Inverse Modeling of NO_x and NH_3 Precursor Emissions Using the Adjoint of CMAQ

Research Prelim Report

Matthew Turner
Advisor: **Professor Daven Henze**
Department of Mechanical Engineering, University of Colorado

April 22, 2010

1 Abstract

Air quality models are utilized by the U.S. EPA to develop emission control regulations. As such, air quality models are critical tools for improving air quality, yet they contain large amounts of uncertainty. Developing a constraint on emissions using the 4D-Variational data assimilation technique is one method of reducing the uncertainty in sources of aerosols. In an effort to provide more accurate predictive capabilities the current Community Multiscale Air Quality (CMAQ) adjoint model will be updated to include an adjoint of aerosol dynamics. This will result in the first ever air quality adjoint that includes an adjoint of aerosol size distribution. In order to determine the influences of specific sources of inorganic $PM_{2.5}$, NO_x and NH_3 measurements from the OMI and TES instruments aboard the Aura satellite will be assimilated using the 4D-Var method in the CMAQ model.

2 Motivation

Since its inception in 1970, the U.S. Environmental Protection Agency (EPA) has been tasked with the protection of human health and the environment. Title I of the Clean Air Act dictates that the EPA is in charge of setting National Ambient Air Quality Standards (NAAQS) for air pollutants that "cause or contribute to air pollution which may reasonably be anticipated to endanger public health or welfare..." [*42* U.S.C.,§*7408*(a)(1)(A)]. EPA's role has broadened to include development of NAAQS, evaluation of the cost and effectiveness of emissions control policies, and monitoring of NAAQS exceedences. One major aspect of this process is the use of air quality models to estimate current and future distributions of fine particulate matter ($PM_{2.5}$).

Particulate Matter (PM) is an air pollutant consisting of a mixture of solid and liquid particles suspended in the air. Knowledge of PM concentrations is important for many reasons, two of which are that PM has an adverse effect on human health, and PM also has an effect on climate change. Many atmospheric chemical transport models (CTM) have been modified recently to include aerosols, however the models still contain considerable uncertainty. There are many different methods to decrease uncertainty in the models, such as developing a better understanding of the physical processes, data assimilation, and using an ensemble of models. This project focuses on the 4D-Variational technique, a sophisticated approach to data assimilation, and how it pertains to atmospheric CTMs, more specifically aerosol modeling. The 4D-Var technique requires the development of an adjoint model for use in inverse modeling of a currently existing CTM.

Effects of PM on Climate Change Aerosols have both a direct and indirect effect on radiative forcing (a perturbation in the radiative energy budget of the Earth, caused by an external source [37]). Aerosols have a direct radiative forcing because they both scatter and absorb solar and infrared radiation in the atmosphere [29].

Indirect radiative forcing from aerosols is caused by changes in cloud properties that are caused by aerosols. Aerosols alter the efficiency at which liquid water, ice and mixed-phase clouds form and precipitate, which alters the properties of the clouds [29] . There is a set amount of water available for clouds. The water can form large droplets within the

clouds, which causes precipitation (a major removal mechanism for aerosols). The addition of PM into the atmosphere causes the water to condense onto the particles. This results in more, but smaller droplets in the clouds, which increases the cloud albedo [38]. In addition to increasing the albedo, this effect tends to decrease the chance of precipitation. If precipitation is suppressed, this results in excess water remaining in the atmosphere.

While aerosols in general can cause both a "cooling" and a "warming" effect in the atmosphere, the overall combination of both direct and indirect effects results in the cooling of the atmosphere.

Health Effects of PM Exposure to PM in ambient air has been shown to be a cause of various adverse health effects, both short-term and long-term. Some adverse health effects include impaired pulmonary function, adverse effects on the cardiovascular system, and perhaps most severe, a one or more year reduction in the life expectancy of the average population [30]. Many studies have been performed on the effects of long-term exposure to PM on mortality and morbidity (Figure 1).

Increases in fine PM and sulfate have been shown to increase mortality, a majority of which were related to an increase in cardiovascular mortality. The Health Effects Institute (HEI) performed an extension of a previous American Cancer Society (ACS) study and discovered significant increases in the relative risk of cardiopulmonary and lung cancer deaths, as well as deaths from all causes resulting from exposure to $PM_{2.5}$ [30]. Recent studies also suggest that long-term exposure to $PM_{2.5}$ is most strongly associated with mortality due to

Study	PM[1] metric[a]	Total mortality		Cardiopulmonary mortality		Lung cancer mortality	
		Excess RR[b]	95% CI (%)	Excess RR	95% CI (%)	Excess RR	95% CI (%)
Six City *(6,11)*	$PM_{2.5}$	13%	4.2 – 23	18%	6.0 – 32	18%	–11 – 57
Six City new *(6)*	$PM_{2.5}$	14%	5.4 – 23	19%	6.5 – 33	21%	–8.4 – 60
ACS *(6)*	$PM_{2.5}$	6.6%	3.5 – 9.8	12%	6.7 – 17	1.2%	–8.7 – 12
ACS new *(6)*	$PM_{2.5}$	7.0%	3.9 – 10	12%	7.4 – 17	0.8%	–8.7 – 11
ACS new *(6)*	$PM_{15-2.5}$	0.4%	–1.4 – 2.2	0.4%	–2.2 – 3.1	–1.2%	–7.3 – 5.1
ACS new *(6)*	$PM_{10/15}$	4.1%	0.9 – 7.4	7.3%	3.0 – 12	0.8%	–8.1 – 11
ACS new *(6)*	$PM_{10/15}$ SSI[c]	1.6%	–0.8 – 4.1	5.7%	2.5 – 9.0	–1.6%	–9.1 – 6.4
ACS extended *(5)*	$PM_{2.5}$ 1979–1983	4.1%	0.8 – 7.5	5.9%	1.5 – 10	8.2%	1.1 – 16
ACS extended *(5)*	$PM_{2.5}$ 1999–2000	5.9%	2.0 – 9.9	7.9%	2.3 – 14	12.7%	4.1 – 22
ACS extended *(5)*	$PM_{2.5}$ average	6.2%	1.6 – 11	9.3%	3.3 – 16	13.5%	4.4 – 23
AHSMOG *(2)*[d]	$PM_{10/15}$	2.1%	–4.5 – 9.2	0.6%	–7.8 – 10	81%	14 – 186
AHSMOG *(12)*[e]	$PM_{2.5}$	8.5%	–2.3 – 21	23%	–3.0 – 55	39%	–21 – 150
Veterans Administration[f]	$PM_{2.5}$	–10.0%	–15 – –4.6				

