1974], meteorology [*Talagrand and Courtier*, 1987], and oceanog-

raphy [*Tziperman and Thacker*, 1989], the adjoint method has only been applied to CTMs relatively recently [*Fisher and Lary*, 1995; *Elbern et al.*, 1997; *Errera and Fonteyn*, 2001]. The treatment, while successful, has been limited to the assimilation and recovery of gas-phase species.

[4] The inclusion of detailed aerosol chemistry and physics has become requisite in atmospheric CTMs. Future implementation of four-dimensional variational analysis (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 (A. Sandu et al., Inverse modeling of aerosol dynamics using adjoints: Theoretical and numerical considerations, submitted to *Mathematics and Computers in Simulation*, 2004) (hereinafter referred to as Sandu et al., submitted manuscript, 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.

[5] 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 four-dimensional (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.

[6] 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 on the basis of 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.

[7] 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 transferring species. A significant fraction of organic aerosol particles is composed 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 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.

[8] 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.

[9] 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 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, and hence it is desirable to assess both tactics when introducing the adjoint method to a new field (Sandu et al., submitted manuscript, 2004).

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

[10] 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]

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

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

$$p(\mu, t) = \sum_{i=1}^{n} p_i(\mu, t),$$

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

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]

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

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.

[11] To solve equation (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. Inverse Problem

[12] The goal of inverse modeling is to estimate model parameters that 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 (equation (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

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

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

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

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 a priori (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 equation (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, as $\mathcal{J}$ depends on χ implicitly through the dependency of p_i on χ given by the forward model. In order to determine $\nabla_\chi\mathcal{J}$ an inverse model must be constructed that can calculate the derivative of the forward solution with respect to the model parameters.

3.1. Adjoint Method

[13] 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, but it also facilitates the simultaneous estimation of large numbers of parameters. One 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. A limitation of the adjoint method itself is that estimates from the solution of the inverse problem are subject to the same systematic and random errors present in the forward model. Unlike the Kalman filter approach, these factors cannot be treated explicitly. Although the method can be used to improve systematic error induced by model parameters, sound application is limited to models for which random errors in the forward solution are small, or at least well characterized.

[14] In sections 3.2 and 3.3 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 the variable dependencies and nonlinearities that are included in the discretized forward model. 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. Alternatively, to derive the continuous adjoint equations by hand, one must linearize the equations first, leading to gradients that can be highly approximate. 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.

[15] 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

[16] 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 equation (1) is

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

$$(5)$$

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"

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

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

[17] 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 equation (2) makes the $\partial H / \partial p_i$ 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

[18] 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].

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

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

$$(7)$$

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

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

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

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

Using the chain rule (in its transposed form), one can expand the right-hand side (RHS) of equation (9)

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

Evaluation of the RHS of equation (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.

[20] Defining the discrete adjoint variable as

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

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:

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

In this manner, the adjoint method is reduced to calculating the product $\{\partial [F_j(p_i^k, g_i^k)] / \partial p_i^{k-1}\}^T \lambda^k$ 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.

[21] Because of the nonlinear nature of $F_j(p_i^k, g_i^k)$ introduced by the dynamic time step and nonlinearities in the growth law, $\partial[F_j(p_i^k, g_i^k)]/\partial p_i^{k-1}$ 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 that minimize these types of computational demands (for example, *Elbern and Schmidt* [1999], or the distributed scheme implemented for a parallel model of A. Sandu et al. (Adjoint sensitivity analysis of regional air quality models, submitted to *Journal of Computational Physics*, 2003)); these techniques could be applied to the aerosol adjoint model as well.

4. Inverse Modeling of Aerosol Size Composition Dynamics

[22] 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. The cost function is then minimized using the L-BFGS-B 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 nonideal effects, Raoult's 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 mole fractions are equivalent to the mass fractions, and the growth rate can be written as

$$H_i = \frac{2\pi D_p D_i}{m\left[1 + (2\ell)/(\alpha D_p)\right]}\left(g_i - \frac{p_i}{\sum_{i=1}^{n} p_i} c_i^\circ\right). \quad (13)$$

[23] 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 $D_i = 1 \times 10^{-5}$ m^2/s. In the aerosol phase each species is initially lognormally distributed: Species 1 is located in the smaller bins, species 2 is located in the larger bins, and species 3 is located across all bins. The gas-phase concen-

Table 1. Test Problem Specifications[a]

Species	g_i, μg/m^3	c_i°, μg/m^3	p_i, μg/m^3	$\bar{D}_p$, μm	σ
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]Initial gas-phase concentrations g_i, pure component surface vapor concentrations c_i°, and parameters of the initial lognormal distribution: total concentration p_i, mean particle diameter $\bar{D}_p$, and standard deviation σ.

trations and pure component vapor concentrations are selected such that species 1 condenses and species 2 evaporates, while the third species is nonvolatile. Figures 1a–1c show the reference run at $t = 0$, 15 min, and 2.5 hours, respectively. Most of the progress toward an equilibrium distribution is made during the first 15 min. Figure 1d 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.

[24] 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 span ∼15 min.

[25] The continuous adjoint equation (5) for the forward model is solved using finite differences. Because of the nonlinearity of equation (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 4 times greater than that required to run the discrete model.

[26] 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. As the primary interest was investigation of formulation of the inverse modeling problem, we did not explore variations in the complexity of the aerosol distribution in order to keep the forward model consistent from case to case. 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.

