PROJECT SUMMARY:

The overall goal of the proposed project is to develop general computational tools, and associated software, for assimilation of atmospheric chemical and optical measurements into chemical transport models (CTMs). These tools are to be developed so that users need not be experts in adjoint modeling and optimization theory (just as users of CTMs need not be experts in numerical solutions of the partial differential equations that underlie such models). These developments will foster a deeper understanding of: (1) inaccuracies in CTMs; (2) sensitivities of CTMs input and parameter uncertainties; and (3) the comparison of model predictions and atmospheric measurements. These computational tools have the promise to move the field of atmospheric chemical modeling to the next plateau of understanding the extent to which model predictions encompass available measurements, an understanding that is currently hampered by the absence of systematic theory and general analysis tools. We plan to apply these techniques and analysis tools both to the interpretation of observational data and to forecasting activities. Our research approach is to:

- Develop novel and efficient algorithms for 4D-Var data assimilation in CTMs;
- Develop general software support tools to facilitate the construction of discrete adjoints to be used in any CTM;
- Apply these techniques to important applications including: (a) analysis of emission control strategies for Los Angeles; (b) the integration of measurements and models to produce a consistent/optimal analysis data set for the AceAsia intensive field experiment; (c) the inverse analysis to produce a better estimate of emissions; and (d) the design of observation strategies to improve chemical forecasting capabilities.

Significant advances in our fundamental understanding of atmospheric chemistry and our ability to anticipate and manage change requires as accurate a representation of the chemical state of the atmosphere as possible. This proposal has as its objective the development and utilization of Information Technology Research (ITR) tools to integrate measurement and modeling analysis with the goal of providing an *optimal* analysis state of the atmosphere. By optimal analysis state we mean an intimate and close integration of modeled and measured quantities, with the two merged together to provide a consistent and *best* estimate of the chemical state of the atmosphere. This improved estimate state better defines the spatial and temporal fields of key chemical components in relation to their sources and sinks. This information is critical in designing cost-effective emission control strategies for improved air quality, for the interpretation of observational data such as those obtained during intensive field campaigns, and to the execution of air-quality forecasting. The development of the tools to integrate measurements and models is also critical to the challenge of a full utilization of the vast amounts of satellite chemical data in the troposphere that are now becoming available, and which will become more prevalent in the coming years. A closer integration of measurements and models will require significant advances in: (i) data assimilation techniques; (ii) numerical algorithms; (iii) software and application-specific data mining strategies; and (iv) interfaces between measurement data and model data. These topics, along with science applications, represent the major research elements of this proposal. The critical requirement for advancing the state of the science is a focused, multidisciplinary effort, as envisioned herein.

Development of a General Computational Framework for the Optimal Integration of Atmospheric Chemical Transport Models and Measurements using Adjoints

1. GOALS

The overall goal of the proposed project is to develop general computational tools, and associated software, for assimilation of atmospheric chemical and optical measurements into chemical transport models (CTMs). These tools are to be developed so that users need not be experts in adjoint modeling and optimization theory (just as users of CTMs need not be experts in numerical solutions of the partial differential equations that underlie such models). These developments will foster a deeper understanding of: (1) inaccuracies in CTMs; (2) sensitivities of CTMs input and parameter uncertainties; and (3) the comparison of model predictions and atmospheric measurements. These computational tools have the promise to move the field of atmospheric chemical modeling to the next plateau of understanding the extent to which model predictions encompass available measurements, an understanding that is currently hampered by the absence of systematic theory and general analysis tools.

2. INTRODUCTION

Significant advances in our fundamental understanding of atmospheric chemistry and our ability to anticipate and manage change requires as accurate a representation of the chemical state of the atmosphere as possible. This proposal has as its objective the development and utilization of Information Technology Research (ITR) tools to integrate measurement and modeling analysis with the goal of providing an *optimal* analysis state of the atmosphere. By optimal analysis state we mean an intimate and close integration of modeled and measured quantities, with the two merged together to provide a consistent and *best* estimate of the chemical state of the atmosphere. This improved estimate state better defines the spatial and temporal fields of key chemical components in relation to their sources and sinks. This information is critical in designing cost-effective emission control strategies for improved air quality, for the interpretation of observational data such as those obtained during intensive field campaigns, and to the execution of air-quality forecasting. The development of the tools to integrate measurements and models is also critical to the challenge of a full utilization of the vast amounts of satellite chemical data in the troposphere that are now becoming available, and which will become more prevalent in the coming years. The tools to facilitate this integration and optimization need to be developed now.

Assimilation of chemical information is only now beginning in air quality/chemistry arenas, but offers the same motivations as those realized in the field of meteorology. Assimilation techniques can be utilized to produce three-dimensional, time varying *optimal* representations of the chemical composition of the atmosphere, that are consistent with the observed physical and chemical states. These optimal analysis states would be of great value to atmospheric chemistry research. In the case of intensive field experiments, for example, the assimilation of data in the forecast operations would improve the predictive capabilities of the CTMs, while in post-analysis the merging and integration of the flight observations and modeled fields would provide a comprehensive, analyzed, self-consistent, 3-D chemical and dynamic data set that all scientists could use to help interpret their measurements. The assimilation techniques can also be used to systematically improve our ability to refine individual science components. For example sensitivity and error analysis embedded in the simulation can be used to provide better estimates of chemical emissions into the atmosphere (inverse modeling), and to provide methodologies to analyze the simulation and to systematically design optimal measurement strategies (i.e., what measurements are needed and where should they be conducted to improve our depiction of the chemical state of the atmosphere.

A schematic diagram of our research focus is presented in Figure 1. A determination of an optimal analysis state requires the use of both model and measurement data. The traditional inter-relationships between predicted and measured quantities, where the measurements are used to provide initial and boundary conditions for the models and to evaluate model performance, and the modeled quantities provide context and interpretation of the measured quantities, are included in this framework. But this proposal goes well beyond these boundaries. The sensitivity and error analysis techniques required to produce the optimal analysis state provide new methodologies for analyzing model simulations and providing the context for comparisons with measurements. This envisioned intimate integration of measurements and models represents formidable ITR challenges. For example, the assimilation of chemical quantities into atmospheric models greatly increases the computational burden (by at least an order of magnitude). As described in the following sections, a closer integration of measurements and models will require significant advances in: (i) data assimilation techniques; (ii) numerical algorithms; (iii) software and application-specific data mining strategies; and (iv) interfaces between measurement data and model data. These four topics, along with science applications, represent the major research elements of this proposal.

The critical requirement for advancing the state of the science is a focused, multidisciplinary effort, as envisioned herein.

This proposed research sets out to develop and apply assimilation techniques for the purposes outlined above. The techniques to be employed arise out of fundamental aspects of optimization and control theory, but their application will require new developments in approaches to implementing these methods, since atmospheric CTMs represent, without doubt, the most ambitious application of this theory yet attempted.

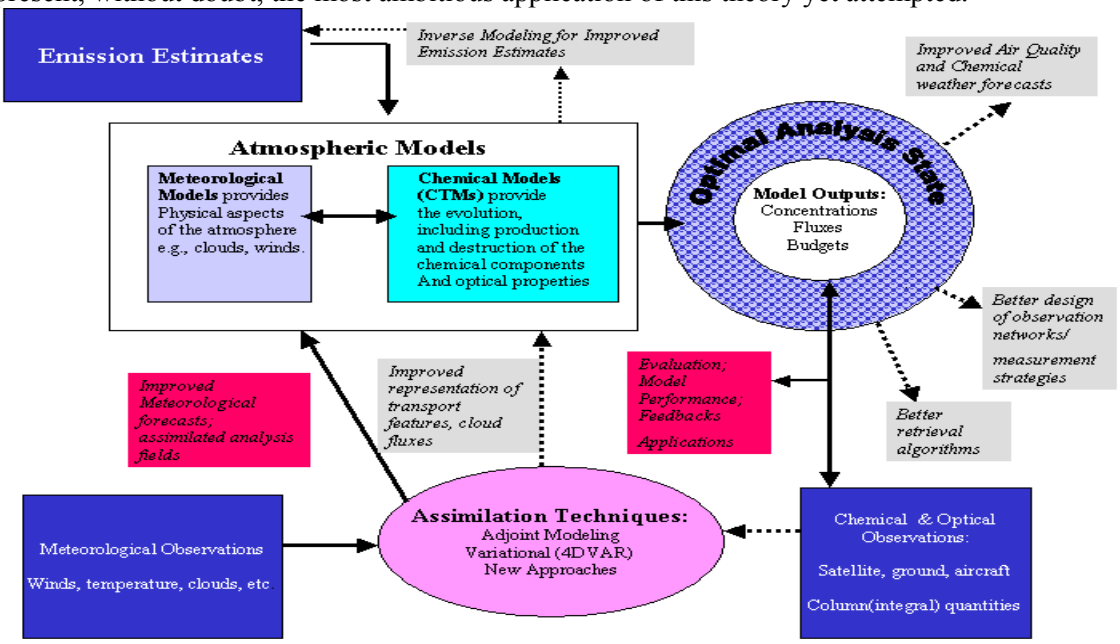

Figure 2. Overview of our proposed research in data assimilation for chemical models. Solid lines represent current capabilities. Dotted lines represent new analysis capabilities that arise through the assimilation of chemical data.

3. BACKGROUND

Significant advancements have been made in recent years in our ability to measure and model atmospheric chemistry. We are now able to measure at surface sites and on mobile platforms (such as vans, ships and aircraft), with fast response times and wide dynamic range, many of the important primary and secondary atmospheric trace gases and aerosols (e.g., carbon monoxide, ozone, sulfur dioxide, black carbon, etc.), many of the critical photochemical oxidizing agents (such as the OH and HO_2 radicals), the chemical composition of individual particles, and integral optical properties including scattering at low relative humidity (mid-visible wavelength), fine-mode fraction of scattering, wavelength dependence of both sub-micron and coarse scattering, single scattering albedo of both sub-micron and coarse-mode aerosol, and hygroscopicity of both sub-micron and coarse aerosol. Not only is our ability to characterize a fixed atmospheric point in space and time expanding, but the spatial coverage is also expanding through growing capabilities to measure atmospheric constituents remotely using sensors mounted at the surface and in aircraft. Remote sensors mounted on satellites are now providing global coverage in the troposphere for some chemical species (e.g., CO, HCHO). Similar advancements are being made in our ability to measure key physical properties associated with these constituents, including optical properties and radiance fields.