[a] Increments are 10 μg/m³ for $PM_{2.5}$ and 20 μg/m³ for $PM_{10/15}$.
[b] Excess RR (percentage excess relative risk) = 100 × (RR–1), where the RR has been converted from the highest-to-lowest range to the standard increment (10 or 20) by the equation RR = exp(log(RR for range) × /range).
[c] PM measured with size-selective inlet (SSI) technology. The other PM measurements in ACS new *(6)* were based on dichotomous sampler with 15-μm and 2.5-μm cut-off points.
[d] Pooled estimate for males and females.
[e] Using two-pollutant (fine- and coarse-particle) models; males only.
[f] Males only, exposure period 1979–1981, mortality 1982–1988 (from Table 7 in Lipfert et al. *(13)*).
Source: US Environmental Protection Agency *(10)*.

Figure 1: Comparison of excess relative risk for mortality from American cohort studies [30]

ischemic heart disease, dysrhythmias, heart failure, and cardiac arrest. For these causes of death, an increased $PM_{2.5}$ exposure of 10 $\mu g/m^3$ was associated with an 8-18% increase in death [32].

The HEI suggested that coarse particles were not significantly associated with mortality [24, 31]. However, other studies, in particular the Adventist Health and Smog study [1], found that exposure to PM_{10} has significant effects on nonmalignant respiratory death and lung cancer deaths. A European study in the Netherlands [23] suggested that exposure to traffic-related air pollution led to an increase in cardiopulmonary mortality. Work from Southern California has shown that in areas of high PM concentrations, lung function growth in children is reduced [12, 13] and lung function growth rate is strongly positively correlated with PM concentrations [3].

Objectives Air quality models are important for EPA efforts to improve health, but models contain large uncertainty. This project focuses on reducing uncertainty in sources of aerosols by developing a constraint on emissions using observations with the 4D-Var data assimilation technique. Through this approach, this project will address the following research objectives:

1. More accurately distinguish between natural and anthropogenic sources of aerosol.
2. More accurately distinguish between local and long-range sources of aerosol.
3. Better predict the effects that policy change will have on the future evolution of atmospheric composition.

Ammonium sulfate and ammonium nitrate consist of approximately half of the average mass concentration in the United States. Our analysis will allow us to quantify their contribution to air quality, from uncertainty in their precursor emissions to effectiveness of their control strategies. Inverse modeling of additional species that contribute to $PM_{2.5}$ is not considered here as their formation mechanism is uncertain, or there are not enough widespread observations to warrant data assimilation at this time. While such species could later be included in the analysis, in this work we focus on species for which data assimilation methods can be applied in order to reduce the uncertainty for all subsequent investigations. Recent inverse modeling work considers an alternative approach to inverse modeling $PM_{2.5}$ sources that suggest a possible unified approach for addressing both NO_x and NH_3 emissions.

This project will result in the first ever air quality adjoint that includes an adjoint of aerosol size distribution code, improving the overall performance of the CMAQ model. Particle size distribution is important for both health and climate concerns, as it is associated with the particle origin, their transport in the atmosphere and their ability to be inhaled into the respiratory system [30]. Adjoint models used for decision making that include aerosol size distribution will eventually be able to utilize the Aerosol Polarimetry Sensor on NASA's soon-to-be-launched Glory satellite.

3 Inverse Modeling and Background

Inverse modeling is a means by which one can constrain model parameters (in this case, aerosol precursor emissions and aerosol spatial variability) through the use of observational

data. In order to perform inverse modeling, two main criteria must be met; one must have large amounts of observational data, and knowledge of the physical system.

Adjoint Models Data assimilation provides a framework for combining observations and models to make an estimate of a system (in this case, the chemical makeup of the atmosphere). The adjoint method allows the residuals to be minimized by varying the control variables [14]. To quantify the error of a model prediction, a cost function is introduced. A cost function is a metric that quantifies the total weighted squared error between the model prediction and the observed data. An adjoint model is a means of calculating the sensitivity of the cost function with respect to all model parameters simultaneously. This approach to inverse modeling is appealing because it allows for both the overall magnitude and spatial distribution of emission to be refined [22].

Adjoints of Eulerian chemical transport models have been developed and used for inverse modeling of emissions [10, 17, 18, 20, 25]. In these applications, the adjoint model is used to calculate the gradients of the cost function with respect to emissions.

Nearly all emphasis to date in chemical data assimilation studies has been on gas-phase species. An adjoint of a fixed size aerosol model has been accomplished for a global coupled chemistry-aerosol model [20] and box model adjoint of aerosol dynamics have been studied [21, 35]. The adjoint of GEOS-Chem was developed specifically for inverse modeling of total aerosol mass and composition. Inclusion of aerosol-phase chemistry in the existing CMAQ-ADJ will result in the first coupled gas-aerosol, regional scale adjoint model to explicitly describe aerosol mass composition and size distribution. [8]

Discrete Adjoint There are two types of adjoint models, discrete and continuous. The discrete adjoint method is a technique that calculates the sensitivities of a cost function with respect to control variables by solving a set of discretized partial differential equations.

In order to quantify the misfit between the model prediction and the observed data, a cost function is introduced

$$J = \frac{1}{2}(X - X_0)^T C_m^{-1}(X - X_0) + \frac{1}{2}\sum_{k=1}^{N}(d - G(P_i))^T C_d^{-1}(d - G(P_i)). \quad (1)$$

The parameter X_0 is the *a priori* estimate of X, the matrix C_m is the error covariance corresponding to the *a priori* term, the d variable corresponds to the observations at time t, G maps the solution from the model to the observational space, the C_d's are the error covariances corresponding to the observations, and P_i is the mass distribution of species i. A smaller J value corresponds to a better fit between the model, the data, and the *a priori* estimate.

Following the derivation by Giering and Kaminski [14], the derivative of the cost function with respect to the control variables can be calculated. The derivative of the cost function with respect to changes in the vector of initial conditions is

$$\nabla_X J(X_0) = A^*(X_0)(F(X_0) - D), \quad (2)$$

where $A^*(X_0)$ represents the adjoint model, and the difference $(F(X_0) - D)$ is often referred to as the adjoint forcing.