[27] Table 2 summarizes the conditions and results of each of the cases considered. The "Recover" column lists which parameters were being assimilated; the numbers refer
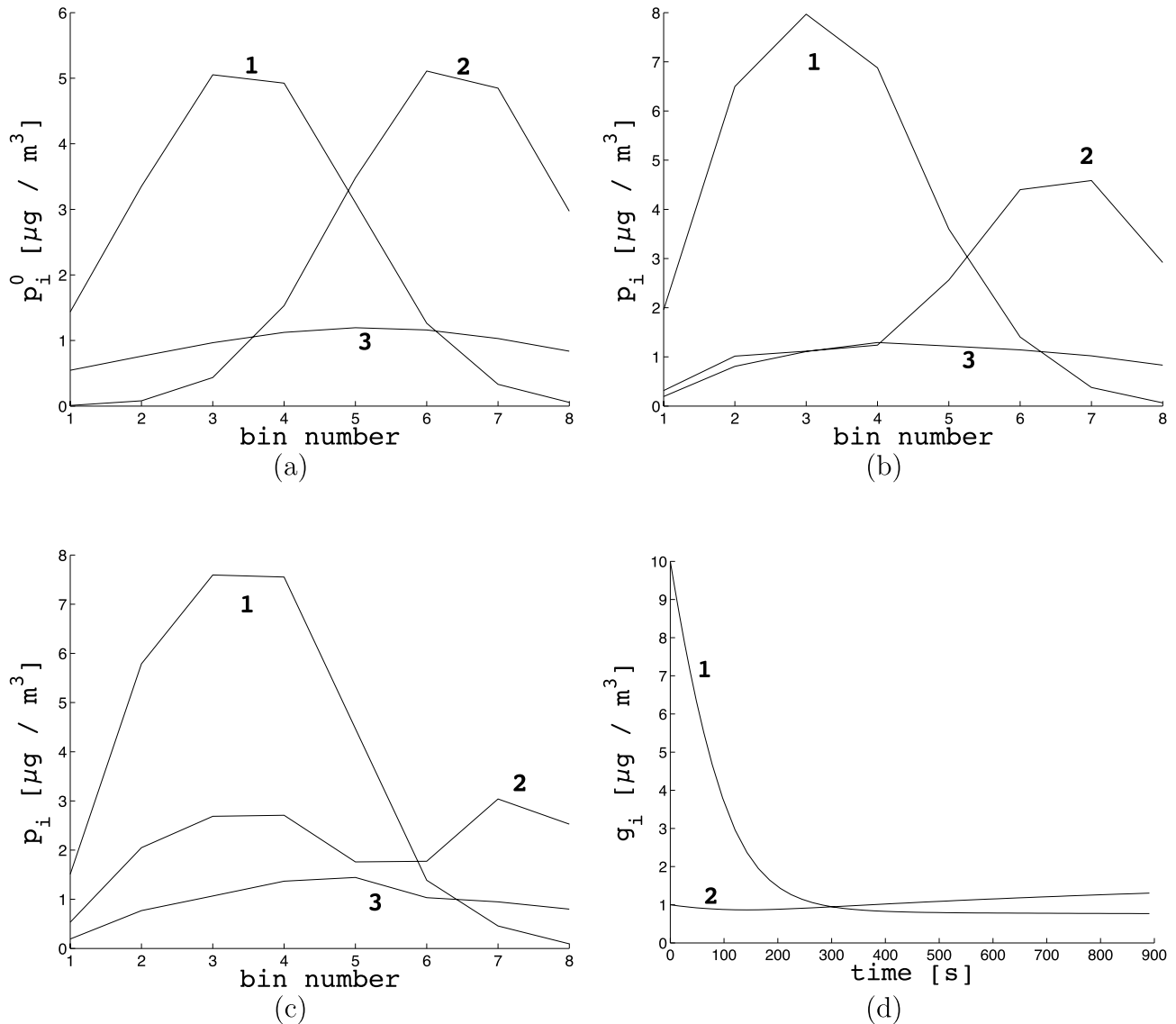
Figure 1. Forward model calculation (reference solution). Species 1 is condensing, species 2 is mostly evaporating, and species 3 is inert. Plotted are the aerosol size distributions at (a) $t = 0$, (b) $t = 15$ min, (c) $t = 2.5$ hours, and (d) the gas-phase concentrations as a function of time.

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, or c for the first, second, and third 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 scalar measure of the relative success of the data assimila-

tion is the percent of the error in the initial guess that is still present after optimization,

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

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.

[28] As the entire assimilation procedure depends critically upon the minimization of $\mathcal{J}$, it is worth digressing momentarily to discuss some features of the cost function that arise in inverse aerosol modeling. Consider the full cost function given in equation (4). Rigorous treatment of the cost function for the test problem would require generation of fictitious error covariance such that R_k and B can be defined. However, realistic values of R_k and B will be highly

Table 2. Conditions and Results of Assimilation Tests Using Notation Outlined in Section 4

Case	Recover p_i^0	Recover c_i°	Observe Bins	Observe Species	Observe Time	Guess p_i^0	Guess c_i°	R p_i^0	R c_i°
1a.i	1–3	-	1–8	1–3	15/15	×(2, 0.5, 1.5)	-	0.07	-
1a.ii	1–3	-	1–8	1–3	15/15	×(2, 0.5, 1.5)	-	0.01	-
1b.i	1–3	-	1–8	1–3	40/40	×(2, 0.5, 1.5)	-	0.19	-
1b.ii	1–3	-	1–8	1–3	40/40	×(2, 0.5, 1.5)	-	0.26	-
1c.i	1–3	-	1–8	1–3	15/150	×(2, 0.5, 1.5)	-	0.34	-
1c.ii	1–3	-	1–8	1–3	15/150	×(2, 0.5, 1.5)	-	0.27	-
1d.i	1–3	-	1–8	1–3	150/150	×(2, 0.5, 1.5)	-	0.21	-
1d.ii	1–3	-	1–8	1–3	150/150	×(2, 0.5, 1.5)	-	0.68	-
2a	-	1–3	1–8	1–3	15/15	-	×2, ×5, +1	-	0.00
2b	-	1–3	1–8	1–3	15/15	-	+20, −10, +5	-	0.00
3a	1	1	1–8	2–3	15/15	×(2,-,-)	×(10,-,-)	0.11	0.01
3b	1	1	1–8	2–3	15/15	5,-,-	×(10,-,-)	0.49	0.02
3c	1	1	1–8	2–3	46/46	5,-,-	×(10,-,-)	0.01	0.00
4a	1–3	-	1–4	1–3	15/15	×(2, 0.5, 1.5)	-	0.84	-
4b	1–3	-	1–4	1–3	46/46	×(2, 0.5, 1.5)	-	0.63	-
4c	2	-	1–4	1–3	"	×(-,0.5,-)	-	0.46	-
4d	1–3	-	$\overline{12},\overline{34},\overline{56},\overline{78}$	1–3	15/15	×(2, 0.5, 1.5)	-	0.13	-
4e	1–3	-	$\overline{123456},\overline{78}$	1–3	15/15	×(2, 0.5, 1.5)	-	0.31	-
4f	1–3	-	$\overline{123456},\overline{78}$	1–3	15/15	5,5,5	-	0.31	-

