

Article

Chemical Data Assimilation—An Overview [‡]

Adrian Sandu ^{1,*} and Tianfeng Chai ²

¹ Computational Science Laboratory, Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0106, USA

² NOAA/OAR/ARL, Silver Spring Metro Center #3, Rm. 3437, 1315 East West Highway, Silver Spring, MD 20910, USA; E-Mail: Tianfeng.Chai@noaa.gov

[‡] The paper is dedicated to the memory of Dr. Daewon Byun, whose work remains a lasting legacy to the field of air quality modeling and simulation.

* Author to whom correspondence should be addressed; E-Mail: Sandu@cs.vt.edu;
Tel.: +1-(540)-231-2193; Fax: +1-(540)-231-9218.

Received: 7 June 2011; in revised form: 9 August 2011 / Accepted: 19 August 2011 /

Published: 29 August 2011

Abstract: Chemical data assimilation is the process by which models use measurements to produce an optimal representation of the chemical composition of the atmosphere. Leveraging advances in algorithms and increases in the available computational power, the integration of numerical predictions and observations has started to play an important role in air quality modeling. This paper gives an overview of several methodologies used in chemical data assimilation. We discuss the Bayesian framework for developing data assimilation systems, the suboptimal and the ensemble Kalman filter approaches, the optimal interpolation (OI), and the three and four dimensional variational methods. Examples of assimilation real observations with CMAQ model are presented.

Keywords: chemical transport modeling; data assimilation; Kalman filter; variational methods

1. Introduction

Chemical data assimilation produces improved estimates of the chemical state of the atmosphere by combining information from three different sources: the physical and chemical laws of evolution

(encapsulated in the model), the reality (as captured by the observations), and the current best estimate of the distribution of pollutants in the atmosphere (encapsulated in the prior) – all with associated errors [1]. Considerable experience with data assimilation has been accumulated in the fields of numerical weather prediction, ocean modeling, and oil reservoir simulation [2–7]. Chemical data assimilation has started to play an important role in the atmospheric composition studies, and many successful applications illustrate its benefits [8–18]. These benefits include improved initial and boundary conditions, and refined top-down emission estimates, all contributing to better air quality forecasts. Chemical data assimilation poses specific challenges related to the multiple physical processes included in models, the stiffness of chemical equations, the sparseness of chemical observations, and the uncertainty in the anthropogenic and natural emission levels.

The chemical interactions take place on a wide range of temporal scales (from milliseconds to days). This makes the system numerically stiff. The concentrations of short lived radical species follow the concentrations of long lived species through quasi steady state relations. After a short time the chemical evolution collapses onto a low dimensional manifold in state space. As a consequence, when meteorological fields are computed off line, ensembles of simulations will tend to converge to the same trajectory. Moreover, a direct adjustment of radical species through data assimilation is not feasible.

In regional air quality simulations, the influence of the initial conditions fades in time, and the concentration fields become largely driven by emission and removal processes (and by lateral boundary conditions in regional simulations). Therefore, to improve the analysis capabilities of CTMs, it is necessary to consider the estimation of emission parameters and lateral boundaries through data assimilation [19,20]. Moreover, both the anthropologic and natural emissions are poorly constrained (*i.e.*, the prior information on emissions is highly uncertain). This makes the top down estimation of emissions a challenging computational problem. Chemical transport models are often characterized by non negligible biases, and data assimilation can benefit from bias correction schemes [21].

Chemical observations are still sparse, as the network is not as extensive as that used in numerical weather prediction. Local observations of chemical and particulate concentrations are strongly influenced by the local variability, yet they are used to constrain large scale three dimensional fields. Recently there has been a considerable growth in the available remote sensing (satellite) data on tracer concentrations. This data is characterized by non-negligible biases; a method to alleviate this issue is proposed in [22], where a single coherent dataset is created from all available ozone column measurements.

An additional difficulty arises from the multiphysics nature of the simulation, where the evolution is driven by multiple competing physical processes. A successful data assimilation system needs to correctly account for error correlations between chemical species (due to chemical interactions) and between chemical and dynamic variables (due to transport processes).

This paper gives an overview of the state of the art in chemical data assimilation. We review chemical transport models in Section 1.1 and chemical observations in Section 1.2. Section 2 is devoted to the formulation of the chemical data assimilation problem in a Bayesian framework. Practical assimilation methods discussed include optimal interpolation (OI) (Section 3.3), suboptimal Kalman filters (Section 3.2), ensemble Kalman filters (Section 3.4), three dimensional variational (3D-Var, Section 3.5) and four dimensional variational data assimilation (4D-Var, Section 3.6). Challenges to

chemical data assimilation such as data inputs, the construction of adjoints, and the construction of error covariance matrices are highlighted in Section 4. Assimilation results with real data and the CMAQ model are presented in Section 5. Section 6 draws conclusions and pinpoints to future directions in chemical data assimilation.