It has been shown that the computation of the cost function and its gradient, in general, takes only two to five times the computation of the cost function [4], [16]. The gradient vector $\nabla_X J(X_0)$ could also be calculated using the finite difference method. However, the adjoint method has two main advantages over the finite difference method:

1. Especially for a large number of parameters, n, the adjoint model saves run time.
2. The computed gradient is exact when using the adjoint model.

Continuous Adjoint [19] The continuous adjoint is formulated by first solving the nonlinear PDEs and then discretizing as a final step [15]. As with the discrete approach, the continuous adjoint also results in discrete adjoint equations, however the equations developed using the continuous adjoint method are not identical to those developed using the discrete approach. This project will only be using the discrete approach to develop adjoint code; further derivation of adjoints for air quality models can be found in Sandu, et al. (2005) [34].

Advantages of Each Approach Using both the continuous and discrete adjoint methods, one ends up with a set of discrete equations. However, in the discrete approach one starts by discretizing the nonlinear PDE. These discretized equations are then linearized and transposed [15]. In the continuous approach, the discretization is the final step, after linearizing the adjoint equations.

The advantages of the discrete approach are [15]:

- The exact gradient is obtained, ensuring that the optimization process can converge fully. Therefore, discrete adjoint gradients agree more with the finite difference gradients than do the continuous adjoint gradients.
- The creation of the adjoint program is conceptually straightforward. This allows for automatic generation of the adjoint code through programs such as the tangent linear and adjoint model compiler (TAMC).

The advantages of the continuous approach are [15]:

- The adjoint variables and the adjoint boundary conditions have much more physical significance than those in the discrete approach. However, as the mesh width is reduced, the discrete adjoint boundary conditions match the continuous boundary conditions [27].
- The adjoint program can be much simpler and require less memory.

However, it has been shown that the difference between the continuous and discrete gradient reduces as the mesh size increases [27].

4 Box Model

One way of characterizing atmospheric models is by their dimensionality. The simplest model is the box model (a zero-dimensional model), where the atmospheric domain consists of a single cell. In box model simulations, concentrations are only a function of time and therefore are the the same everywhere. As box models are the simplest type of atmospheric model, they provide a good starting point for one who is new to the field.

As an introduction to both chemical transport models and adjoint models, the first three months of this project focused on a stratospheric box model that uses the Kinetic Pre-Processor (KPP). KPP is a tool that converts a specification of the chemical mechanism into FORTRAN or C code which implements the time derivative of concentrations and its Jacobian, together with a suitable numerical integration scheme [6]. The model originally contained only ten chemical reactions, each with a constant reaction rate. During the development of the model an additional 19 chemical reactions were added (Table 1), as well as functions to calculate reaction rates for a given temperature, low and high pressure limits, and surface area (where applicable).

One can see that after the addition of the chemical reactions, not all species are gas phase (i.e., NO_3^-). As there were originally no aerosol dynamics within the model, some preliminary code was developed to calculate coagulation due to Brownian motion. However, the aerosol dynamics code was never integrated into the box model, as I obtained access to NASA's Pleiades super computer and was able to begin work with the three-dimensional chemical transport model CMAQ.

In order to verify that the adjoint for the box model was developed properly, numerous simulations were performed in order to obtain ample finite different and adjoint gradient data to allow for an accurate comparison (Figure 2). An adjoint model is deemed accurate if adjoint gradients and finite difference gradients are nearly identical. One can see from Figure 2 that for this box model the adjoint and finite difference gradients of reaction rates with respect to various species concentrations are exactly the same (to within numerical precision).

5 CMAQ

The air quality model currently used by the EPA is the CMAQ modeling system [2]. It is designed to approach air quality as a whole by modeling multiple air quality issues. CMAQ is a three-dimensional Eulerian chemical transport model of gaseous and aerosol air pollution. It is comprised of three main components: species emissions, transport, and physiochemical transformations. CMAQ is developed to have multi-scale capabilities so that separate models are not required for urban scale and regional scale air quality modeling.

Air Quality Models In air quality studies, we are concerned with a wide variety of topics such as: determining the contribution of a source to the concentration of pollutants in an area, determining the most cost-effective way to reduce pollutant concentrations, where to place a future source in order to minimize its environmental impacts, etc. Air quality models that involve descriptions of emission patterns, meteorology, chemical transformations, and

Chemical Reaction
$NO_2 + O_3 = NO_3 + O_2$
$NO_3 + NO_2 + M = N_2O_5 + M$
$NO_3 + h\nu = NO_2 + O$
$NO_3 + h\nu = NO + O_2$
$N_2O_5 + H_2O = 2HNO_3$
$O^1D + H_2O = 2OH$
$O^1D + CH_4 = OH + CH_3$
$HO_2 + NO = NO_2 + OH$
$O + O_2 + M = O_3 + M$
$HO_2 + O = OH + O_2$
$HO_2 + O_3 = OH + 2O_2$
$OH + O_3 = HO_2 + O_2$
$Cl + O_3 = ClO + O_2$
$ClO + O = Cl + O_2$
$ClO + NO = Cl + NO_2$
$HO_2 + ClO = HOCl + O_2$
$Br + O_3 = BrO + O_2$
$BrO + ClO = BrCl + O_2$
$BrO + ClO = ClOO + Br$

Table 1: Chemical reactions added to stratospheric box model.

removal processes are essential for answering these questions.

CMAQ Aerosol Module CMAQ version 4.7 utilizes a new aerosol module called AERO5. The AERO5 module employs a modal approach to represent particulate [5,26]. The particle size distribution is represented as a superposition of three lognormal sub-distributions, called modes: Aitken mode, Accumulation mode, and Coarse mode. The size of aerosol particles affects their lifetime in the atmosphere as well as their physical and chemical properties [36]. Particles in the Aitken mode have an aerodynamic diameter of 0.1 μm and below ($PM_{0.1}$). The Aitken mode represents fresh particles either from nucleation or from direct emission. The Accumulation mode consists of particles with an aerodynamic diameter between 0.1 μm and 2.5 μm ($PM_{0.1-2.5}$). This mode represents particles that are either directly emitted or have grown from the Aitken mode through coagulation. Particles in the coarse mode have aerodynamic diameters greater than 2.5 μm but less than 10 μm ($PM_{2.5-10}$). In the default configuration, CMAQ estimates concentrations of aerosols in each of the three modes (ex. ASO4I, ASO4J, ASO4K correspond to sulfate aerosol in the Aitken mode, Accumulation mode, and Coarse mode respectively). In addition to estimating concentrations of each aerosol in each mode, CMAQ also calculates A25J (a sum of all particulate matter with an aerodynamic diameter of 2.5 μm and below). All three modes are subject to wet and dry deposition, in addition to condensation. Equilibrium concentrations of inorganic aerosols are