Chemical transport models (CTMs) are designed to describe the fate and transport of atmospheric chemical constituents associated with the gas and aerosol phases. CTMs have advanced to the point where they now specifically follow on the order of one hundred chemical species, interacting through chemical mechanisms involving hundreds of chemical reactions. In addition, the transport aspects of CTMs are now run in close interaction with dynamic meteorological models (e.g., Grell, et al., 1993), with the tendency to have the chemical and dynamical aspects of the models merged into a single computational system. Because of the growing recognition of the role that aerosols play in radiation, meteorology/climate, and health impact areas, CTMs are now including a more detailed description of aerosol dynamics, and calculate size-resolved aerosol composition, radiances, and photolysis rates interactively with the cloud and aerosol fields. Computational power and efficiencies have advanced to the point where CTMs can simulate pollution distributions in an urban air shed with spatial resolution of a few kilometers, or can cover the entire globe with horizontal grid spacing of less than 100 kilometers (less than 1 degree). Contemporary CTMs (see section 4.3.1 for details regarding the models to be

used in this study) when applied in prognostic studies are able to provide quantitative information on the distributions of many of the key trace gases and aerosol constituents in the atmosphere. CTMs have become an essential element in atmospheric chemistry studies, including important applications such as providing science-based input into best alternatives for reducing pollution levels in urban environments, and assessments into how we have altered the chemistry of the global environment.

However, quantitative aspects of model-based atmospheric chemistry analyses and forecasts are hampered by the fact that comprehensive CTMs are often poorly constrained due to a variety of reasons including: incomplete emissions information; lack of key measurements to impose initial and boundary conditions; missing science elements; and poorly parameterized processes. Improvements in the analysis capabilities of CTMs require them to be better constrained through the use of observational data. The close integration of observational data is recognized as essential in weather/climate analysis and forecast activities, and this is accomplished by a mature experience/infrastructure in meteorological data assimilation. By data assimilation we are referring to the process by which model predictions utilize measurements to produce an optimal representation of the state of the atmosphere.

To be precise, consider an atmospheric CTM represented as a system of coupled partial differential equations

$$\frac{\partial y}{\partial t} = F(y) + \varepsilon, \qquad y(t_0) = y_{0,} \qquad y\,|_{boundary} = b, \tag{1}$$

where y is the state vector representing the concentration field, F the advection-diffusion-reaction operator, and ε a random field representing the modeling error (which includes errors in meteorological fields, missing chemistry, etc.). The observational data are given as a sequence of state measurements y_k^o taken at discrete moments in time t_k, k=1,2,… ; the measurements are assumed to be linearly related with the state vector

$$y_k^o = H_k y_k + \varepsilon_k^o \tag{2}$$

and corrupted by observational errors ε^o. Here *H* is the model equivalent operator. Data assimilation tools provide an optimal estimate of the evolving state of the atmosphere by combining the estimated state of the atmosphere - as simulated by the model (1) - with the measurement of the state - as given by observations (2).

Assimilation techniques fall within the general categories of variational (3D-Var, 4D-Var) and Kalman filter based methods, which have been developed in the framework of optimal estimation theory. The Kalman Filter methods are stochastic methods developed under the minimum variance principle (minimize the trace of the error covariance matrix) that attempt to forecast in a dynamically consistent manner not only the state but also the error covariance matrix. They have the advantage that information on the model error is included into analysis. In the Kalman Filter approach the observations are processed sequentially. The variational data assimilation approach seeks to minimize a cost functional that measures the distance from measurements and the "background" estimate of the true state. In the 3D-VAR method the observations are processed sequentially in time. The 4D-VAR generalizes this method by considering observations that are distributed in time. These methods have been successfully applied in meteorology and oceanography [Navon, 1998], enabling both meteorological forecasts and the production of meteorological analysis several times a day. These analyses have proved invaluable in understanding the atmospheric circulation and monitoring and assessing changes in the circulation on both long and short timescales.

Assimilation techniques are beginning to be applied to atmospheric chemistry models [Menard, et al., 2000; Elbern, et al., 1999; Mulholland & Seinfeld, 1995]. Most efforts have focused on assimilation of satellite data using variants of the Kalman filter technique. These efforts include the assimilation of aerosol optical depths for the Indian Ocean Experiment (INDOEX) [Collins et al, 2001] using the Model of Atmospheric Chemistry and Transport (MATCH); the assimilation of column CO from Measurement of Air Pollution from Space (MAPS) [Lamarque et al., 1999 and Kasibhatla et al., 2000], CO from the Japanese Advanced Earth Observing (ADEOS) satellite [Clerbaux et al., 2001], and from Measurements of Pollution in the Troposphere (MOPITT) [Lamarque et al., 2001]; the near real-time assimilation of ozone using the Total Ozone Mapping Spectrometer (TOMS) total ozone column and the Solar Backscatter Ultraviolet (SBUV) partial ozone column using the GEOS (Goddard Earth Observing System) Data Assimilation System (DAO) [Stajner et al., 2001 and Riishojgaard et al., 2000], the assimilation of MLS ozone [Khattotov et al., 2000 and Levelt et al, 1998] and the assimilation of the tropospheric ozone column [Lamarque et al., 2001]. These studies have consistently shown a marked improvement in the concentrations of assimilated species when compared against independent datasets. Since the publication of Lamarque et al.(1999) the NCAR assimilation system has been improved to include an estimate of the error

associated with the assimilated CO distribution. A two dimensional Kalman Filter for a stratospheric tracer (pure advection) model has been implemented at Data Assimilation Office, NASA Goddard Space Flight Center to assimilate methane observations from the UARS [Menard et al., 2000]. For the tracer analysis a Lagrangian formulation of the Kalman filter showed to be a more natural framework than the Eulerian filter due to the conservative properties of field, error variance, and error covariance dynamics [Lyster et al., 2000]. However, even for this simplified model the computational cost is significant.

These techniques are just beginning to be used in non-linear atmospheric chemical applications. When chemical transformations and interactions are considered, the complexity of the implementation and the computational cost of the data assimilation are highly increased. Recently 4D-Var assimilation techniques have been applied to chemical transport models, where the initial conditions chemical concentrations are estimated from observations over the assimilation period [Fischer and Lary, 1995; Elbern et al., 1997; Elbern and Schmidt, 1999 and 2001].

Current assimilation methods have their strengths as well as weaknesses. Only for perfect (zero model error) linear models with Gaussian statistics, both 4D-Var and Kalman Filter techniques provide identical analyses at the end of the 4D-Var-analysis interval [Jazwinski 1970, Li and Navon 2001]. In realistic nonlinear cases, the fundamental differences between these approaches are driven by the assumptions invoked and the mechanisms used to obtain an optimal estimate, in ways that are poorly understood. There are a number of computational and theoretical issues that need to be explored to utilize the full power of data assimilation schemes. The key challenges are related to: 1) Complexity and cost of data assimilation is greatly increased over meteorological assimilation; 2) Issues of error growth within CTMs undergoing assimilation have not been explored and remain open research questions; 3) The model error used in KF and EKF is hard to specify. There are practical problems related with the model error that necessitate a fine tuning parameter analysis and the influence of the model bias on the analysis results; 4) The computational cost of the full Kalman Filter (also the Extended Kalman Filter-EKF) is of order n^3 where n is the dimension of the state vector (spatial grid dimension times the number of species). For comprehensive atmospheric chemistry models n may be easily of order 10^6-10^7, which for the present computing resources makes the implementation impractical. This high computational cost has prevented the Kalman filter based algorithms from being implemented in their full form. Intensive research is dedicated to developing sub optimal Kalman filters and smoothers (Todling 1994); and 5) In order to perform the 4D-Var computations the forward trajectory must be available in reverse order such that a large amount of memory must be allocated for storage during the forward run. Intensive research is focused on developing optimal storage strategies and check-pointing schemes [Griewank 2000]. The storage for atmospheric chemistry applications is one order of magnitude larger than in meteorological implementations [Elbern and Schmidt, 1999].

New approaches are needed and will be explored in this research. In the proposed research we plan to focus on 4D-Var assimilation and adjoint analysis. We believe that these techniques offer a theoretically sound and computationally feasible framework for the integration of models and measurements. In the context of assimilation within CTMs, the optimal 4D-Var data assimilation problem looks for the (initial) state y_0 that minimizes the cost functional

$$\Im(y) = \frac{1}{2}(y - y^b)^T B_0^{-1}(y - y^b) + \frac{1}{2}\sum_{k=0}^{N} (H_k y_k - y_k^o)^T R_k^{-1}(H_k y_k - y_k^o) \tag{3}$$

where y^b is the background concentration, y_k is the model (1) output at t_k when y_0 is the initial value, B and R are the covariance matrices of the background and observational errors.

The inverse modeling problem looks for the vector p of model parameter values that minimizes the model-observation gap as measured by the cost functional

$$\Im(y,p) = \frac{1}{2}(p - p^b)^T B^{-1}(p - p^b) + \frac{1}{2}\sum_{k=0}^{N}(H_k y_k - y_k^o)^T R_k^{-1}(H_k y_k - y_k^o) \tag{4}$$

where p^b and B are the background value and the error covariance matrix of the parameters.

The minimization problem associated with these approaches is computationally demanding, but can be efficiently solved using adjoint modeling to compute the gradient of the cost functional. This is achieved with the backward-in-time loop

Step 1. Initialize $\nabla\Im = 0$

Step 2. for k = N,1,-1 do $\nabla\Im = (\frac{\partial y_k}{\partial y_{k-1}})^T [H_k^T R_k^{-1}(H_k y_k - y_k^o) + \nabla\Im]$ (5)

Step 3. $\nabla\Im = B_0^{-1}(y_0 - y^b) + H_0^T R_0^{-1}(H_0 y_0 - y_0^o) + \nabla\Im$

The advantage of the adjoint algorithm as compared to forward methods is that it avoids matrix-matrix multiplications. Compared with sequential techniques the 4D-Var has the advantage that it can be cast in a manner completely consistent with the model equations and the four-dimensional distribution of the observations over the assimilation window. This makes the cyclic 4D-Var very suitable for numerical forecasting. Also for real-time applications the availability of results may be delayed since the analysis process starts after all the observations are available. While relatively inexpensive from the computational point of view, the 3D-Var takes as input the observational and forecast error statistics. However, in practice little is known about these statistics. Variational methods may also be used to estimate various parameters of the model such as boundary values, and emission rates.

We believe that with the multidisciplinary and concentrated research plan proposed here, we will be able to develop the computational techniques and software tools needed to perform 4D-Var analyses in CTMs, and apply them to important problems in atmospheric chemistry. Details are presented below.

4. RESEARCH ELEMENTS

The research elements of this proposal are focused on the development of a general computational framework to facilitate the close integration of measurements and models in CTMs. We plan to focus our efforts on the development of assimilation methodologies based on 4D-Var data assimilation and adjoint modeling to the general and closely related problems of ***Adjoint Sensitivity Analysis, Optimal Variational Data Assimilation***, and ***Inverse Modeling***. We plan to apply these techniques and analysis tools both to the interpretation of observational data and to forecasting activities. Our research approach is to:

1. Develop novel and efficient algorithms for 4D-Var data assimilation in CTMs;
2. Develop general software support tools to facilitate the construction of discrete adjoints to be used in any CTM;
3. Apply these techniques to important applications including: (a) analysis of emission control strategies for Los Angeles; (b) the integration of measurements and models to produce a consistent/optimal analysis data set for the AceAsia intensive field experiment; (c) the inverse analysis to produce a better estimate of emissions; and (d) the design of observation strategies to improve chemical forecasting capabilities.