case dependent in any assimilation involving real data; hence they will be implemented herein less formally in order to focus on construction of the adjoint model in general. Within the twin experiment framework the observations can be considered to be exact and independent; hence R_k reduces to the identity matrix. For assimilation cases in which we limit ourselves to observations in only a subset of the species or bins, the corresponding diagonal element of R_k^{-1} will be zero. Since all the weight factors are then either 0 or 1, this can be equivalently represented by writing the summations in equation (4) over only the observed species/bins. In most cases considered, the cost function does not penalize departure from the background estimates since we know that the observations are correct while the initial guesses are wrong. Leaving out the first term of equation (4) is equivalent to letting B go to ∞. Exceptions to this arise in case 4, where we have reason to believe that the background estimate of the initial distribution is functionally more appropriate than the converged solution. For such cases, which are, in general, underdetermined, preconditioning of the cost function (i.e., including the penalty term $(1/2)(\chi - \chi_b)^T B^{-1}(\chi - \chi_b)$) may be appropriate.

[29] Finally, let us consider issues that arise for real aerosol inverse modeling. Even for inverse modeling studies of real systems, R_k and B are commonly taken to be diagonal [*Mendoza-Dominguez and Russell*, 2000, 2001]. Furthermore, it is often assumed that all elements of B are equal so that the entire matrix can be characterized by a single parameter, the so-called ridge regression parameter. For aerosol inverse modeling these assumptions may not be valid. Significant observational error covariance will exist between species that are not measured independently but are inferred on the basis of charge equilibrium (for example, nitrate concentrations are often inferred from the measured amounts of sulfate and ammonium). Furthermore, it will be likely that the background terms for some species (for example, sulfates) will be known with relatively small variance, while others will have very large variance (secondary organic aerosol (SOA)); hence B will likely not be simply a scalar multiple of the identity matrix. Overall, inverse aerosol problems are likely to be ill conditioned because of the model resolution in the size domain being much more refined than the observational resolution. One possible alternative that avoids having to introduce additional bias via R_k and B is to simply halt the optimization process before the cost function is completely minimized, as conjugate gradient methods will minimize along the largest regular vectors first.

4.1. Case 1: Recovery of Initial Distributions

[30] 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 three species being measured in all 8 bins and all the surface concentrations considered known. Cases 1a.i, 1b.i, and 1c.i used the discrete adjoint model while cases 1a.ii, 1b.ii, and 1c.ii used the continuous adjoint model. The reference, guessed, and optimized initial distributions for cases 1a.i and 1a.ii are shown in Figure 2. 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 cases 1b.i and 1b.ii (given in Table 2 but not plotted) 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 hours, but observations are still taken every ∼15 min. Figure 3 shows 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 (∼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 hours, the discrete model performs much better than the continuous model (case 1d; see Figure 4).
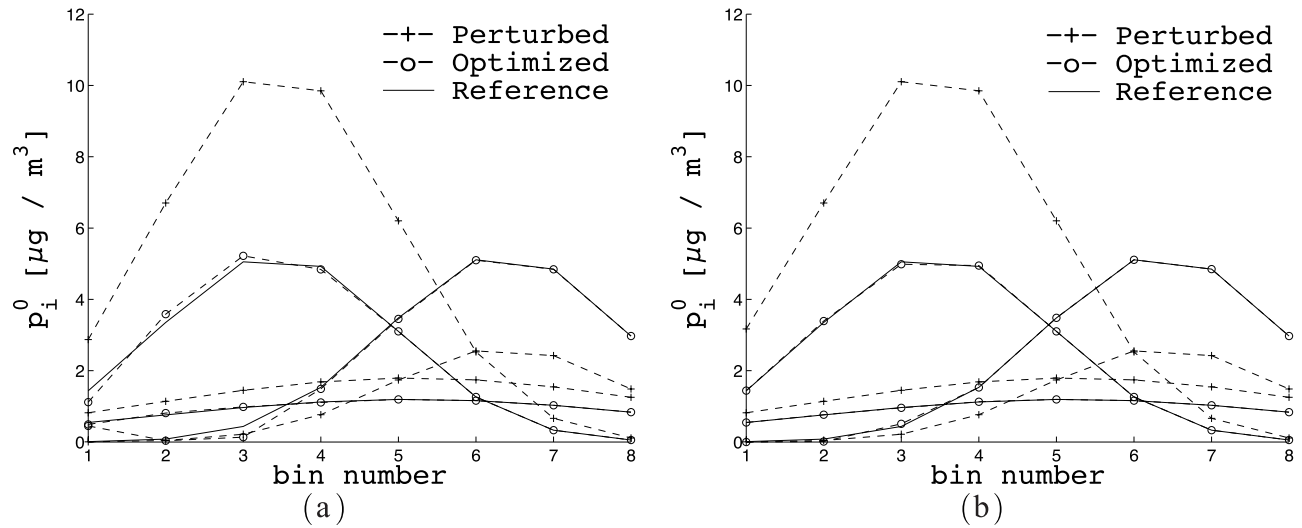
Figure 2. Case 1a: Simultaneous recovery of the initial distribution of all three species from an observation at the final time (15 min) using (a) the discrete adjoint model (case 1a.i) and (b) the continuous adjoint model (case 1a.ii). The continuous model performs slightly better, primarily in the lower bins for species 1 and 3.