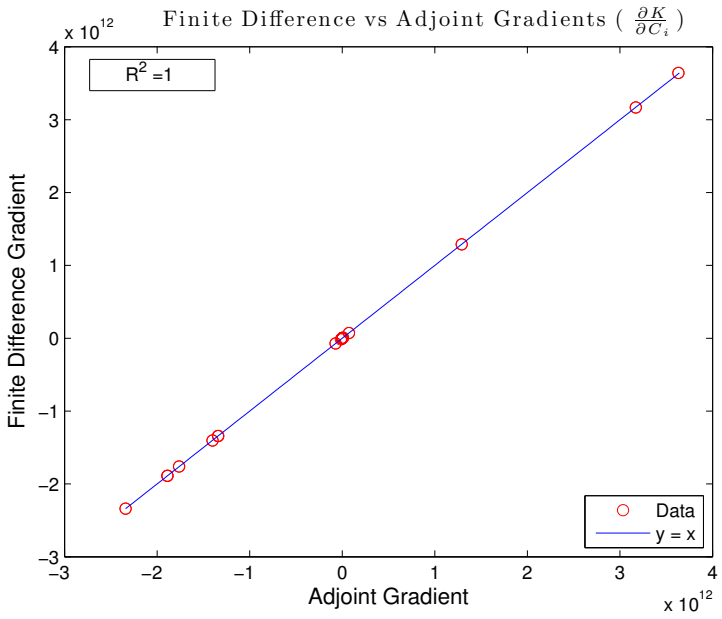
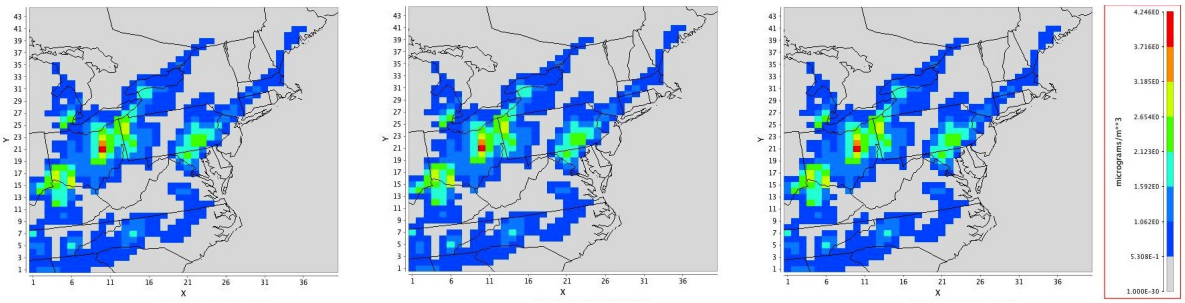


Figure 2: Verification of box model adjoint.

solved using ISORROPIA [28].

The current version of CMAQ, which utilizes the new aerosol module AERO5 by default, contains substantial scientific improvements over previous versions. Previous versions of CMAQ considered coarse-mode particles to be dry and inert. Additionally, components in the coarse mode could not evaporate or condense. CMAQ version 4.7, however, allows semi-volatile aerosol components to condense and evaporate from the coarse mode and non-volatile sulfate to condense on the coarse mode. CMAQ version 4.7 also uses dynamic mass transfer to simulate the coarse mode. This is because particles in the coarse mode are often not in equilibrium with the gas-phase.

Forward Model Simulations Forward model simulations were performed using CMAQ version 4.7 on an Eulerian grid with a horizontal resolution of 36 km $\times$ 36 km with 24 vertical layers. Simulations were initially run in serial, as it is easier to configure the model to run on a single processor. In order to ensure that the model was configured properly, simulations were performed using benchmark input files (initial conditions, boundary conditions, etc.) provided with the model. The input data corresponds to a 24 hour period starting at 00:00 UTC on July 22, 2001. After verifying that the serial simulation matches the output of the benchmark data, the model was setup to run in parallel. Following the parallel simulation that used the benchmark input data, results were compared for the benchmark data (Figure 3(a)), the serial simulation data (Figure 3(b)), and the parallel simulation data (Figure 3(c)). Further comparison of the data shows that the serial and parallel models are set up properly, as the results match to within numerical precision (Figure 4).



(a) Benchmark (b) Serial (c) Parallel

Figure 3: Comparison of accumulation mode ammonium concentrations in layer 1 at 00:00 UTC on July 23, 2001, after a 25 hour simulation, for (a) benchmark simulation, (b) serial simulation results, (c) and parallel simulation results.

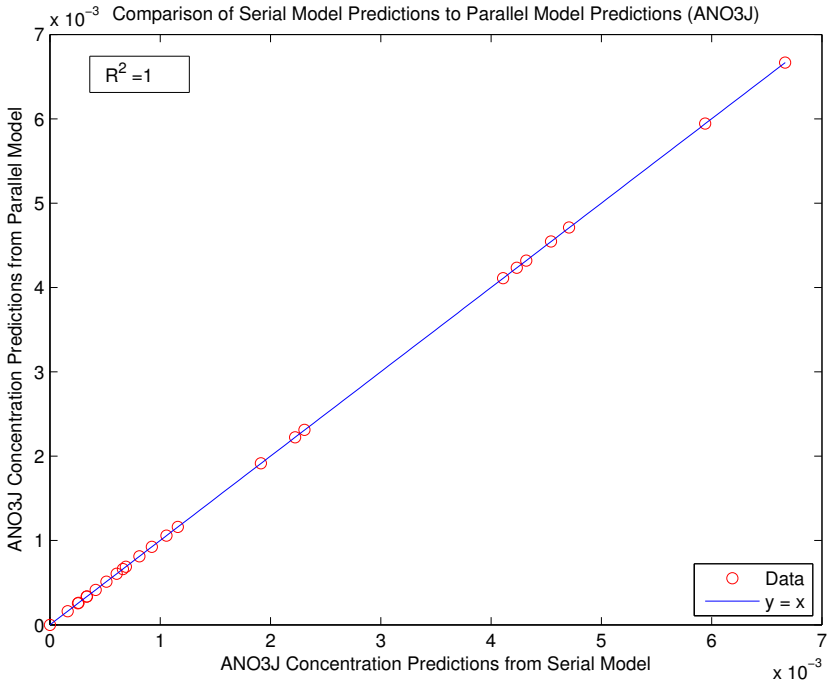


Figure 4: Verification of parallel simulation output for accumulation mode nitrate.