This research program is ambitious, but is what is needed to make optimal use of the sparse data and complicated atmospheric chemistry models. This research builds extensively upon our experiences with atmospheric chemistry models, measurements of chemical and optical quantities, symbolic processing for automatic code generation, the design of efficient and stable numerical methods, adjoint modeling and assimilation, in finding optimal measurement distributions, and parallelization of CTMs. Our expertise will be extended and incorporated in a comprehensive set of computational tools for adjoint analysis and assimilation. The main research elements are detailed below. The time-line and management aspects of the research plan are presented in the Management Section.

4.1. Algorithms

Data assimilation can be accomplished successfully only if numerical errors are small, and the gap between model output and measurements is due mainly to insufficient information (e.g. for initial and boundary conditions, physical process parameterizations, etc). A successful use of the 4D-Var framework requires efficient solutions to the following computational issues: construction of adjoints, evaluation of gradients by reverse integration, checkpointing, and optimization. An important part of this research will focus on the development of novel efficient algorithms for 4D-Var data assimilation. In what follows we will refer to "adjoint of a model" as a short designation for the adjoint of the tangent linear operator associated with the nonlinear (continuous or discrete) model.

4.1.1 Improved numerical methods

The chemical 4D-Var data assimilation requires an iterative optimization procedure, where at each iteration the chemical adjoint model is integrated (for derivative information), and the relevant parameters are adjusted such that the cost functional decreases. The assimilation process is therefore considerably more expensive than running the model in forward mode. The great increase in computational burden due to chemical data assimilation places much more importance on the need to increase the efficiency of numerical methods.

The adjoint formulation (5) requires the derivative of one step of the numerical method (dy_k/dy_{k-1}). The numerical algorithms for the forward method be fast and accurate, but in addition they should allow an inexpensive calculation of step derivatives dy_k/dy_{k-1}. The resulting reverse integration of the adjoint should also

be fast and accurate. Therefore new constraints are placed on the numerical methods. In this research we will focus on developing such methods for nonlinear models, including chemical kinetics and aerosol optical properties and dynamics. An example is the adjoint of the Rosenbrock method we developed in [Daescu et.al., 2000]. To improve upon the performance will plan to test Runge-Kutta time stepping methods for computing (dy_k/dy_{k-1}), and we will investigate the propagation of derivatives with Gear methods. An alternative to building the adjoint of the discrete equation is to consider a (backward-in-time) numerical integration of the continuous adjoint equation. This may lead to a simpler formulation, but raises questions related to the convergence of the optimization procedure, the formulation of continuous adjoint for complex boundary conditions, etc. This formulation and the related issues will also be investigated.

In addition, the assimilation procedure makes explicit use of the model representativeness error, which includes the modeling and numerical errors. In this research we plan to ensure that the forward model output will produce the concentrations together with the estimated numerical errors [Sandu, 1997; Verwer et. al., 1999]. The estimators of these errors can be implemented as a difference of solutions obtained at different resolutions (coarse/fine grids or large/small time steps), or as a difference of solutions obtained with numerical methods of different orders of accuracy.

4.1.2. Construction of discrete adjoints for specific CTM elements

While the formulation of adjoints for the linear transport part is a relatively well-studied problem, for other elements of modern CTMs only a handful of results are currently available. We will develop adjoint formulations and solution algorithms for the following areas.

Adjoint models for stiff systems. Ordinary differential equations simulating gas and liquid phase chemistry and interphase mass transfer are stiff, and implicit discretization formulas are needed to ensure stability of the numerical integration. Implicit formulas solve linear systems of equations at each step; building an efficient adjoint model (of the discrete system) is therefore a challenging task. There are few results available on the theory and implementation of adjoint models for stiff systems. In [Daescu et. al., 2000] we developed an adjoint formulation for Rosenbrock integration methods. In the proposed work we will extend this work and formulate adjoint models for the popular Gear and Runge-Kutta methods, which are fully implicit and therefore require an iterative solution process. Sparsity of the chemical Jacobians will be exploited to obtain computational efficiency. The reuse of LU decompositions from the direct integration will be considered in order to further minimize the computational time. Attention will be devoted to reduced chemical models that can decrease the complexity of adjoint formulations.

Adjoint models for aerosol and optical properties simulations. Particulate matter (aerosol) processes have become a priority focus area in environmental science. Aerosol dynamics (coagulation, growth, nucleation) is described by a set of integro-differential equations [Jacobson, 1999]. When aerosol processes are included in the models (solving for their size distribution and their chemical composition) the overall computational time increases by an order of magnitude. Building efficient adjoint models for aerosol dynamics and chemistry is a challenging and (to our knowledge) largely unexplored direction. The proposed work will continue our preliminary efforts [Sandu et. al., 2001a,b,c,d] to build and solve adjoint aerosol models. We will explore adjoint formulations of the integro-differential dynamic equations and of the coupled dynamic-chemistry equations; and adjoint formulations for the chemical equilibrium solvers widely used in aerosol modules.

In addition to the aerosol dynamics problem, there is the important problem of aerosol optical properties. Today we have the ability to measure integral optical properties associated with aerosol scattering and absorption with high resolution and excellent accuracy (and knowledge of uncertainty - which is critical to assimilation). These measured properties reflect intimate couplings between aerosol size and composition, atmospheric radiative transfer, and meteorological parameters such as relative humidity and clouds. There is a big need and a great opportunity to figure out how to incorporate these integral optical properties into the data assimilation and sensitivity analysis framework. We will develop adjoint formulations of the coupled CTM, radiative transfer and optical property estimation components of CTMs, and use these to explore the sensitivity of optical properties to the total aerosol concentration and the mixture of components. A key goal will be to identify how well these optical measurements constrain the chemical models and what would be the highest leverage measurements to add to the suite for improving the constraint. This work will build upon research into sensitivity analysis in coupled CTM and radiative transfer models using automatic differentiation [He and Carmichael, 1999; and He et al., 2000], and will be extended to adjoint and 4D-Var techniques.

Adjoint models for nested and adaptive grids. Non-uniform grids are necessary to control the spatial discretization error by placing additional spatial resolution in the areas of higher physical or chemical activity. A

common used technique is nesting finer grids within the coarser computational grid (an approach taken by WRF). A more involved approach dynamically and automatically adjusts the grid as the areas of intense activity change during the simulation. The use of adaptive grids in atmospheric chemistry-transport models was previously studied at the University of Leeds [Tomlin, 2000], at CWI, Amsterdam [van Loon, 1996; Blom and Verwer, 1994], as well as in industry [SAIC, 2001; Sarma, 2000]. The development of the mesh generation/management software is outside the scope of this project. We will build on existing adaptive mesh generators like SUMAA3D [SUMAA3D, 2001], GRUMMP [Ollivier-Gooch et. al., 2001], Geompack [Joe, 2001], QMG [Vavasis, 2000]. We will study the construction and implementation of adjoint models with nested grids and with unstructured adaptive grids for different types of elements (parallelepiped, tetrahedral, hexagonal). The use of adaptive grids is expected to introduce new challenges in the building of adjoint models via automatic differentiation, similar to the ones raised by adaptive step sizes with time integration. The checkpoints must store grid information, and the backward integration must recreate the sequence of grids used by the direct integration.

4.1.3. Optimization algorithms

The optimization algorithm is another key element of the assimilation process. The optimization problem is nonlinear and is huge; assimilating a concentration field requires the minimization of the cost functional over a space of millions of variables. Each iteration of the optimization algorithm involves (at least) one backward integration of the adjoint model - therefore iterations are expensive and fast convergence is highly desirable. Popular unconstrained optimization methods used in the context of 4D-Var assimilation of transport processes are Newton, truncated Newton [Wang, et al. 1995] and LBFGS. Presently we are using the LBFGS [Daescu et.al., 2000]. A promising newer development is the Adjoint Newton algorithm [Wang, et al. 1997; Wang, et al. 1998]. For assimilating chemical fields it is natural to search for a solution under the constraint that concentration values are nonnegative. Our research will focus on the analysis and development of hybrid methods for large-scale optimization that combines quasi-Newton methods with Hessian-free Newton (HFN) methods in a dynamic manner. While the gradient of the cost functional may be computed using the adjoint method, automatic differentiation or finite differences may be used to compute the products of the Hessian matrix times a vector. An efficient approach for computing Hessian-vector products uses the second order adjoint model (forward over reverse [Le Dimet et. al. 2001]). Our experience shows that the cpu time to compute an exact Hessian times vector product is bounded by a constant (~10) times the cpu time of the forward run. Truncated Newton (TN) methods have been used in meteorological data assimilation experiments [Wang, et al. 1995] but are still believed to be too expensive for atmospheric chemistry models. Interlacing iterations of the limited memory BFGS method [Liu and Nocedal, 1989] and truncated Newton methods allows for the information on the curvature of the cost function collected in the limited memory matrix to be used for preconditioning the inner conjugate gradient iteration of the HFN methods. These "enriched methods", which attempt to improve the performance of both algorithms without increasing the computational cost have already shown promising results for large-scale optimization [Morales and Nocedal, 2000] and will be utilized in this research.

4.1.4. Adjoint models as tools for sensitivity studies, parameter estimation , and targeting

In air quality modeling uncertainties in the model initial state, emission and deposition rates, and boundary values, to name only a few, must be considered. Most of the previous work was dedicated to the problem of finding an optimal initial state and only recently has the feasibility of adjoint modeling for emission rate assessment begun to be explored (Elbern et al. 2000). A mathematical framework of the adjoint parameter estimation, identifiability and regularization issues with applications to meteorology and oceanography are described by Navon (1998). In Vukicevic and Hess (2000) we constructed an adjoint to the transport (advection, deep convection and shallow and boundary layer transport) in HANK (the NCAR regional CTM) as well as to the wet and dry deposition. This adjoint was used to examine the sensitivity of the long-range transport of Asian emissions to Hawaii with regards to emission strength, wet deposition and convective mass flux and boundary conditions. We demonstrated how to use adjoint information to construct sensitivity maps of the crucial transport parameters versus time, and to analyze the transport pathways of tracer plumes to Hawaii accounting for all relevant model processes (i.e., including convection and diffusion). In this proposed research we will further develop and apply these techniques to atmospheric chemistry models in the 4D-Var data assimilation context. This research activity will focus on algorithms for: 1) improving the models performance by providing an optimal set of parameters; 2) identifying highly sensitive spatial regions where changes in parameters (e.g. emission rates) have the greatest impact on the model forecast; 3) identification of emission rates.