[31] While even the longer assimilation periods considered here are much shorter (temporally) than can be expected for assimilations involving actual data and real species, this is only an artifact of the arbitrary environmental conditions used for this test case. A more relevant (and general) measure of the assimilation period is the number of numerical integration steps taken between observations. Examining assimilation intervals of 50, 150, and 500 time steps over a length of up to 500 steps covers a wide range of potential models and sets of observational data. For example, a local urban aerosol model that is run for a few days typically employs time steps on the order of minutes and is compared to observations taken during intervals on the order of hours. For large-scale regional models that are run for months, time

steps are typically on the order of hours, and observation intervals are on the order of days.

[32] In order to test the validity of the tangent linear approximations inherent in the adjoint model over the assimilation period, the gradient was also calculated using finite differences with a perturbation of 10^{-9}. Figure 5 shows the relative reduction in the cost function after the first optimization step, $\Delta = [(\mathcal{J}^0 - \mathcal{J}^1)/\mathcal{J}^0] \times 100\%$, as a function of the total number of steps in the assimilation period. The adjoint gradient becomes increasingly inaccurate beyond $\sim$250 steps. As the aerosol distribution approaches equilibrium, the assimilation becomes increasingly difficult.

[33] In addition to comparing the ability of the two types of adjoint models to recover the initial distributions, it is
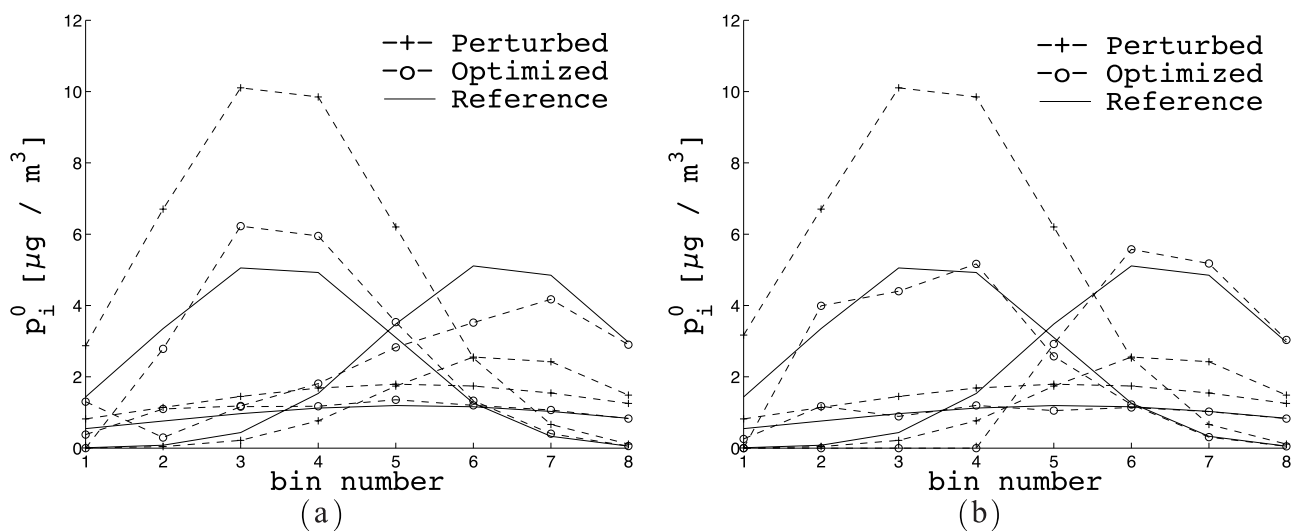

Figure 3. Case 1c: Simultaneous recovery of the initial distributions of all three species from 10 observations taken every 15 min over the course of 2.5 hours using (a) the discrete adjoint model (case 1c.i) and (b) the continuous adjoint model (case 1c.ii). Overall performance is similar between the two approaches.
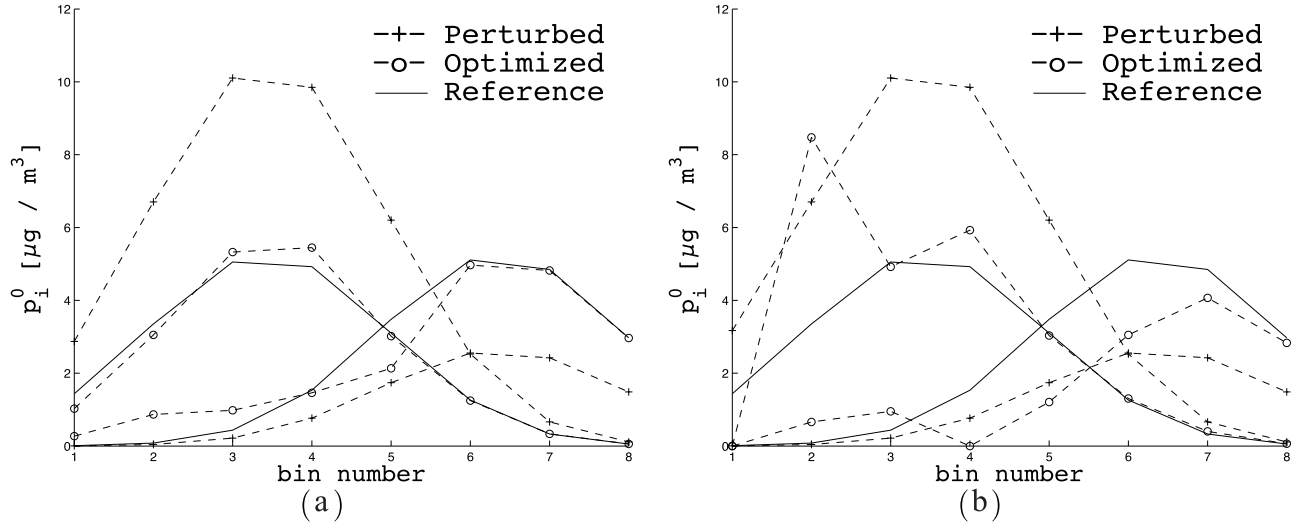
Figure 4. Case 1d: Simultaneous recovery of the initial distributions of all three species using only one observation after 2.5 hours. Results are shown for (a) the discrete adjoint model (case 1d.i) and (b) the continuous adjoint model (case 1d.ii), from which the superior performance of the former for this case is quite evident. Species 3 is omitted from the plots for clarity.