Aerosol Adjoint An adjoint of the entire AERO5 module is to be developed during the process of this project. A new version of CMAQ is scheduled to be released soon, which will include drastic changes to the AERO5 module (changes to the coding of the module, not the science behind it). Because of the impending modifications, it has been advised that the development of the aerosol adjoint be postponed until the newest version of CMAQ is released. There are a few subroutines within the AERO5 module, however, that are likely to remain unchanged (specifically, the N2O5PROB subroutine). The N2O5PROB subroutine

calculates the N_2O_5 heterogeneous reaction probability, often referred to as γ in literature [7]. The adjoint of this subroutine has been developed in order to calculate the sensitivity of the N_2O_5 heterogenous reaction probability with respect to sulfate, nitrate, and ammonium concentrations. Thus, the cost function in this case is simply $J = \gamma$.

CMAQ utilizes four different methods to solve for γ, depending on the value of the user-defined parameter GPARAM. If the user selects a value of 1 for GPARAM, CMAQ uses a constant value of 0.1 for γ, as recommended by Dentener and Crutzen [9]. More recent literature has shown that a constant value of 0.1 is an upper estimate of γ. If the user wishes to employ equations presented by Riemer et al (2003) [33], GPARAM is set to 2 or 3. In this case, γ is calculated according to [33]

$$\begin{aligned} \gamma &= \gamma_1 f + \gamma_2(1-f) \\ f &= \frac{ASO_4}{ASO_4+ANO_3} \end{aligned} \tag{3}$$

where γ_1 and γ_2 are calculated according to Table 1 in Evans and Jacob (2005) [11]; ASO_4 and ANO_3 correspond to aerosol sulfate and aerosol nitrate concentrations, respectively.

The default setting in the current version of CMAQ is to calculate γ as a function of relative humidity, temperature, particle composition, and phase state [7]. In the default setting, the model first checks whether the ambient relative humidity is below the crystallization relative humidity for the given inorganic particle composition. The model also checks whether the ambient relative humidity exceeds the relative humidity for ice formation. If it is determined that the particles are frozen, γ is assigned a constant value of 0.02. If the particles are determined to be dry (RH below crystallization RH), the N_2O_5 heterogeneous reaction probability is calculated according to equations (6-13) of Davis et al. (2008). Experimentation has shown that the uncertainty in γ values increases as RH increases. Therefore, when the particles are determined to be neither frozen or dry (in aqueous phase), γ is calculated using an alternative parameterization as described in equations in Appendix A of Davis et al. (2008)

In order to verify that the adjoint has been written properly, one needs to compare the calculated adjoint gradient to the finite difference gradient. When calculating the finite difference gradient, I used a two-sided finite difference approach. Gradient data was generated using aerosol concentrations in the first vertical layer of the grid cell that contains Los Angeles, CA. The concentrations were determined by running the forward model with a 36 km horizontal grid and twenty-four total vertical layers. The simulation was run between 0:00 UTC on June 3, 2004 and 0:00 UTC on June 7, 2004 with a time step of one hour. The resulting plot of finite difference gradients versus adjoint gradients is shown in Figure 5.

One can see that the adjoint gradients match the finite difference gradients nearly exactly, meaning that the adjoint was developed correctly. It should be noted that there are some points where the adjoint gradient does not match the finite difference gradient. The outlier that corresponds to a finite difference gradient of 0 and an adjoint gradient of -0.019 is the calculated gradient with respect to nitrate corresponding to the initial conditions. The finite difference method fails in this case as the finite difference perturbation was chosen to

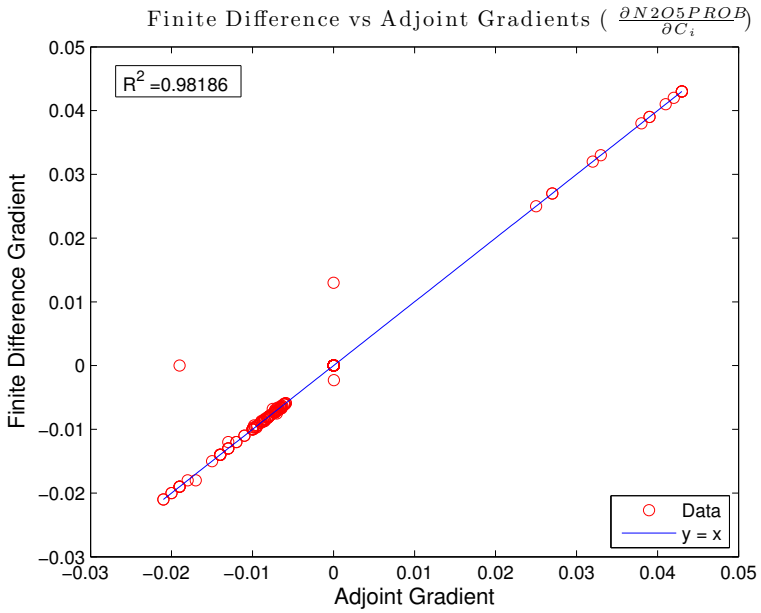


Figure 5: Verification of adjoint of N2O5PROB subroutine under default conditions.

be 0.1% of the concentration, which results in a perturbation that is outside the numerical precision of the computer. The other two outliers both have adjoint gradients of zero and non-zero finite difference gradients. The outlier above the $y = x$ line is a gradient calculated with respect to ammonium, while the outlier that lies below the $y = x$ line is a gradient calculated with respect to sulfate. For both of these points, modifying the code to use a finite difference perturbation of 0.01% results in finite difference gradients much closer to zero. Therefore, it can be concluded that these outliers are caused by nonlinearities in the forward model.