The spatio-temporal distribution of the observations plays an essential role in the data assimilation process. The adaptive (targeting) observations methods search for locations were additional observational resources should be deployed in order to minimize the error in the analysis and subsequent forecasts. Strategies for targeting observations were considered mostly in numerical weather prediction [Langland et al. 1999, Palmer et al. 1998] and applications to atmospheric chemistry are just beginning. Sensitivity studies on the observing network in the context of data assimilation [Rabier et al. 1996, Pu et al. 1997] have shown that the impact of data on the analysis estimate is highly determined by the location of the observations relative to dynamically sensitive regions of the atmosphere. The adjoint technique is an efficient approach to identify sensitive regions where uncertainty in the model parameters can lead to large errors in the model forecast. The adjoint of the forecast model may be used to implement targeting methods such as the gradient sensitivity method [Langland et al. 1999] and singular value decomposition technique [Gelaro et al. 1999]. Our research will focus on testing and developing new targeting methods in the context of 4D-variational data assimilation. Recently [Daescu and Carmichael, 2001] we proposed a novel adjoint-based algorithm for adaptive observations using a periodical update of the sensitivity values to include the interaction with previously located observations and observations taken from fixed locations. Preliminary results performed with a 2D transport-chemistry test model indicate that at a computational cost equivalent with the cost of few forward model integrations an adaptive observational path may be determined with significant benefits for the model forecast. In this proposal we plan to apply these techniques to 3-dimensional studies as discussed later.

4.2. Software Tools

Another important goal of this project is to develop general software tools to facilitate the close integration of measurements and models. Specifically, we will develop a software framework to support 4D-Var data assimilation for CTM's. The framework will be generic in the sense that any (forward) chemistry-transport model can be used as a basis for the assimilation or inverse modeling process. The framework will provide tools for: 1) construction of the adjoint model; 2) handling large datasets; 3) checkpointing support; 4) optimization; 5) analysis of results; 6) remote access to data and computational resources.

We will apply/test/demonstrate the new tools in specific studies using our chemical transport models. However, the outcomes of this research will be general tools and methodologies that are widely applicable and not model-specific. Figure 2 gives an overview of the software tools envisioned and their use in optimal state analysis. We now explain in detail the components of the framework.

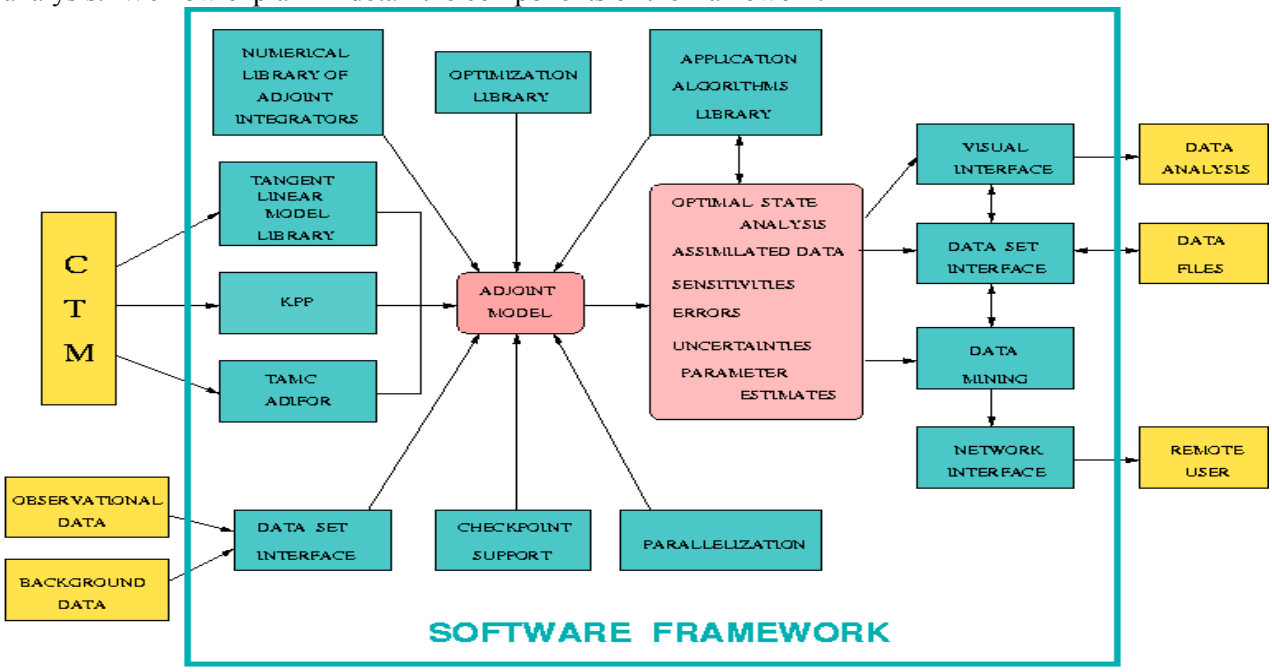

Figure 2. Software framework and tools to be developed.

4.2.1. Improved modeling software technology

Software technology is an important component of large-scale models. A difficult-to-follow design of the forward model negatively impacts the ability to build an efficient adjoint. Therefore, as a starting step, we will explore software design techniques and tools that will make the direct models easier to use and maintain, and will

facilitate the development of adjoints and their applications. A layered software design will be employed for the chemistry-transport model and the resulting adjoint. Low-level layers will manage the distributed data structures over multiple processors; science models, numerical methods and the corresponding adjoint models will be implemented at upper-level layers. This approach will greatly assist error control, management of adaptive meshes, interoperability with the meteorological model, employment of standard numerical algorithms, interface with on-line visualization tools, and portable parallelization. We will also study the possibility of reformulating CTMs within general PDE-solving environments [Overture; DAGH; KeLP] that offer a ready-to-use programming infrastructure and easy access to standard numerical libraries.

4.2.2. Tools for building adjoint models.

The framework will contain software tools that assists the construction of the adjoint of a given CTM in a quick and error free manner. Some tools will be developed during this research (KPP, adjoint model libraries), while other are generic and publicly available (TAMC, Adifor).

Construction of transport adjoint model using automatic differentiation. Automatic differentiation tools [Carmichael et. al., 1997] and symbolic preprocessing [Damian-Iordache, 1996] may be used to implement the adjoint code that allows for flexibility. For the linear transport equations good adjoint models can be obtained using automatic differentiation software like TAMC. Black-box application of automatic differentiation however is likely to produce inefficient code for the adjoints when stiff, nonlinear chemistry is included [Elbern, et. al. 1997]. This research will investigate creative applications of automatic differentiation that produce efficient code, yet require minimal user intervention. Specifically, automatic differentiation will be used to generate the transport part of the adjoint model. The computationally intensive numerical solvers will be implemented separately in an efficient manner (see below).

Construction of chemical adjoint models using the kinetic preprocessor. The kinetic preprocessor (KPP) [Damian-Iordache, 1996] of Damian, Sandu, Carmichael et. al. represents the novel generation of software tools we started to develop for air quality modeling. KPP considers the chemical equations written in standard chemical form; it performs a complex analysis of the kinetic system (including sparsity) and produces a ready-to-run collection of Fortran/C routines able to simulate the time evolution of the chemical system. It incorporates numerical integration techniques developed by the PIs. KPP currently produces the chemical kinetic (prognosis) model; it will be extended to also construct optimal forward and adjoint sensitivity models for chemical kinetics, and also for chemical equilibrium models. KPP will generate the variational (forward) and adjoint equations (this requires direct calculation of products of the form Jacobian times vector, without explicitly building the Jacobian). Sparse techniques will be implemented to handle adjoint equations. KPP will be extended to generate second order derivatives, useful in the optimization process. The KPP numerical methods libraries will be added methods for direct decoupled integration (Dunker, 1984) and for backwards integration of the adjoint [Elbern, et. al. 1997]. Optimization routines will also be included in the KPP numerical libraries to produce a self-contained package for data assimilation on chemical box models.

A library of standard tangent linear models. In addition to chemical systems, whose tangent linear models can be conveniently generated by KPP, different other subsystems could be implemented together with their tangent linear models. The user may choose to use one of the available modules as an alternative to applying automatic differentiation to the existing model.

A library of adjoint integration software. Building blocks will be developed for numerical adjoint computations. These are needed since the direct application of automatic differentiation acts both on the model and on the numerical method and are likely to produce inefficient code. Special attention will be paid to the implementation of the linear algebra solvers used in reverse adjoint integration. The natural decoupling of the problem (like in the direct decoupled method) allows for the reuse of LU decompositions; and the inherent sparsity permits further computational savings. The library will contain: 1) standard transport algorithms (of finite volume and discontinuous Galerkin type) together with the corresponding adjoint discrete models; 2) standard methods for implicit time integration of stiff systems (of Gear, Runge-Kutta and Rosenbrock type), together with their adjoint formulations; 3) solvers for particle dynamics and their adjoint formulations.

4.2.3. Handling Large Data Sets

Earth-science data assimilation is characterized by immense complexity and large data sets. It is important to develop interfaces between models and databases that enable the automatic retrieval of background concentrations and observational data, and automatic construction of assimilated data sets.

Data Set Storage Formats. Meta-data formats are available for storing large atmospheric chemical datasets. Widely used formats like [NetCDF] with the [IOAPI] interface and [HDF5] provide array-oriented data access

and define machine-independent formats for representing scientific data. A standardized (set of) format(s) will be used for observational, input, and assimilated data sets, as well as for the checkpoint information. A standard interface to the data format of choice will be developed and provided within the framework.

Data interpolation tools. Data must be interpolated from the data file grid to the computational grid. A set of interpolation routines will be provided, making it possible for a generic model to easily utilize the existent data sets. Different models may use different coordinate systems (e.g. sigma vs. altitude vs. pressure for the vertical direction, etc). The interpolation routines will also include conversions of data between different coordinate mappings. Estimates of the interpolation errors are necessary for assimilation purposes (building the functional) and will be produced by the interpolation routines.

Construction of the cost functional. A library that assists building the cost functionals in (2) and (3) will be implemented. One function of the library is to extract the observations y_k^o from the data files. The other function is to compute the model equivalent $H_k(y_k)$ from the grid data, i.e. to define the measurement operators H_k for the available measurement types. These data types may be one space dimension as is the case for surface measurements, two space dimensions as is the case for column quantities such as total column ozone or aerosol optical depths from satellites, or three-dimensional as is the case for *in situ* measured quantities on aircraft.

4.2.4. Checkpointing capabilities

The adjoint sensitivity calculation requires a reverse integration of the adjoint system. For this the forward trajectory must be saved, more exactly one must periodically store the state of the system during forward integration (checkpoints). The checkpoint information contains the concentrations at each gridpoint, as well as information about the grid. If a coupled meteorology-chemistry code is used than the meteorological variables (wind fields, temperatures and humidity) also need to be stored. The checkpoints are therefore large temporary data sets. The framework will provide interfaces for creating and reading checkpoints, using optimal storage schemes [Griewank 1992].

4.2.5. The optimization library

A library of optimization routines will be made available within the framework. The library will include efficient implementations of both standard, existing algorithms (LBFGS, TN) and of newly developed hybrid algorithms (as discussed in section 4.1.3.). The optimization routines will be directly interfaced with the direct and adjoint models.

4.2.6. The application algorithms library

An application-specific set of algorithms will be implemented. Such algorithms will support particular applications of adjoints for data assimilation and inverse modeling like optimal emission reconstruction; optimal emission control; optimal location of (adaptive) observations, etc. The users of the framework will be able to program their own applications at this level.