important to compare the computational expense of each approach. The total optimization expense ratio is η_{tot}, where

$$\eta_{tot} = \frac{\text{total computational time(discrete)}}{\text{total computational time(continuous)}}. \quad (15)$$

Let t_f be the time for the forward calculation, t_b be the time for the backward calculation, and N_J be the number of cost function evaluations during minimization. Noting that the total computational time for each test is approximately equal
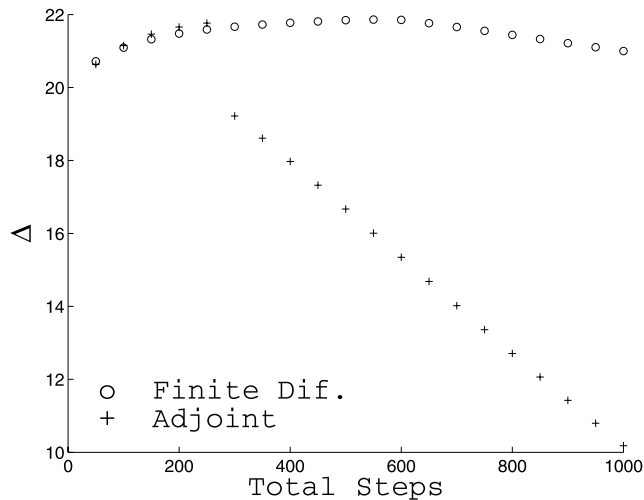
to $N_J \times (t_f + t_b)$, this ratio can be further broken down into a product of ratios that are fairly consistent in magnitude throughout each test and whose smallness indicates the degree to which the discrete calculation is preferable.

$$\eta_{tot} = \eta_J \eta_f \eta_b, \quad (16)$$

where

$$\eta_J = \frac{N_J(d)}{N_J(c)},$$

$$\eta_f = \frac{t_f(d)}{t_f(c)},$$

$$\eta_b = \frac{\{1 + [t_b(d)]/[t_f(d)]\}}{\{1 + [t_b(c)]/[t_f(c)]\}}.$$

The values of each ratio are given in Table 3. Considering t_b/t_f to be a measure of the efficiency of the backward calculation with respect to the forward calculation, the large values of η_b indicate that the backward calculation is much more efficient for the continuous model than the discrete model. However, as indicated by η_J, 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.

Figure 5. Simultaneous recovery of the initial distributions of all three species using only one observation at the final time step (x axis). Plotted is the relative reduction of the cost function after the first minimization step, $\Delta = (\mathcal{J}^0 - \mathcal{J}^1)/\mathcal{J}^0 \times 100\%$, as a function of the assimilation period. The accuracy of the gradient computed using the adjoint method (plus symbols) is seen to decay in comparison to that from the finite difference calculation (open circles) as the distribution approaches equilibrium.

Table 3. Timing Ratios for Comparing the Discrete Adjoint Model to the Continuous Adjoint Model, as Defined by Equation (16)[a]

Case	η_{tot}	η_J	η_f	η_b
1a	2.2	0.8	0.3	8.3
1b	1.4	0.9	0.2	8.3
1c	1.6	0.6	0.3	8.2

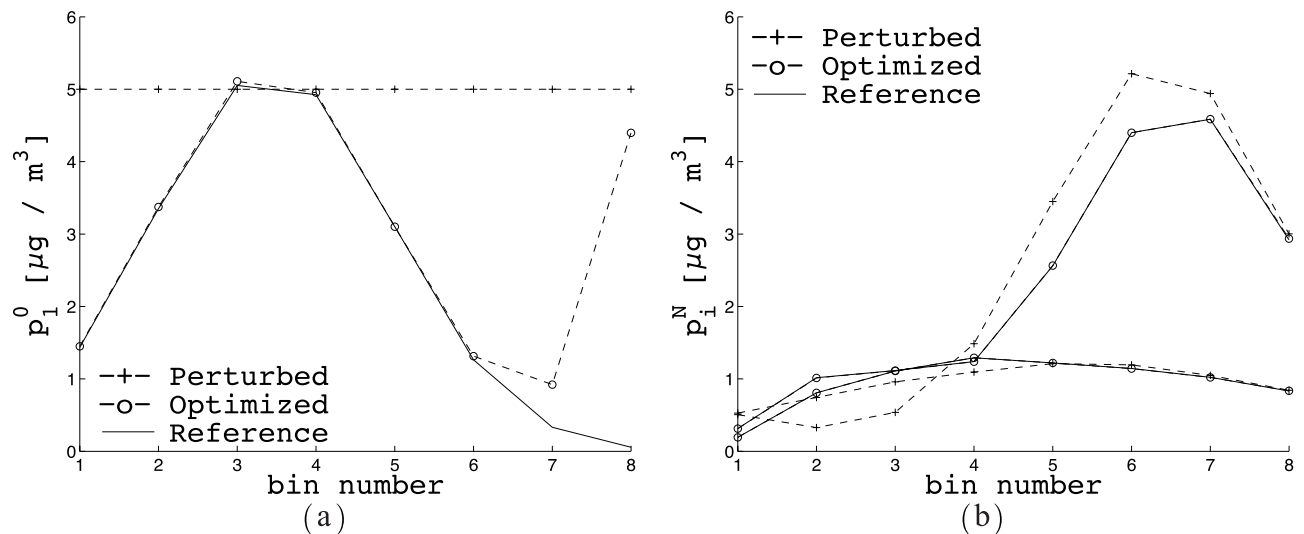[a]Values less than 1 indicate the discrete model is preferable.