Analysis of the gradients (Figure 6) reveals that a majority of the data results in an ammonium gradient of zero, to within numerical precision. In order to understand why the resulting gradients are zero, one needs to examine the equations from which the adjoint code was developed. For the default calculation of γ, the only location where the ammonium concentration appears is equation 11 from Davis et al. (2008)

$$x_2 = \max(0, \min(1 - x_3, \frac{A}{N + S} - 1)), \tag{4}$$

where x_3 represents the molar concentration of nitrite normalized by the summed concentration of nitrate and sulfate; and A, N, and S are the molar concentrations of ammonium, nitrate, and sulfate, respectively. When $\frac{A}{N+S}$ is less than one, which is the case for a majority of the data used to verify the adjoint, the value of x_2 will be independent of the ammonium concentration resulting in a zero gradient with respect to ammonium.

Figure 6 also shows near zero values for the gradient with respect to sulfate. However, analysis of the data reveals that a majority of these gradient values are on the order of 10^{-6}.

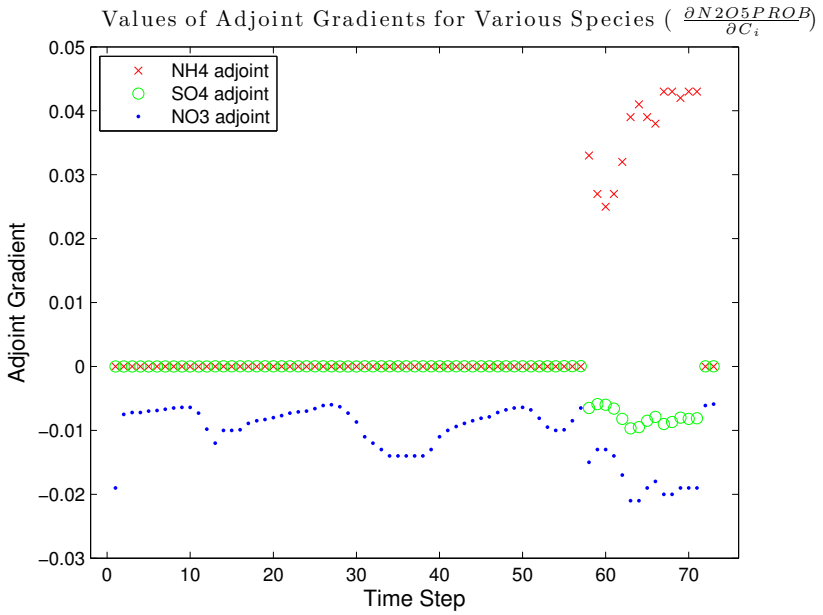


Figure 6: Calculated adjoint gradient values.

It can be concluded that, while perturbations in sulfate concentrations do have an effect on γ, the effect is much smaller than both perturbations in ammonium and nitrate. In addition, one can conclude from the figure that an increase in sulfate or nitrate concentrations will result in a lower N_2O_5 reaction probability, while an increase in ammonium concentration either has no effect or results in an increase in the reaction probability.

As mentioned previously, there are multiple different ways in which the model will calculate the N_2O_5 reaction probability depending on the value of the user-defined parameter GPARAM. In an effort to ensure that the adjoint model will work for all cases, adjoint code was also developed for the cases when the user defines GPARAM to be either 2 or 3. As with all adjoints, the model must be verified by comparing the adjoint gradients to the finite difference gradients (Figure 7). Similar to the default case, the adjoint gradients match the finite difference gradients well, with few outliers. The point that corresponds to a finite difference gradient of zero, with an adjoint gradient of approximately -0.03 is the calculated gradient with respect to nitrate corresponding to the initial conditions. The initial nitrate concentration for this simulation was 2.0×10^{-30}. The finite difference method again fails due to the perturbation being outside the numerical precision of the computer.

Analysis of the gradients (Figure 8) again shows that perturbations in nitrate concentrations are negatively correlated to γ. As in the default case, the gradients with respect to sulfate appear to be nearly zero. However, analysis of the data reveals that the gradients are positive and on the order of 10^{-4}. Similar to the default case, perturbations to nitrate concentrations result in a decrease in γ, and the correlation between γ and changes in nitrate concentration is much stronger than changes in sulfate.

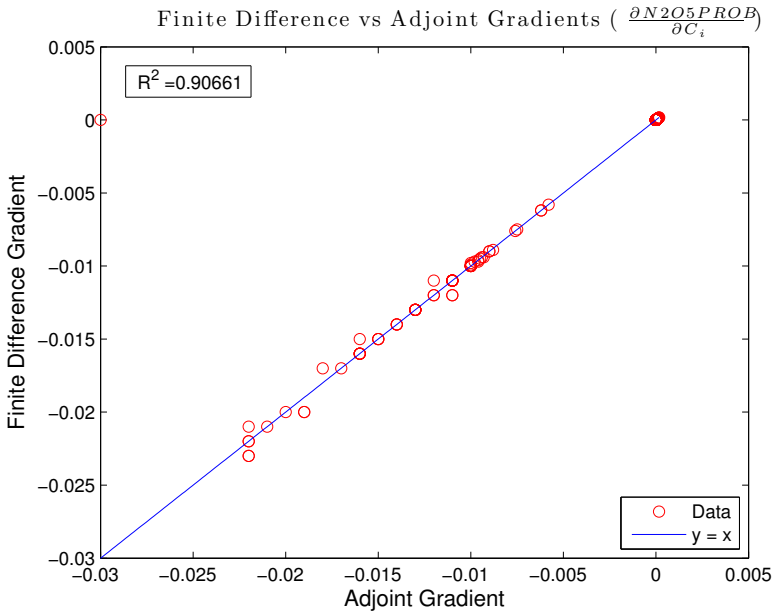


Figure 7: Verification of adjoint of N2O5PROB subroutine when GPARAM equals 2 or 3.