4.2.7. Parallelization

In order to manage the huge computational burden of data assimilation parallelization, of the model and its adjoint becomes very important. In [Miehe et. al. 2001a & 2001b] we developed a communication library for the parallelization of the direct model, which was successfully applied to STEM-III. This work will be extended to include the parallelization of the adjoint model and of the optimization algorithm. The primary platform for the development work will be Beowulf clusters at Michigan Tech and at Iowa. The layered software architecture will be exploited to isolate the data decomposition and message passing components of the implementation. This will allow the code to run efficiently on a variety of platforms that support MPI. We will investigate parallelization techniques for the adjoint model that use the same data decomposition as the direct model. Checkpointing strategies which use the local (node) storage will be considered; this decreases the communication overhead, and allows larger data sets to be stored, meaning that fewer repeated computations are needed by the adjoint. The performance data collected between the software layers will provide the information needed to optimize the code with respect to data decomposition, load balancing, granularity and do so across the significant parameter space spanned by the Beowulf clusters and supercomputers at NCAR.

4.2.8. Visualization interfaces

Visualization tools will play an important role in the realization and utilization of optimal estimates of the chemical analysis state. The size and complexity of the analysis data sets have greatly increased over the last decade are expected to continue to grow. Tools that integrate advanced visualization hardware and interactive software to create collaborative virtual environments that allows the user to create, view, navigate, and interact with data, models, and images in an immersive 3-D environment, are needed. We plan to build upon our experience in using virtual reality tools for atmospheric applications. We have experience applying virtual reality

tools, through the use of an ImmersaDesk obtained through a NASA project focused on collaborative research environments for environmental applications. Most recently we have utilized new tools developed by personnel from the Atmospheric Sciences Data Center (ASDC) at NASA's Langley Research Center (the Virtual Global Explorer and Observatory (VGEO) software), developed by VRCO, in the recently completed NASA TRACE-P field experiment, a major activity within NASA's Earth Science Enterprise [(GTE) http://www-gte.larc.nasa.gov]. We demonstrated how model products from multiple models and observations from a variety of platforms could be merged together and visualized in a single virtual environment (See Figure 3). In this study we plan to build upon these experiences, as well as ongoing efforts by the visualization group at UCAR, to make visualization an integral part of the assimilation and collaborative analysis framework.

We plan to devise innovative uses of visual tools to allow the meaningful comparison of measured, prognosed, and assimilated fields. This will put the measurements in the context and will enhance the interpretation of observational and simulation data. Specifically, we will use the Vis5D system (http://www.ssec.wisc.edu/~billh/vis5d.html) that can visualize both gridded and irregularly located data, accepts streamed data over the network, and uses multiple threads for parallelization. Cave5D supports cave environments. Our visualization subsystem will be developed on top of the Vis5D core, which can be accessed directly via its API. The visualization subsystem will contain data visualization interfaces, visual data mining interfaces (see below), and will work with the framework's standard data structures and data file formats.

4.2.9. Data mining aspects

Interfaces will be developed for visual data mining of the assimilated fields. We will use the Vis5D capabilities to allow the users to request cross-sections of the data, isosurfaces, zoom in on regions of interest, or follow trajectories (of virtual flights) through the data. Users could also look for sensitivity information and influence function values. The interface will be web-based, which will enable the sharing of results with remote community members.

4.2.10. The role of high bandwidth connections

The University of Iowa, Caltech, Michigan Technological University, the University of Washington, and the University of Minnesota are Internet-2 institutions, and NCAR is also on a high-speed network. This will allow observational data from a remote location to be assimilated using a local model, and without having to download and store the full data set in advance. Production runs at one or more locations will send the results directly at the visualization site. Thus data from multiple sources could be integrated into a single data analysis platform.

The assimilated fields created will stored for future use. A database of measured, predicted, and assimilated fields will be made available online. An XML interface will be employed to organize this database and facilitate data access via the web. The database will contain visual demos of each dataset; these visual "labels" will be available online, enhancing the database search.

4.3 Research Applications

4.3.1 Models

The techniques and tools described above will be developed, tested and used in applications in contemporary CTMs. Three well known CTMs have been developed by the investigators in this proposal and are available for use. These include the CIT (California Institute of Technology Urban Air Shed) Model [Meng, et al. 1998]; the STEM model [Song & Carmichael, 2001] and the HANK model [Hess at al., 2000]. These models are presently being used in a variety of urban and regional scale prognostic analysis and forecasting activities. As part of an on-going collaboration between Professors Seinfeld and Carmichael under the AceAsia experiment, elements of the CIT and STEM models are being joined into a single analysis model (STEM-CIT). This model will serve as the primary test-bed for this project. This model combines detailed transport and gas phase chemistry, aerosol size and composition, and radiative and optical property estimation in a single modeling system. However the tools that we develop will be general and can be used by any CTM (urban, regional or global).

To demonstrate and facilitate the transfer of these tools to the general modeling community, we plan to employ these techniques in the WRF (Weather Research Forecasting) model now under development. This model is an ideal platform for incorporating data-assimilation techniques so as to make accurate chemical forecasts and predictions. The overall goal of the WRF (Weather Research Forecasting) Model project is to develop a next-generation mesoscale forecast model and assimilation system that will advance both the understanding and prediction of important mesoscale precipitation systems, and promote closer ties between the research and operational forecasting communities. The model is being developed as a collaborative effort among several NSF, Department of Commerce and DOD sponsored institutions together with the participation of university scientists. The unique aspect of this model is it is designed as both a research tool and an operational forecast model by

NCEP (The National Centers for Environmental Prediction), who provide the operational support for the civilian weather prediction for the United States. The WRF model is currently operational and is being run in real time at NCAR and a number of other institutions. Currently WRF-chem, the coupled chemistry component of WRF, is under development, with a working group comprised of Peter Hess (head), and representatives from several Universities, NOAA, the EPA, the North Carolina Supercomputing center and the Air Force. WRF-chem is intended to be an advanced state-of-the art coupled chemistry-meteorology model. It is expected that as WRF replaces the MM5 model, WRF-chem will become widely used by many of the institutions that currently use meteorology generated by the MM5 to generate meteorology for chemistry transport models (i.e., the EPA, university researchers and other institutions). The WRF-chem prototype model is expected to be completed by September 2002. NOAA will use this model for real time chemical weather forecasts.

4.3.2 Analysis of Air Pollution in the Los Angeles Air Basin

We will apply the adjoint methodology to the analysis of air pollution in the Los Angeles air basin (the South Coast Air Basin of California). This airshed remains the most polluted in the United States; it is the airshed for which the most detailed source emissions inventories exist and for which the most comprehensive urban-scale monitoring and modeling have been carried out. Historically many of the advances in atmospheric modeling have been implemented and tested on models applied to this air basin. Such advances include the full integration of size- and composition-resolved aerosols with gas-phase ozone chemistry (Meng et al., 1997). A strenuous test of the adjoint methodology will be its extension to three-dimensional atmospheric gas-aerosol models. At present, the most advanced application of the adjoint methodology is to three-dimensional ozone models by the group at Institute Pierre-Simon Laplace, C.N.R.S., Paris. (This work was reported on recently at a symposium at the University of Houston. It has not yet been published.) We will apply the adjoint formalism to three-dimensional gas-aerosol models, which has not heretofore been attempted. The mathematics of doing so is significantly more complex than that for gas-phase photochemical models, as partial differential equations governing evolution of the aerosol size distribution must be solved coupled to the gas-phase species conservation equations. The focal region will be the Los Angeles airshed.

In analyzing gas-phase ozone photochemistry, we will evaluate the sensitivity of peak ozone concentration in the airshed to variations in emissions of VOCs and NOx at different source regions. The use of the adjoint sensitivity method allows one to address directly this issue of sensitivity to source emissions rather than through brute force, one-at-a-time perturbations in sources throughout the region, which is the current way in which these ozone sensitivities are obtained. We will then extend the formalism to evaluate sensitivities of peak aerosol levels to gas-phase emissions in various parts of the airshed. Both of these applications have direct relevance to practical air pollution control decisions. This project offers the potential to provide computational tools that would add considerably to the way in which ozone and aerosol control strategies are evaluated for urban regions.

4.3.3 The Analysis of the AceAsia Experiment

The intimate connection between measurements and models in atmospheric chemistry research can be illustrated through the example of large-scale field studies. Aircraft experiments, involving multiple aircraft and as many as 100 scientists, play an important role in atmospheric chemistry. Aircraft experiments are designed to characterize atmospheric chemistry and transport in association with features of specific scientific interest (e.g., transport of pollutants in association with frontal outflow) and to test certain aspects of our understanding. The difficulty in planning such experiments lies in the fact that features of interests are usually transient in nature. Flight planning relies on forecasting what features of interest are in the region and where they are expected to be at the time of the flight. Traditionally, flight planning has relied on meteorological forecasts alone. CTMs are just beginning to be used in forecast-mode to enhance flight planning by enabling the representation of important three-dimensional atmospheric chemical structures (such dust storm plumes, and polluted air masses associated with large cities or widespread biomass burning events) and how they evolve over time. With this added information, it is hoped that the expensive field-deployed resources (facilities and people) can be employed/utilized more effectively, and science successes maximized. CTM forecasts play the additional important role of providing a contextual representation of the experiment, and facilitate a quick analysis of the field results. The data obtained along the flight tracks for specific experiments (of typical duration of 8 hrs) provide the “real” representation of the atmosphere at those specific points in time and space. The models predict the time evolution and three-dimensional structures within the entire region of operation during the proposed measurement period. When measured and modeled data are viewed together, the context of the observations is elucidated. For example, one can see the sources of the air mass intercepted and measured by the aircraft and where it was headed. In addition, when viewed over the entire field period (typically weeks), the combined data

set of measured and modeled quantities allows for an assessment of how typical the observed features are (i.e., is the observed feature unusual or do events like it occur with some regularity). This is illustrated in Figure 3, where modeled dust and CO distributions are shown along with a flight path and measured CO values.