Figure 6. Case 3b: Recovering the initial distribution of species 1 from an observation of species 2 and 3 at the final time (15 min). Shown are the aerosol size distributions (a) of species 1 at $t = 0$ and (b) of species 2 and 3 at $t = 15$ min.

[34] 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_f will decrease and η_b 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_f. In this simple model, η_f is $\sim 1/4$ because the average time step taken in the discrete model is ~ 4 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_f, 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 utilize 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

[35] 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 $c_1^\circ = 20 \ \mu g/m^3$, 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

[36] The third scenario addresses a common question encountered in aerosol measurement: On the basis of accurate information on 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

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

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. This being said, it is interesting to note that it is not necessary to precondition the cost function in order to converge to the correct distribution because the problem is overdetermined in this case.

[37] 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 $\mu g/m^3$ 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 6a). To understand why this would
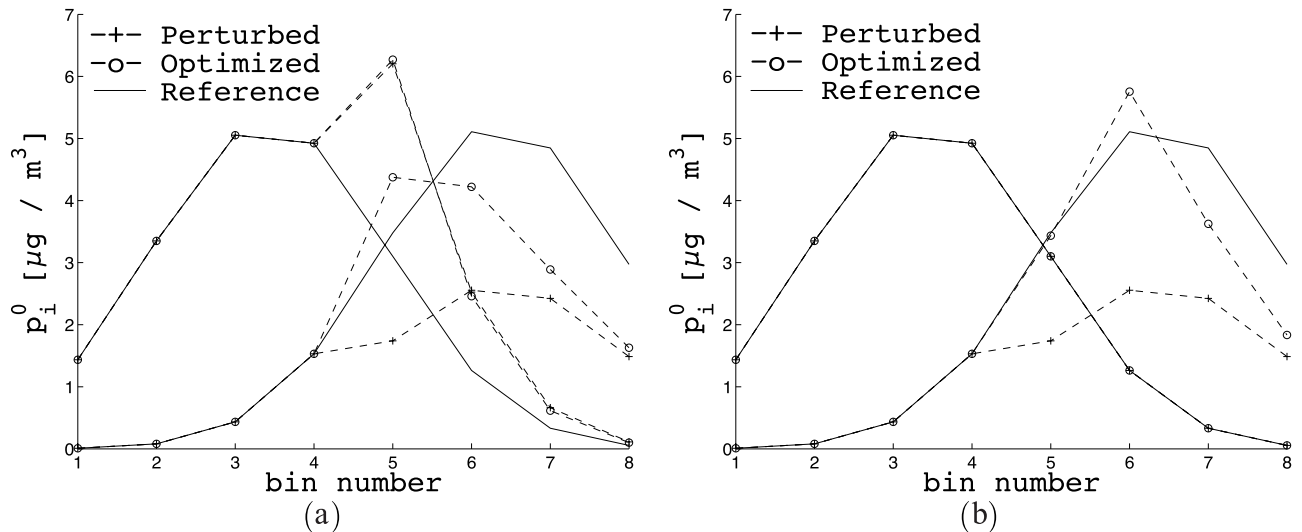
Figure 7. Case 4: Recovering initial distributions using only data from the smaller four size bins. The results for case 4a (recovery of all three initial distributions simultaneously) and case 4c (recovery of only the initial distribution of species 2) are shown in Figures 7a and 7b, respectively. Species 3 is omitted from the plots for clarity.

be the case, it is useful to recall that the driving term for the discrete adjoint model is $\partial[\mathcal{J}(p_i^N)]/\partial p_i^N$. 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 6b, and we see that 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), as indicated by Figure 5.

4.4. Case 4: Recovery From Partial Distributions

[38] 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,

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

On the basis of this information the initial concentrations in the larger bins were determined and are shown in Figure 7a. 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 7b). Tests 4d–4f 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

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

where $\bar{j}$ is the index of the $\bar{s}$ lumped bins. In case 4d each pair of adjacent bins is averaged, while in case 4e the observed distribution is of only two bins: one that contains particles whose diameter is smaller than 2.76 μm and one that contains 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 8. In order to avoid optimizing to erroneously smooth or jagged distributions, the solution can be constrained by including the penalty
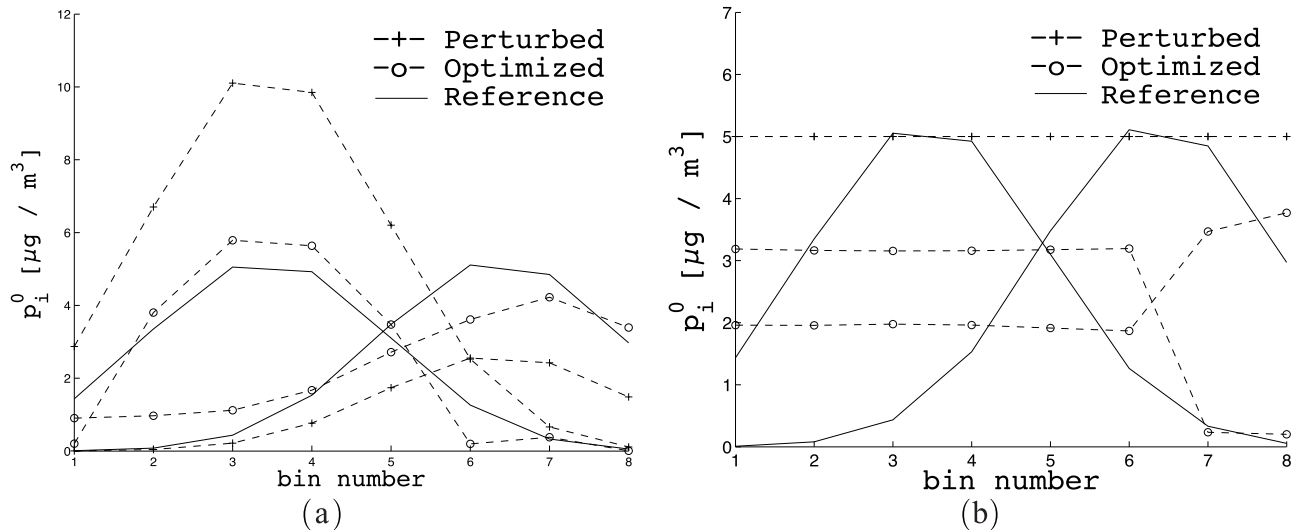
Figure 8. Case 4: Simultaneous recovery of the initial distribution of all three species from observations of the total particulate concentration in bins 1–6 and bins 7–8 using (a) a lognormal initial guess (case 4e) and (b) a flat initial guess (case 4f). Species 3 is omitted from the plot for clarity.