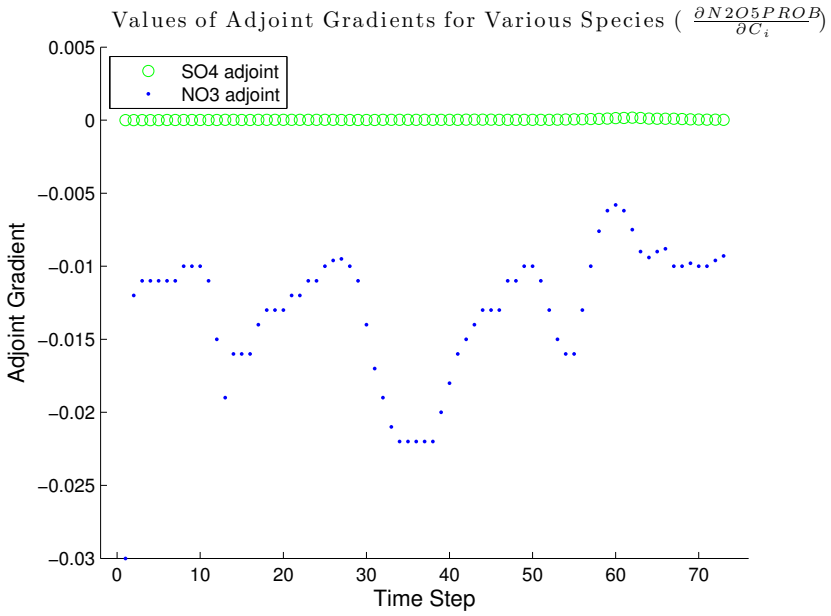


Figure 8: Calculated adjoint gradients when GPARAM equals 2 or 3.

6 Future Work

Adjoint Development Due to the fact that there are to be substantial changes to the coding of the aerosol processes in the next release of CMAQ, it has been suggested that we

hold-off on developing the adjoint for a majority of the subroutines in the model until the newest version of the model is released (anticipated to be released within the next month). Therefore, the next step in this project is still to complete the development of the adjoint of all aerosol processes in the model. This will include an adjoint of aerosol size distribution, which will result in the first ever aerosol size distribution adjoint, improving the overall performance of the CMAQ model. The adjoint of each subroutine will be verified by comparing finite difference gradients to adjoint gradients. Also, the aerosol adjoint will be verified as a whole by again comparing gradients. Furthermore, these additions will be developed in parallel in order to reduce the computational time required to run the model. Parallel code utilizes multiple processors to carry out many calculations simultaneously. Parallelization is very desirable for chemical transport models, because the models are divided into cells. The aerosol processes in each individual cell has no dependence on the other cells, allowing for parallel computation.

Sensitivity Analysis Following the completion of the adjoint model, sensitivity analysis will be performed. It is anticipated that emissions contribute substantially to error in comparing model performance to observations. However, there are model concerns besides emissions that may dominate. Adjoint models provide a means of diagnosing dominant sources of model sensitivity. This is because the sensitivity of the model response is calculated with respect to all model parameters, not just emissions.

Inverse Modeling Inverse modeling will be performed on the domain to constrain spatial variability and seasonal cycles of NH_3 emission. Since the initial conditions provide a small influence when compared to emissions over the course of several months, inversions for each season and year can be performed separately. Inverse modeling for constraining NO_x emissions will use data from months when NO_x emissions from mobile sources are most distinguishable from other sources. Inverse modeling of NO_x will be performed on the U.S. domain, where emissions and met field data will be provided by the EPA's Atmospheric Modeling Division.

Observations Inverse modeling requires a large amount of data. A combination of speciated aerosol concentrations taken from a network of air quality surface monitoring stations, and satellite measurements taken from the TES and OMI instruments aboard the Aura satellite will be used. Ammonium measurements will be taken from the CASTNet of monitoring stations, which provides weekly average measurements of sulfate, ammonium, and nitrate. Tropospheric NH_3 vertical columns from the TES instrument aboard the Aura satellite will also be used. TES has provided global surveys of ammonia from 2004 to 2009, and continues to conduct Special Observations over NH_3 hotspots. Tropospheric NO_2 vertical profiles from the OMI instrument will be used.

Quantification of Success The validity of the inverse modeling solutions will be investigated. Through the use of cross-validation techniques, the dependence of the solution on potential biases in the observations will be explored, where observations from one set of observations are withheld from being used in the inversion. The observations that were withheld are then used to evaluate the performance of the forward model using the inverse

modeling solution.

References

[1] DE Abbey, N Nishino, WF McDonnell, RJ Burchette, SF Knutsen, WL Beeson, and JX Yang. Long-term Inhalable Particles and Other Air Pollutants Related to Mortality in Nonsmokers. *Am. J. Resp. Crit. Care*, 159(2):373–382, Jan 1999.

[2] KW Appel, AB Gilliland, G Sarwar, and RC Gilliam. Evaluation of the Community Multiscale Air Quality (CMAQ) Model Version 4.5: Sensitivities Impacting Model Performance Part I - Ozone. *Atmos. Environ.*, 41(40):9603–9615, Jan 2007.

[3] EL Avol, WJ Gauderman, SM Tan, SJ London, and JM Peters. Respiratory Effects of Relocating to Areas of Differing Air Pollution Levels. *Am. J. Resp. Crit. Care*, 164(11):2067–2072, Jan 2001.

[4] W Baur and V Strassen. The Complexity of Partial Derivatives. *Theor. Comput. Sci.*, 22(3):317–330, Jan 1983.

[5] FS Binkowski and SJ Roselle. Models-3 Community Multiscale Air Quality (CMAQ) Model Aerosol Component - 1. Model Description. *J. Geophys. Res.-Atmos.*, 108(D6):4183, Jan 2003.

[6] V Damian, A Sandu, M Damian, F Potra, and GR Carmichael. The Kinetic Preprocessor KPP - a Software Environment for Solving Chemical Kinetics. *Comput. Chem. Eng.*, 26(11):1567–1579, Jan 2002.

[7] JM Davis, PV Bhave, and KM Foley. Parameterization of N2O5 Reaction Probabilities on the Surface of Particles Containing Ammonium, Sulfate, and Nitrate. *Atmos. Chem. Phys.*, 8(17):5295–5311, Jan 2008.

[8] KL Denman, G Brasseur, A Chidthaisong, P Ciais, PM Cox, RE Dickinson, D Hauglustaine, C Heinze, E Holland, D Jacob, U Lohmann, S Ramachandran, PL da Silva Dias, SC Wofsy, and X Zhang. 2007: Couplings Between Changes in the Climate System and Biogeochemistry. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge University Press.*, pages 1–90, Jun 2007.

[9] F.J Dentener and P.J Crutzen. Reaction of N2O5 on Tropospheric Aerosols - Impact on the Global Distributions of NOx, O3, and OH. *J. Geophys. Res. Atmos.*, 98(D4):7149–7163, Jan 1993.

[10] H Elbern, H Schmidt, O Talagrand, and A Ebel. 4D-Variational Data Assimilation with an Adjoint Air Quality Model for Emission Analysis, Jan 2000.