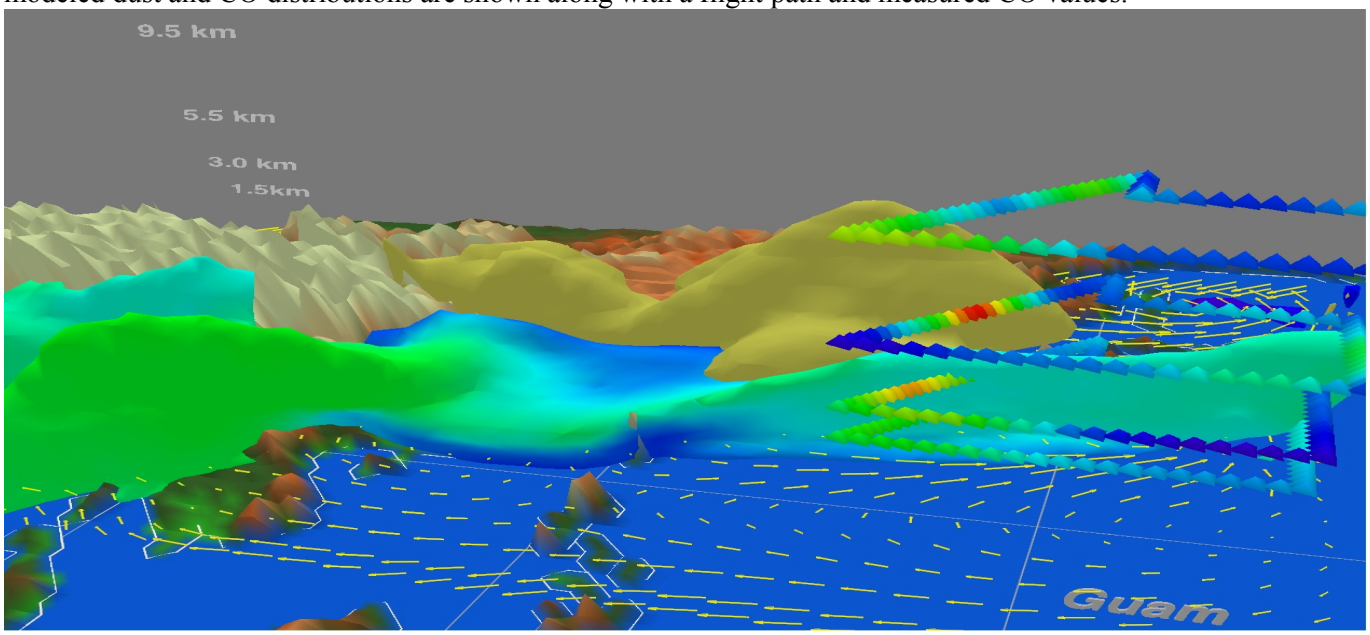

Figure 3. The use of visualization to help the integration of measured and modeled data. Shown are measured CO along the flight path, the brown isosurface represents modeled dust (100 ug/m3), and the blue isosurface is CO (150 ppb) shaded by the fraction due to biomass burning (green is more than 50%).

The recent NSF-funded large field experiment ACE-Asia (Aerosol Characterization Experiment) provides an excellent case study to examine the effectiveness of the data assimilation methods and tools. This experiment involved multiple ground based measurement sites, three airplanes, one ship, and the coordination with various satellite measurements. AceAsia produced a high density of data from multiple independent measurements, and a wide variety of different types of data. These include: a) in-situ measurements from ground stations, on board ships, and on various aircraft platforms; b) remote sensed data from satellites; and c) vertical distributions from ground based lidars. This data spans vast temporal and spatial scales, and illustrates the needs and challenges of integrating these data into a consistent data analysis set. Here we plan to apply the tools developed to integrate the measured and modeled data into a comprehensive data set for analysis by measurers and modelers. Examples of specific questions to be pursued are discussed below.

Aerosol direct forcing analysis of ACE-Asia. Uncertainty concerning aerosol forcings of climate is a dominant limitation in our current ability to interpret the industrial-era climate record and predict the future climate. This uncertainty arises to a very great extent from the complexity and variability of the aerosol phenomenon, which leads to great difficulty in implementing observational strategies adequate to constrain the forcing. The ACE-Asia experiment was designed to improve knowledge of Asian aerosols and to focus on one part of the forcing problem – namely, direct (or clear-sky) radiative effects. Application of the adjoint/variational technique to the ACE-Asia data set will include assessment of direct radiative forcing by Asian aerosols and attribution of that forcing to major chemical components (dust, sulfate, and carbonaceous.) The variational approach will provide a unique opportunity to systematically address many questions of interest to the aerosol measurement community. How well do the measurements used in ACE-Asia constrain the calculated forcings? Which measurements are most important, in terms of both type of measurement and their spacing? Which sources of measurement error were responsible for the largest uncertainties? Given practical limitations, what set of measurements (again, in terms of both type and sample frequency) would most effectively constrain the forcing? Or, put another way, if science or policy objectives require that we be able to measure regional-scale direct forcings to a specified accuracy, what density, frequency, and type of measurements would be needed to achieve this? Answers to such questions will be critical both to assessing current knowledge and to designing optimal research strategies in the future.

Here we plan to focus first on optical properties. Very large amounts of consistently acquired data are available from ACE-Asia (one ship, up to 7 surface stations, up to 4 aircraft - though not all parameters are

measured at all platforms). Using the adjoint sensitivity analysis and data assimilation techniques we plan to identify how well these optical measurements constrain the chemical models and what would be the highest leverage measurements to add to the suite for improving the constraint.

Coupling *in-situ* and satellite data for quantifying climate forcing by aerosols. A major transition in aerosol research is underway. Beginning with improved nadir and multiangle radiometers aboard NASA's Terra satellite (MODIS and MISR), scanning absorption spectrometry (SCIAMACHY/ENVISAT-1), polarization radiometry (POLDER/ADEOS-II), enhanced occultation radiometry (SAGE-III), and active lidar (ICESat/GLAS, CALIPSO), current decade is witnessing a quantum leap in both the quantity and quality of aerosol data from satellites. This satellite data will play an increasingly dominant role, yet unless it is coupled to *in-situ* measurements – necessary, for example, to provide chemical characterization and thus attribution to sources - its value will be severely limited. For these reasons, designing strategies for coupling *in-situ* and satellite observations stands as a major challenge confronting the aerosol research community. To meet this challenge, extreme differences in spatial/temporal resolution and coverage will have to be addressed. In terms of chemical sampling, for instance, surface data will clearly not suffice, since radiatively important aerosols are distributed throughout the column; however, airborne platforms are extremely limited, for cost reasons, to either very simple routine measurements or very sparse comprehensive measurements. Thus, this is a complex optimization problem, well suited to the sensitivity analysis provided by the adjoint/variational approach. The ACE-Asia experiment will serve as an outstanding test-bed for addressing this problem, since state-of-the-art *in-situ* measurements were deployed across a variety of ground, ship, and airborne platforms and great care was taken to acquire data coincident with all of the new-generation satellites now in orbit.

4.3.4 Forecasting and Related Activities

Forecasts of atmospheric chemical concentrations are becoming of increased importance. At the moment a variety of methods (statistical, expert systems, and CTM-based) are being used to forecast air quality and support air pollution abatement strategies in certain regions (over 120 cities in the US are issuing air pollution forecasts). An increased ability to forecast air pollution has obvious important societal consequences. Furthermore the use of chemical forecasting in support of comprehensive atmospheric chemistry and air pollution studies is becoming the standard mode of operation. In addition, forecasting provides one of the best model evaluations, as the model cannot be tuned in advance. The purpose of this focus area is to investigate the effect of our data assimilation tools to improve our ability to produce chemical forecasts. We will address the effect of assimilating surface and satellite measured concentrations on chemical forecasts. In addition we will address the question of the optimal measurement strategies for making chemical forecasts.

One activity will focus again on Los Angeles. This air shed has an excellent measurement network, and provides the best opportunity to evaluate the effectiveness of data assimilation on model performance. Here we will perform a systematic evaluation of the impact of the assimilation of individual chemical species and set of species on model prediction skills as outlined in our paper [Daescu et al., 2001]. We will test and demonstrate the operational aspects of our assimilation tools, and also provide guidance on what species to focus assimilation efforts on to improve operational forecast of air quality.

Another key element in chemical data assimilation is the fact that while there exists an infrastructure of meteorological measurements (e.g., surface observation network, the global radiosonde network, and a host of dedicated satellite observations) to be assimilated into meteorological models, the chemical measurements are much more sparse. The number of observations available for chemical assimilation is typically several orders of magnitude smaller than the number of variables in the model. Thus, the spatial and temporal distribution of the observations plays an essential role in the effectiveness of the data assimilation process. A critical question for the future of chemical forecasting is the design of observational strategies to support these efforts. The tools we plan to develop can be effectively used to help design such strategies. Specifically we plan to apply the 4D-Var tools and influence function analysis to the problem of optimal network designs and adaptive measurement strategies (see Dasceu et al., 2001). These studies will look at measurements to support long term forecasting efforts at a fixed location (such as urban air quality), and also at designing strategies for intensive field operations (e.g., identification of critical quantities to be measured, additional spatial locations to conduct measurements, and to critically deploy resources – if dropsondes are to be deployed from the aircraft–where and when).

Another aspect to be addressed is related to emission estimates in support of forecasting activities. Emissions are a critical element in the atmospheric chemistry and air quality studies. They often represent a significant source of the uncertainty in the analysis chain. The 4D-Var and adjoint techniques developed in this research can

be deployed to optimal estimate of the emissions distribution (i.e., the inverse problem discussed previously). We plan to test and demonstrate these capabilities. Biomass burning is an excellent example of the critical role of emissions. Air quality in many parts of the world is negatively impacted by emissions associated with biomass burning activity. The forecasts of the smoke intensity and future location of the plume depends explicitly on the location and magnitude of sporadic fires. Assimilation of satellite data on fires, burned area, and tropospheric constituents such as CO and aerosol optical depths can be used to produce optimal estimates of the emissions. Here we plan to collaborate with Dr. Wei Min Hao of the USFS who is interested in forecasts of forest fire emissions using up-to-the minute satellite observations of burned areas within the WRF-chem model.

5. INTEGRATION OF RESEARCH AND EDUCATIONAL ACTIVITIES AND DISSEMENATION OF RESULTS

The research conducted in this proposal will engage directly ~10 graduate students and post doctoral fellows at six different institutions (and indirectly many more students within these research groups) in the area of chemical data assimilation. We plan to engage the students in a variety of educational activities. For example, we plan to hold monthly seminars via the use of video conferencing. We also plan for annual research workshops, where we get together to present results and plan additional activities. We also plan that during the course of this research and monthly seminar meetings we will be developing lecture materials related chemical assimilation and its applications. We plan to produce a textbook on the subject by the end of the project period.

In addition to the workshop plans described above, the algorithmic and software tools developed during this research will be published in specialized journals. The software tools will be made available on-line via the Internet. We expect that users from both the research community and operational community will benefit from this research. We also plan to develop Web-based visual interfaces to allow remote access to assimilated data fields. We also plan to make available our assimilation and analysis tools to a broad community. We have already demonstrated this ability through our KPP software. This tool is presently being used by groups worldwide. This group contains atmospheric chemistry and air pollution communities, but also industry and researchers in pharmaceutics, including CWI (Amsterdam), TNO (Amsterdam), DIMA - Institut Francais du Petrole (Paris), Politecnica di Milano, NERI (Roskilde, Denmark), Institut for Tropospheric Research (Leipzig, Germany), Technical University of Madrid (Spain), BISA (Brussels), Solaise group (IFP - Lyon), Universitaet zu Koeln, Inst. fuer Geophysik und Meteorologie (for project EURAD), VITO - Flemish Institute for Technological Research, CETAP - Center for Remote Sensing and Atmospheric Processes (Boeretang, Belgium), Max-Planck Institute fur Chemie (Mainz, Germany).