term in the cost function. While this approach biases the final estimate, this may be appropriate when there is sufficient information known about the true distribution to quantitatively estimate the error covariance matrix B of the initial guess.

5. Conclusions

[39] 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.

[40] 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 using the 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.

[41] 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 when observations were known in only a subset of the size range or when the initial estimates were exceptionally poor. 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 given ample observations and reasonable initial estimates, the adjoint method can be used to recover information about a dynamic, size-resolved, and chemically resolved aerosol distribution under a variety of conditions.

Appendix A: Derivation of Continuous Adjoint Equation

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

$$
\mathcal{J} = \int_{t^0}^{T} \int_{0}^{\infty} J_0[p_1(\mu, t), p_2(\mu, t), \dots, p_n(\mu, t)] \, d\mu \, dt
$$
$$
- \sum_{i=1}^{n} \int_{t_0}^{T} \int_{0}^{\infty} \lambda_i(\mu, t) (\text{LHS}_{p_i} - \text{RHS}_{p_i}) \, d\mu \, dt, \quad (A1)
$$

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

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

where $t^k \in \Omega$ and Ω is the set of discrete time points t^k for which data are known. Taking the variation of equation (A1), we get

$$\delta \mathcal{J} = \int_{t^0}^{T} \int_0^\infty \sum_{i=1}^n \frac{\partial J_0}{\partial p_i} \delta p_i(\mu, t) \, d\mu \, dt$$
$$- \int_{t^0}^{T} \int_0^\infty \sum_{i=1}^n \delta \lambda_i(\mu, t) \left(\text{LHS}_{p_i} - \text{RHS}_{p_i} \right) d\mu \, dt$$
$$- \int_{t^0}^{T} \int_0^\infty \sum_{i=1}^n \lambda_i(\mu, t) \delta \left(\text{LHS}_{p_i} - \text{RHS}_{p_i} \right) d\mu \, dt. \quad (A3)$$

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

$$\delta \mathcal{J} = \int_{t^0}^{T} \int_0^\infty \sum_{i=1}^n \frac{\partial J_0}{\partial p_i} \delta p_i(\mu, t) \, d\mu \, dt$$
$$- \int_{t^0}^{T} \int_0^\infty \sum_{i=1}^n \delta \lambda_i(\mu, t) \left(\text{LHS}_{p_i} - \text{RHS}_{p_i} \right) d\mu \, dt$$
$$- \int_{t^0}^{T} \int_0^\infty \sum_{i=1}^n \lambda_i(\mu, t) \delta \left[\frac{\partial p_i(\mu, t)}{\partial t} \right.$$
$$\left. - H_i(\mu, p_1, p_2, \ldots, p_n, t) p(\mu, t) + \frac{1}{3} \frac{\partial}{\partial \mu} (H p_i) \right] d\mu \, dt. \quad (A4)$$

Then we can rewrite equation (A4) as

$$\delta \mathcal{J} = \sum_{i=1}^n \int_{t^0}^{T} \int_0^\infty \delta p_i(\mu, t) \frac{\partial J_0(\mu, p, t)}{\partial p_i} \, d\mu \, dt$$
$$- \sum_{i=1}^n \int_{t^0}^{T} \int_0^\infty \delta \lambda_i(\mu, t) \left(\text{LHS}_{p_i} - \text{RHS}_{p_i} \right) d\mu \, dt$$
$$- \sum_{i=1}^n \int_{t^0}^{T} \int_0^\infty \lambda_i(\mu, t) \frac{\partial [\delta p_i(\mu, t)]}{\partial t} \, d\mu \, dt$$
$$+ \sum_{i=1}^n \int_{t^0}^{T} \int_0^\infty \lambda_i(\mu, t) H_i(\mu, p, t) \sum_{j=1}^n \delta p_j(\mu, t) \, d\mu \, dt$$
$$+ \sum_{i=1}^n \int_{t^0}^{T} \int_0^\infty \lambda_i(\mu, t) \sum_{j=1}^n \frac{\partial H_i(\mu, p, t)}{\partial p_j} \delta p_j p(\mu, t) \, d\mu \, dt$$
$$- \frac{1}{3} \sum_{i=1}^n \int_{t_0}^{T} \int_0^\infty \lambda_i(\mu, t)$$
$$\cdot \frac{\partial}{\partial \mu} \left[\sum_{j=1}^n \frac{\partial H(\mu, p, t)}{\partial p_j} \delta p_j p_i + H(\mu, p, t) \delta p_i \right] d\mu \, dt. \quad (A5)$$

If we choose the final condition $\lambda(\mu, T) = 0$ and integrate the third term on the right-hand side of equation (A5) by parts, this term becomes

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

Likewise, letting $\lambda_i(0, t) = 0$, $p_i(+\infty, t) = 0$, the sixth term on the right-hand side of equation (A5) can be written as