[11] MJ Evans and DJ Jacob. Impact of New Laboratory Studies of N2O5 Hydrolysis on Global Model Budgets of Tropospheric Nitrogen Oxides, Ozone, and OH. *Geophys Res Lett*, 32(9):L09813, Jan 2005.

[12] WJ Gauderman, Frank Gilliland, Hita Vora, Edward Avol, and Daniel Stram. Association Between Air Pollution and Lung Function Growth in Southern California Children: Results from a Second Cohort. *Am. J. Resp. Crit. Care*, 166:76–84, Jun 2002.

[13] WJ Gauderman, Rob McConnel, Frank Gilliland, Stephanie London, Duncan Thomas, and Edward Avol. Association Between Air Pollution and Lung Function Growth in Southern California Children. *Am. J. Resp. Crit. Care*, 162:1383–1390, Sep 2000.

[14] R Giering and T Kaminski. Recipes for Adjoint Code Construction. *ACM T. Math. Software*, 24(4):437–474, Jan 1998.

[15] MB Giles and NA Pierce. An Introduction to the Adjoint Approach to Design. *Flow Turbul. Combust.*, 65(3):393–415, 2000.

[16] A Griewank. Achieving Logarithmic Growth of Temporal and Spatial Complexity in Reverse Automatic Differentiation. *Optim. Method Softw. 1*, pages 35–54, Nov 2007.

[17] A Hakami, DK Henze, JH Seinfeld, T Chai, Y Tang, G.R Carmichael, and A Sandu. Adjoint Inverse Modeling of Black Carbon During the Asian Pacific Regional Aerosol Characterization Experiment. *J. Geophys. Res. Atmos.*, 110(D14):D14301, Jan 2005. Deriving adjoint equations.

[18] A Hakami, DK Henze, JH Seinfeld, K Singh, A Sandu, S Kim, D Byun, and Q Li. The Adjoint of CMAQ. *Environ. Sci. Technol.*, 41(22):7807–7817, 2007. Notes.

[19] J Haywood and O Boucher. Estimates of the Direct and Indirect Radiative Forcing Due to Tropospheric Aerosols: A Review, Jan 2000.

[20] DK Henze, A Hakami, and JH Seinfeld. Development of the Adjoint of GEOS-Chem. *Atmos. Chem. Phys.*, 7(9):2413–2433, 2007.

[21] DK Henze, JH Seinfeld, W Liao, A Sandu, and GR Carmichael. Inverse Modeling of Aerosol Dynamics: Condensational Growth. *J. Geophys. Res. Atmos.*, 109(D14):D14201, Jan 2004.

[22] DK Henze, JH Seinfeld, and D Shindell. Inverse Modeling and Mapping US Air Quality Influences of Inorganic PM 2.5 Precursor Emissions Using the Adjoint of GEOS-Chem. *Atmos. Chem. Phys. Discuss*, 8(4):15,031–15,099, 2008.

[23] G Hoek and et al. The Association Between Mortality and Indicators of Traffic-Related Air Pollution in a Dutch Cohort Study. *Lancet*, (360):1203–1209, Sep 2002.

[24] D Krewski and et al. Re-analysis of the Harvard Six- Cities Study and the American Cancer Society Study of Air Pollution and Mortality. *Cambridge, MA, Health Effects Institute*, pages 1–97, Oct 2000.

[25] PT Martien and RA Harley. Adjoint Sensitivity Analysis for a Three-Dimensional Photochemical Model: Implementation and Method Comparison. *Environ. Sci. Technol.*, 40(8):2663–2670, Jan 2006. Deriving adjoint equations.

[26] MR Mebust, BK Eder, FS Binkowski, and SJ Roselle. Models-3 Community Multiscale Air Quality (CMAQ) Model Aerosol Component - 2. Model Evaluation. *J. Geophys. Res. Atmos.*, 108(D6):4184, Jan 2003.

[27] SK Nadarajah and A Jameson. A Comparison of the Continuous and Discrete Adjoint Approach to Automatic Aerodynamic Optimization. *AIAA*, (0667), 2000.

[28] A Nenes, SN Pandis, and C Pilinis. Continued Development and Testing of a New Thermodynamic Aerosol Module for Urban and Regional Air Quality Models. *Atmos. Environ.*, 33(10):1553–1560, Jan 1999.

[29] Intergovernmental Panel on Climate Change. Working Group I. Climate Change 2001: the Scientific Basis. page 98, Jan 2001.

[30] World Health Organization. Health Risk of Particulate Matter from Long-range Transboundary Air Pollution. *World Health Organization, European Centre for Environment and Health, Bonn Office*, pages 1–113, Apr 2006.

[31] CA Pope, RT Burnett, MJ Thun, EE Calle, D Krewski, K Ito, and GD Thurston. Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution. *Jama-J. Am. Med. Assoc.*, 287(9):1132–1141, Jan 2002.

[32] CA Pope, RT Burnett, GD Thurston, MJ Thun, EE Calle, D Krewski, and JJ Godleski. Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution - Epidemiological Evidence of General Pathophysiological Pathways of Disease. *Circulation*, 109(1):71–77, Jan 2004.

[33] N Riemer, H Vogel, B Vogel, B Schell, I Ackermann, C Kessler, and H Hass. Impact of the Heterogeneous Hydrolysis of N2O5 on Chemistry and Nitrate Aerosol Formation in the Lower Troposphere Under Photosmog Conditions. *J. Geophys. Res. Atmos.*, 108(D4):4144, Jan 2003.

[34] A Sandu, DN Daescu, G.R Carmichael, and T Chai. Adjoint Sensitivity Analysis of Regional Air Quality Models. *J. Comput. Phys.*, 204(1):222–252, 2005.

[35] A Sandu, W Liao, GR Carmichael, DK Henze, and JH Seinfeld. Inverse Modeling of Aerosol Dynamics Using Adjoints: Theoretical and Numerical Considerations. *Aerosol Sci. Technol.*, 39(8):677–694, Jan 2005.

[36] John H Seinfeld and Spyros N. Pandis. Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, (2nd Edition). *John Wiley & Sons, Inc.*, Jan 2006.

[37] Susan Solomon and Intergovernmental Panel on Climate Change Working Group I. Climate Change 2007: The Physical Science Basis. Jan 2007.

[38] S Twomey. Pollution and Planetary Albedo. *Atmos. Environ.*, 8(12):1251–1256, Jan 1974.