We also plan to interact closely with other related data assimilation efforts through the participation and organization of data assimilation workshops. For example, next summer of group of people at NCAR is running a summer school on data assimilation and network design for carbon cycle and biogeochemistry. The plan is to build a simple data assimilation system for CO_2 on the regional scale and use HANK and its adjoint as the basis. The attendees will try to design a measurement network design using the synthetic data and the data assimilation system provided. There is also a plan to propose to host at NCAR an ASP summer colloquium on data assimilation in 2003. The present focus is on meteorological data assimilation, but the addition of a chemical component could be added.

The developed tools will be available as part of WRF-chem modeling system and will be made available to the community through this community model. That the WRF model can be downloaded off the NCAR web site and the WRF-chem model will also be able to be downloaded in the future. We also plan to conduct chemical data assimilation workshops in connection with the WRF-chem activities. There is a great deal of interest in this model, and many development and user group meeting are planned in support of the WRF-chem development over the next several years. We will hold workshops to teach people how to use the assimilation tools developed here in WRF-chem. Though these activities, the techniques developed in this proposal will be widely distributed and used.

Finally, the 4D-Var and adjoint techniques advanced in this research will be of use to fields outside of atmospheric chemistry and air quality. The fields of chemical reactor design and combustion utilize very similar model formulations and numerical techniques. More broadly PDE control problems are beginning to explore uses of 4D-Var and adjoint techniques.

MANAGEMENT PLAN & RESEARCH TEAM

This proposal is a collaboration between: Professor Gregory R. Carmichael at the Center for Global & Regional Environmental Research and Department of Chemical Engineering, The University of Iowa; Professor John Seinfeld, Department of Chemical Engineering, at the California Institute of Technology; Professor Adrian Sandu, from the Department of Computer Science at Michigan Technology University; Dr. Peter Hess at the National Center for Atmospheric Research; Dr. Tad Anderson, Department of Atmospheric Sciences, The University of Washington; and Dr. Dacian Daescu, Postdoctoral Fellow at the Institute for Mathematics and its Applications (IMA) at the University of Minnesota. This proposal builds upon the extensive modeling and analysis activities and experiences of the investigators. Members of the team have extensive experience in building and applying contemporary coupled chemical transport models (CTMs). The CIT, STEM, and HANK models have been used extensively to analyze observational data, and are being used in forecast modes to support major field studies (TOPSE, ACE-Asia and TRACE-P). The research team also includes the disciplines of chemical engineering, computer science, applied mathematics, and atmospheric science and meteorology, and expertise in chemical and aerosol field and laboratory measurements, data assimilation, computational science, numerical analysis, and high-performance computing. The research will involve ~10 PhD and postdoctoral fellows.

Professor Carmichael will coordinate the overall management of the project, and he will also be responsible for the application and testing of the assimilation techniques within the STEM model and their application to the ACE-Asia studies and to improve forecasting capabilities, and will coordinate the high performance computing elements. Dr. Hess will be responsible for the integration of these techniques into the WRF model, to applications to inverse modeling of emissions, and to the coordination or workshops at NCAR. Professor Sandu will have overall responsibility for the numerical algorithms and software components. Professor Seinfeld will lead the applications for the LA airshed and AceAsia, and provide oversite to the aerosol modeling elements. Dr. Tad Anderson will have primary responsibility for comparison of model and measured optical properties and the evaluation of how to best use optical measurements in data assimilation, He will also provide guidance to the project on the use of satellite products. Dr. Daescu will play a leading role in developing improved 4D-Var and adjoint modeling algorithms, and on better techniques for the design of optimal measurement strategies. He is presently at IMA where for the next few years they have a focus on assimilation methods for geoscience applications, where he focusing his research efforts on 4D-Var and adjoint techniques. (His participation is at no cost to this project.)

This research effort builds upon established collaborations and on-going efforts. For example: Carmichael, Seinfeld and Anderson are all participants in the AceAsia experiment and are working together on the analysis and interpretation of the data; Hess and Carmichael are on the WRF-chem working group that is advising the development of this model; and Sandu, Daescu and Carmichael are working closely together on numerical and software issues related to improved air quality modeling. In addition research underway in active research grants related to AceAsia analysis and improved numerical techniques for aerosol modeling will complement the proposed work, which brings together this group with a distinct focus on assimilation analysis. Several of these on-going projects are currently funded by NSF (see Current & Pending Section). However none of the completed projects have dealt directly with the subject of this proposal. As a result we have not included a formal Prior Support section.

The project is planned as a five-year effort. Efforts related to improved numerical methods, software tools and science applications would begin immediately, and continue throughout the five-year period. A general overview of the work plan is presented in Table 1. The primary research activities related to improved numerical methods and science applications will feed the software tools development. We already have in place some elements of the software system for assimilation in our KPP toolbox. Our numerical methods development will be designed as to produce software tools for inclusion in this framework. Software elements will be added all along the project timeline. The science applications will also serve to develop various aspects of the software system. Issues related to data formats, visualization, etc., will best be addressed in the context of the application analysis. The general strategy is as follows: 1) to assemble an initial set of the tools to enable science applications to begin; 2) begin the application studies using the available tools; 3) continue to develop and refine the assimilation techniques, and 4) tackle more complex science applications as the tools become more sophisticated.

This research effort will utilize various means to facilitate collaboration. We plan to hold monthly seminars via the use of video conferencing. We also plan for annual research workshops, where we get together to

present results and plan additional activities. Science meeting for the investigators and their students will be held at least once per year. It is also envisioned that individuals may spend extended time periods in another collaborators laboratory. We also plan to exploit Internet 2 capabilities at the host institutions to test and move the large data sets associated with this research. We plan to carry out the computational aspects of the project on Beowulf clusters at Michigan Tech. and the University of Iowa, and to perform large applications at the Supercomputing facilities at NCAR.

Table 1. Timeline of proposed research activities.

RESEARCH ACTIVITY	***Started***	***Year 1***	***Year 2***	***Year 3***	***Year 4***	***Year 5***
Improved Numerical Techniques						
Inexpensive step derivatives	***x***	←-----	--------	-----→		
Efficient tracking of numerical errors		←-----	--------	----→		
Backwards in time integrators			←----	-------	-----→	
Construction of Discrete Adjoints						
Nested and adaptive grids				←-----	------	----→
Stiff systems	***x***	←-----	-------	-----→		
Aerosol dynamics	***x***			<----	--------	-----→
Aerosol optical properties			←---	-------	-----→	
Improved Optimization Algorithms	***x***		←----	-------	----→	
Adjoint tools for Sensitivity Analysis	***x***	←----	--------	--------	------	---→
Adjoint tools for inverse modeling		←-----	--------	-----→		
Adjoint tools for targeting measurements	***x***			←-----	-------	-----→
Software Tools						
Design of the General Framework	***x***	←-----	-----→			
Adjoint building tools	***x***	←-----	-------	--------	--------	-----→
Standard Libraries	***x***	←-----	--------	--------	--------	----→
Data handling	***x***	←----	------	-------	-------	----→
Formats	***x***					
Construction of cost functionals	***x***		←----	--------	-----→	
Libraries	***x***		←-----	-----	-------	----→
Parallelization	***x***	←-----	-------	-----→		
Science Applications						
Air Quality of LA	***x***					
Adjoint sensitivity analysis of emissions		←-----	-----→			
Optimal emission reductions				←----	--------	---→
AceAsia	***x***					
Optimal analysis state		←-	-------	-------	-------	---→
Adjoint sensitivity of aerosol optical propert.		←-----	--------	--------	----→	
Assimilation of satellite data			←-----	-------	--------	----→
Forecasting Applications						
Inverse emissions modeling		←-----	-------	-------	----→	
Improved air quality forecasts in LA				←---	--------	-----→
Targeting measurements	***x***			←-----	-------	----→

X – designates that some activity is already taking place upon which this research will build. For example, the AceAsia applications we already are running the CTM to support data analysis. However no assimilation activities are underway.

REFERENCES

Blom, J.G. and J.G. Verwer, VLUGR3: A Vectorizable Adaptive Grid Solver for PDEs in 3D. I. Algorithmic Aspects and Applications, *Appl. Numer. Math.*, Vol. 16, p. 129-156, 1994.

Carmichael, G.R., A. Sandu and F.A. Potra, Sensitivity Analysis for Atmospheric Chemistry model via Automatic Differentiation, *Atmospheric Environment*, Vol. 31, p. 475--489, 1997.

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

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

DAGH: Distributed Adaptive Grid Hierarchy, http://www.cs.utexas.edu/users/dagh.

Daescu, D. G.R. Carmichael, and A. Sandu, Adjoint Implementation of Rosenbrock Methods Applied to Variational Data Assimilation Problems. *J. of Computational Physics*, Vol. 165, p. 496--510, 2000.

Daescu, D.N. and G.R. Carmichael, An adjoint sensitivity method for the adaptive location of the observations in air quality modeling, Submitted to *J. Geophys. Res.*, 2001.

Damian-Iordache, V., KPP - A symbolic preprocessor for chemistry kinetics, Master Thesis, Department of Computer Science, The University of Iowa, Iowa City, IA 52246, 1996.

Damian, V., A. Sandu, M. Damian, F. Potra, and G.R. Carmichael, The Kinetic PreProcessor KPP -- A Software Environment for Solving Chemical Kinetics, submitted to *Computers and Chemical Engineering*, Computer Science Technical Report CSTR-01-07, Michigan Technological University, 2001.

Data Assimilation Office, NASA, Overview of GEOS Data Assimilation System (DAS), http://dao.gsfc.nasa.gov/subpages/atbd_chapters/node11.html, 2001.

Dunker, A.M., The decoupled direct method for calculating sensitivity coefficients in chemical kinetics, *J. Chemical Physics*, Vol. 81, p. 2385, 1984.

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

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

Elbern, H., H. Schmidt, and A. Ebel, Variational data assimilation for tropospheric chemistry modeling, *J. Geophys. Res*, 102, 15967-15985, 1997.

Elbern, H., H. Schmidt, O. Talagrand, A. Ebel, 4D-variational data assimilation with an adjoint air quality model for emission analysis, *Environmental Modeling and Software*, 15, 539-548, 2000.

Fisher, M., and D. J. Lary, Lagrangian four-dimensional variational data assimilation of chemical species, *QJRMS*, 121, 1681-1704, 1995.

Francois-Xavier Le Dimet, I. M. Navon, and Dacian N. Daescu, Second-Order Information in Data Assimilation, *Monthly Weather Review*, to appear, 2001.

Gelaro, R., R. Langland, G.D. Rohaly, and T.E. Rosmond, An assessment oof the singular vector approach to targeted observations using the FASTEX data set, *Q.J.R. Meteor. Soc.*, 125, 3299-3328, 1999.

Grell, G., J. Dudhia, and D.R. Stauffer, A description of the fifth generation Penn State/NCAR Mesoscale Model (MM5), NCAR/TN-398+IA, 116 pp., Natl. Center for Atmos. Res., Boulder, CO, 1993.

Griewank, A., Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation, *Optimization Methods and Software*, 1:35-54, 1992.

Griewank, A., Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, *Frontiers in Applied Mathematics*, 19, 2000.