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

If $p(\mu, t)$ is the solution of equation (1), $\text{LHS}_{p_i} - \text{RHS}_{p_i} = 0$, then

$$\delta \mathcal{J} = \sum_{i=1}^n \int_{t^0}^{T} \int_0^\infty \delta p_i(\mu, t) \frac{\partial J_0(\mu, t)}{\partial p_i} \, d\mu \, dt$$
$$+ \sum_{i=1}^n \int_0^\infty \lambda_i(\mu, t^0) \delta p_i(\mu, t^0) \, d\mu$$
$$+ \sum_{i=1}^n \int_{t^0}^{T} \int_0^\infty \frac{\partial \lambda_i(\mu, t)}{\partial t} \delta p_i(\mu, t) \, d\mu \, dt$$
$$+ \sum_{i=1}^n \int_{t^0}^{T} \int_0^\infty \lambda_i(\mu, t) H_i(\mu, p, t) \sum_{j=1}^n \delta p_i(\mu, t) \, d\mu \, dt$$
$$+ \sum_{i=1}^n \int_{t^0}^{T} \int_0^\infty \lambda_i(\mu, t) \sum_{j=1}^n \frac{\partial H_i(\mu, p, t)}{\partial p_j} p(\mu, t) \delta p_j \, d\mu \, dt$$
$$+ \frac{1}{3} \sum_{i=1}^n \int_{t_0}^{T} \int_0^\infty \frac{\partial \lambda_i(\mu, t)}{\partial \mu} \sum_{j=1}^n \frac{\partial H(\mu, p, t)}{\partial p_j} \delta p_j p_i \, d\mu \, dt$$
$$+ \frac{1}{3} \sum_{i=1}^n \int_{t_0}^{T} \int_0^\infty \frac{\partial \lambda_i(\mu, t)}{\partial \mu} H(\mu, p, t) \delta p_i \, d\mu \, dt. \quad (A8)$$

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

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

References

Bott, A. (1989), A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes, *Mon. Weather Rev.*, *117*, 1006–1016.

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

Cacuci, D. G. (1981a), Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach, *J. Math. Phys. N. Y.*, *22*(12), 2794–2802.

Cacuci, D. G. (1981b), Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses, *J. Math. Phys. N. Y.*, *22*(12), 2803–2812.

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

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

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

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

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

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

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

Gelbard, F., and J. H. Seinfeld (1980), Simulation of multicomponent aerosol dynamics, *J. Colloid Interface Sci.*, *78*(2), 485–501.

Gelbard, F., Y. Tambour, and J. H. Seinfeld (1980), Sectional representation for simulating aerosol dynamics, *J. Colloid Interface Sci.*, *76*(2), 541–556.

Giering, R., and T. Kaminski (1998), Recipes for adjoint code construction, *Trans. Math. Software*, *24*(4), 437–474.

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

Giles, M. B., M. C. Duta, J.-D. Müller, and N. A. Pierce (2003), Algorithm developments for discrete adjoint methods, *AIAA J.*, *41*(2), 198–205.

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

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

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

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

Marchuk, G. (1974), *Numerical Solution of the Problems of the Dynamics of the Atmosphere and the Ocean* (in Russian), Gidrometeoizdat, St. Petersburg, Russia.

Mendoza-Dominguez, A., and A. G. Russell (2000), Iterative inverse modeling and direct sensitivity analysis of a photochemical air duality model, *Environ. Sci. Technol.*, *34*(23), 4974–4981.

Mendoza-Dominguez, A., and A. G. Russell (2001), Emission strength validation using four-dimensional data assimilation: Application to primary aerosol and precursors to ozone and secondary aerosol, *J. Air Waste Manage. Assoc.*, *51*, 1538–1550.

Meng, Z., D. Dabdub, and J. H. Seinfeld (1998), Size-resolved and chemically resolved model of atmospheric aerosol dynamics, *J. Geophys. Res.*, *103*(D3), 3419–3435.

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 (2003a), Inverting for emissions of carbon monoxide from Asia using aircraft observations over the western Pacific, *J. Geophys. Res.*, *108*(D21), 8828, doi:10.1029/2003JD003397.

Palmer, P. I., D. J. Jacob, L. J. Mickley, D. R. Blake, G. W. Sachse, H. E. Fuelberg, and C. M. Kiley (2003b), Eastern Asian emissions of anthropogenic halocarbons deduced from aircraft concentration data, *J. Geophys. Res.*, *108*(D24), 4753, doi:10.1029/2003JD003591.

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

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

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

Song, C. H., and G. R. Carmichael (2001), A three-dimensional modeling investigation of the evolution processes of dust and sea-salt in east Asia, *J. Geophys. Res.*, *106*(D16), 18,131–18,154.

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

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

Tziperman, E., and W. C. Thacker (1989), An optimal-control/adjoint-equations approach to studying the oceanic general circulation, *J. Phys. Oceanogr.*, *19*, 1471–1485.

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

Yanenko, N. N. (1971), *The Method of Fractional Steps*, Springer-Verlag, New York.

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

Zhu, C., R. H. Byrd, P. Lu, and J. Nocedal (1994), L-BFGS-B: A limited memory FORTRAN code for solving bound constrained optimization problems, technical report, Dep. of Electr. Eng. and Comput. Sci., Northwest. Univ., Evanston, Ill.

————————

G. R. Carmichael, Center for Global and Regional Environmental Research, University of Iowa, Iowa City, IA 52242, USA. (gcarmich@icaen.uiowa.edu)

D. K. Henze and J. H. Seinfeld, Department of Chemical Engineering, California Institute of Technology, Mail Code 210-41, 1200 E. California Boulevard, Pasadena, CA 91125, USA. (daven@its.caltech.edu; seinfeld@caltech.edu)

W. Liao and A. Sandu, Department of Computer Science, Virginia Polytechnic Institute and State University, 660 McBryde Hall (0106), Blacksburg, VA 24061, USA. (liao@cs.vt.edu; sandu@cs.vt.edu)