Haywood, J. M., V. Ramaswamy, et al. A limited-area-model case study of the effects of sub-grid scale variations in relative humidity and cloud upon the direct radiative forcing of sulfate aerosol, *Geophys. Res. Lett.* **24**: 143-146, 1997.

He, S. and G.R. Carmichael, Sensitivity of photolysis rates and ozone production in the troposphere to aerosol properties, *J. Geophys. Res.*, 104: 26,307-26,324, 1999.

He, S., G.R. Carmichael, A. Sandu, B. Hotchkiss, and V. Damian-Iordache, Application of ADIFOR for air pollution model sensitivity studies, *J. Env. Modeling*, 15: 549-557, 2000.

HDF5: The Hierarchical Data Format, http:// hdf.ncsa.uiuc.edu

IOAPI: Models-3/EDSS Input/Output Applications Programming Interface http://envpro.ncsc.org/EDSS/ioapi

Jacobson, M.Z., *Fundamentals of atmospheric modeling*, Cambridge University Press, 1999.

Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, 1970.

Joe, B., Geompack, http://members.attcanada.ca/~bjoe/index.htm, 2001.

Kasibhatla, P., R. Menard, L. P. Chang, J. Logan, and P. Novelli, Assimilation of Measurement of Air Pollution from Satellites (MAPS) carbon monoxide measurements using a parameterized three-dimensional chemical transport model, submitted to *J. Geophys. Res*., 2000.

Kasten, R. Visibility in the prephase of condensation, Tellus **21**: 631-635, 1969.

KeLP: Kernel Lattice Parallelism, http://www-cse.ucsd.edu/groups/hpcl/scg/kelp.html.

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

Lamarque, J.-F., B.V. Khattatov, J.C. Gille, and G.P. Brasseur, Assimilation of Measurement of Air Pollution from Space (MAPS) CO in a global three-dimensional model, *J. Geophys. Res.*, 104(D21), 26,209-26,218, 1999.

Lamarque J.-F., B. Khattatov, and J. Gille, Constraining tropospheric ozone column through data assimilation, submitted to *J. Geophys. Res*., August 2001.

Langland, R.H., R. Gelaro, G.D. Rohaly, and M.A. Shapiro, M.A., Targeted observations in FASTEX: Adjoint-based targeting procedures and data impact experiments in IOP17 and IOP18, *Q.J.R. Meteorol. Soc.*, 125, 3241-3270, 1999.

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

Li, Z. and I.M. Navon, Optimality of 4D-Var and its relationship with the Kalman filter and Kalman smoother, *Q.J.R. Meteorol. Soc.*, Vol.127, No. 572 - January 2001, Part B, p. 661--684, 2001.

Liu, D.C. and J. Nocedal, On the limited memory BFGS method for large scale optimization, *Math. Programming*, Vol. 45, p. 503-528, 1989.

Lyster, P. M. et al., A Lagrangian trajectory filter for constituent data assimilation, Report DAO, Goddard Laboratory for Atmospheres, January 2000.

Lyster, P. M., S. Cohn, R. Menard, L.-P. Chang, S.-J. Lin, S.-J, and R. G. Olsen, Parallel implementation of a Kalman filter for constituent data assimilation, *Monthly Weather Review*, Boston, 125, 1674-1686, 1997.

Menard, R. and L.-P. Chang, Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part II: validated results and analysis of variance and correlation dynamics, *Monthly Weather Review,* 128, 2672-2686, 2000.

Menard, R. S. E. Cohn, L.-P. Chang and P. M. Lyster,Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part I: Formulation, *Monthly Weather Review*, Boston, MA, 128, 2654-2671, 2000.

Meng, Z., D. Dabdub, and J. H. Seinfeld, Chemical Coupling Between Atmospheric Ozone and Particulate Matter, *Science* 277, 116-119, 1997.

Meng, Z., D. Dabdub, and J.H. Seinfeld, Size-Resolved and Chemically Resolved Model of Atmospheric Aerosol Dynamics, *J. Geophys. Res.*, **103**, 3419-3435, 1998.

Miehe, P., A. Sandu, PAQMSG: A Parallelization Library for Air Quality Models on Structured Grids, http://www.cs.mtu.edu/~asandu/Software/Parallel, 2001a.

Miehe, P., A. Sandu, G.R. Carmichael, Y. Tang, and D. Daescu, A Communication Library for the Parallelization of Air Quality Models on Structured Grids, submitted to *Atmos. Envir.* Computer Science Technical Report CSTR-01-05, Michigan Technological University, 2001b.

Morales, J.L. and J. Nocedal, Enriched Methods for large-scale unconstrained optimization, To appear in *Computational Optimization and Applications*, 2000.

Mulholland, M. and J. H. Seinfeld, Inverse Air Pollution Modeling of Urban-scale Carbon Monoxide Emissions, *Atmospheric Environment* 29, 497-516, 1995.

Navon, I.M., Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography, *Dynamics of Atmospheres and Oceans*, Special issue in honor of Richard Pfeffer, 27, No. 1-4, 55-79, 1998.

NetCDF: network Common Data Form, http://www.unidata.ucar.edu/packages/netcdf

Ollivier-Gooch, C., et al., GRUMMP --- Generation and Refinement of Unstructured, Mixed-Element Meshes in Parallel, http://tetra.mech.ubc.ca/GRUMMP/, 2001.

Overture: Object-Oriented Tools for Solving CFD and Combustion Problems in Complex Moving Geometry, http://www.llnl.gov/CASC/Overture

Palmer, T.N., Gelaro, R., Barkmeijer, J., and Buizza, R., Singular vectors, metrics, and adaptive observations, *J. Atmos. Sci.*, 55, 633-653, 1998.

Petch, J.C. Using a cloud resolving model to study the effects of sub-grid scale variations in relative humidity on direct sulphate aerosol forcing. *Q. J. R. Meteorol. Soc.* **126**: 1-9, 2001.

Pu, Z., E. Kalnay, J. Sela, and I. Szunyogh, Sensitivity of forecast error to initial conditions with a quasi-inverse linear method, *Monthly Weather Review*, 125, 2479-2503, 1997.

Rabier, F., E. Klinder, P. Courtier, and A. Hollingswirth, A., Sensitivity of forecst errors to initial conditions, *Q.J.R. Meteorol. Soc.*, 122, 121-150, 1996.

Riishojgaard, L. P., I. Stajner and G.-P. Lou, The GEOS Ozone Data Assimilation System, *Advances in Space Research*, 25(5), 1063-1072, 2000.

SAIC, The Operational Multiscale Environment model with Grid Adaptivity Omega[tm], http://www.saic.com/omega/, 2001.

Sandu, A., F.A. Potra, V. Damian, and G.R. Carmichael, Efficient implementation of fully implicit methods for atmospheric chemistry, *J. Computational Physics*, Vol. 129, p. 101--110, 1996.

Sandu, A., Numerical aspects of air quality modeling, Ph.D. Thesis, Applied Mathematical and Computational Sciences, The University of Iowa, Iowa City, IA, 1997.

Sandu, A., M. van Loon, F.A. Potra, G.R. Carmichael, and J. G. Verwer, Benchmarking stiff ODE solvers for atmospheric chemistry equations I - Implicit vs. Explicit, *Atmospheric Environment*, Vol. 31, p. 3151--3166, 1997.

Sandu, A., Positive numerical integration methods for chemical kinetic systems, *J. of Computational Physics*, accepted, 2001.

Sandu, A., A Spectral method for solving aerosol dynamics, submitted to *Applied Numerical Mathematics*, Computer Science Technical Report CSTR-01-04, Michigan Technological University, 2001.

Sandu, A., A Newton-Cotes quadrature approach for solving the aerosol coagulation equation, *Atmospheric Environment*, accepted, 2001.

Sandu, A. and C.T. Borden, A Framework for the Numerical Treatment of Aerosol Dynamics, submitted to *Applied Numerical Mathematics*, Computer Science Technical Report CSTR-01-02, Michigan Technological University, 2001.

Sandu, A., D. Daescu and G.R. Carmichael, Adjoint Data Assimilation for Aerosol Dynamics, B. Sportisse editor, Proceedings of 2nd International Conference on Air Pollution Modeling and Simulation (APMS01), Geosciences. Springer, 2001.

Sarma, A., An Application of Adaptive Grid Refinement to Atmospheric Modeling, IMA Workshop on Atmospheric Modeling, Minneapolis, March 15-19, 2000.

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

Stajner, I., L. Risshojgaard, and R Rood, The GEOS ozone data assimilation system: specification of error statistics in press, *Q.J.R. Meteorol. Soc.*, 2001.

SUMAA3D Project, Scalable Unstructured Mesh Algorithms and Applications, Argonne National Laboratory, Penn State University, Virginia Tech, and The University of British Columbia http://www-unix.mcs.anl.gov/sumaa3d/, 2001.

Todling, R. and Cohn, S.E., Suboptimal schemes for atmospheric data assimilation based on the Kalman filter, *Monthly Weather Review*, Vol. 122, p. 2530--2557, 1994.

Tomlin, A., Resolution of Pollutant Profiles in the Troposphere Using an Adaptive 3-D Unstructured Finite Volume Scheme, IMA Workshop on Atmospheric Modeling, Minneapolis, March 15-19, 2000.

van Loon, M., Numerical Methods in Smog Prediction, Ph.D. thesis, CWI, Amsterdam, 1996.

Vavasis, S., QMG: mesh generation and related software, http://www.cs.cornell.edu/Info/People/vavasis/qmg-home.html, 2000.

Verwer, J., E.J. Spee, J. G. Blom, and W. Hunsdorfer, A second order Rosenbrock method applied to photochemical dispersion problems, *SIAM Journal on Scientific Computing*, Vol.20, p. 1456--1480, 1999.

Vukicevic, T., and P.G. Hell, Analysis of tropospheric transport in the Pacific basin using the adjoint technique, *J. Geophys. Res*., 105, 7213-7230, 2000.

Wang, Z., I.M. Navon, X. Zou and F.X. LeDimet, A Truncated-Newton optimization algorithm in meteorology applications with analytic Hessian/vector products, *Computational Optimization and Applications*, Vol. 4, p. 241--262, 1995.

Wang, Z., K. Droegemeier, L. White and I.M. Navon, Application of a New Adjoint Newton Algorithm to the 3-D ARPS Storm Scale Model Using Simulated Data, *Monthly Weather Review*, Vol.125, No.10, 2460-2478, 1997.

Wang, Z., K. Droegemeier and L. White, The Adjoint Newton Algorithm for Large-Scale Unconstrained Optimization in Meteorology Applications, *Computational Optimization and Applications*, 10(3):283--320, 1998.

WRF – The Weather Research Forecasting Model (http://wrf-model.org).

Yienger, J.J., G.R. Carmichael, M.J. Phadnis, T.A. Holloway, M.K. Galanter, W.J. Moxim, and H. Levy II, The episodic nature of air pollution transport from Asia to North America, *J. Geophys. Res*., 105, 26,931-26,945, 2000.