1.1. Chemical Transport Models

An atmospheric chemical transport model (CTM) solves the mass balance equations for concentrations $\mathbf{x}_{(i)}$ of tracer species $1 \leq i \leq s$. The tracer species can be in gas, liquid, or particulate phases, and their concentrations are continuously changed by multiple physical and chemical processes,

$$\begin{aligned} \frac{\partial \mathbf{x}_{(i)}}{\partial t} &= -u \cdot \nabla \mathbf{x}_{(i)} + \frac{1}{\rho} \nabla \cdot (\rho K \nabla \mathbf{x}_{(i)}) + \frac{1}{\rho} f_{(i)}(\rho \mathbf{x}) + E_{(i)}, \quad 1 \leq i \leq s, \quad t_0 \leq t \leq t_F \\ \mathbf{x}_{(i)}(t_0, x) &= \mathbf{x}_{(i)}^{\text{initial}}(x) \\ \mathbf{x}_{(i)}(t, x) &= \mathbf{x}_{(i)}^{\text{inflow}}(t, x) \quad \text{on} \quad \Gamma^{\text{inflow}} \\ K_{nn} \frac{\partial \mathbf{x}_{(i)}}{\partial n} &= 0 \quad \text{on} \quad \Gamma^{\text{outflow}} \\ K_{nn} \frac{\partial \mathbf{x}_{(i)}}{\partial n} &= V_{(i)}^{\text{deposition}} \mathbf{x}_{(i)} - Q_{(i)} \quad \text{on} \quad \Gamma^{\text{ground}} \end{aligned} \quad (1)$$

Here u represents the wind velocity vector, K is the turbulent diffusion tensor, and ρ is the air density. These variables are typically prescribed from simulations with a numerical weather prediction model, or are part of the prognostic variables for meteorological models with online chemistry. The concentrations $\mathbf{x}_{(i)}$ are expressed as a mole fraction (e.g., the number of molecules of tracer per 1 billion molecules of air); the absolute concentration of tracer i is $\rho \mathbf{x}_{(i)}$ (molecules/cm³). $f_{(i)}$ is the rate of transformations of species i and depends on all other concentrations at the same spatial location. Such local transformations are determined by gas and liquid phase chemical kinetics, by inter-phase mass transfer, by aerosol dynamic processes (coagulation and growth), by thermodynamic processes, *etc.* The elevated emissions of species i are $E_{(i)}$ and the ground level emissions are $Q_{(i)}$. The deposition velocity is $V_{(i)}^{\text{deposition}}$. The model has prescribed initial conditions $\mathbf{x}^{\text{initial}}$, and is subject to Neumann boundary conditions [23] at the ground level boundary Γ^{ground} . Dirichlet boundary conditions [23] are imposed at the inflow boundary Γ^{inflow} (along the top and, for regional models, along the lateral boundary as well). A no diffusive flow condition is imposed at the outflow boundary Γ^{outflow} (along the top and, for regional models, along the lateral boundary as well).

The numerical solution to (1) is represented by the discrete model

$$\mathbf{x}_i = \mathcal{M}_{t_{i-1} \rightarrow t_i}(u, K, E, Q, V^{\text{deposition}}, \mathbf{x}^{\text{inflow}}; \mathbf{x}_{i-1}), \quad i = 1, 2, \dots; \quad \mathbf{x}_0 = \mathbf{x}^{\text{initial}} \quad (2)$$

In (2), the solution \mathbf{x}_i is the discrete state vector containing the concentrations of chemical species sampled at the grid points at time t_i . The discrete initial condition \mathbf{x}_0 is obtained by sampling $\mathbf{x}^{\text{initial}}$ at the grid points. The model solution operator \mathcal{M} depends on model parameters such as emission rates, deposition velocities, and boundary fluxes. In principle, all the model parameters, as well as the initial conditions \mathbf{x}_0 , can be retrieved through data assimilation if there are enough observations. However, we have to limit the number of model parameters to be determined as the observations in reality are lacking to accurately constrain the problem.

While there are numerous CTMs available for both regional and global applications, the community Multiscale Air Quality (CMAQ) model is primarily used to provide examples in Section 5. As an open-source community model, CMAQ is widely used by the air quality community worldwide and continuously updated with support from the U.S. Environmental Protection Agency (EPA) and Community Modeling & Analysis System (CMAS) [24]. CMAQ model was developed by the U.S. EPA to meet the needs of both environmental managers and scientists to improve their ability to evaluate the impact of air quality management practices and to probe, understand, and simulate chemical and physical interactions in the atmosphere. CMAQ model has been designed to model multiple air quality issues, such as tropospheric ozone, fine particles, toxics, acid deposition, and visibility degradation as a whole; and it has capabilities to solve air quality problems in multiple scales including the urban and regional scales [25,26]. CMAQ has been used in numerous chemical data assimilation studies. A 4D-Var data assimilation system, including an adjoint model for CMAQ, has been developed for version 4.5 [27]. The assimilation of AIRNow ozone observations [28] proved to be beneficial for improving ozone predictions [29]. Zubrow *et al.* [30] presented an ensemble adjustment Kalman filter (EAKF) approach using a single tracer version of the CMAQ model to assimilate surface measurements of carbon monoxide and showed its ability to provide skillful model results.

1.2. Chemical Observations

Measurements of atmospheric chemical fields have been significantly increasing for the past years throughout the world. Many ground-based networks have been established to routinely monitor the air quality on the surface level. For instance, in the U.S. the AIRNow network has been reporting ozone and fine particle observation (PM_{2.5}, *i.e.*, particulate matter less than 2.5 micrometers in diameter) in near-real-time [28]. However, surface measurements have to be combined with vertical profiles to obtain the three-dimensional states of the atmospheric constituents. Jeuken *et al.* shows that assimilating ozone columns alone has little impact on the shape of the vertical ozone profile which is mainly determined by the transport [31]. To complement the surface measurements, there are observations regularly taken by balloon and lidar networks. In adding the vertical profiles the chemical and dynamical processes of the atmospheric chemistry can be better understood. Such networks include SHADOZ (Southern Hemisphere Additional Ozonesondes) which has operated since 1998 [32–34]. Lidar networks contribute to the atmospheric chemistry studies by providing vertically resolved data in an extended area. Using observations from 14 Japan National Institute for Environmental Studies (NIES) lidars [35] and RAMS/CFORS-4DVAR assimilation system, Yumimoto *et al.* investigated mineral dust transport and emission in East Asia during a spring dust event in 2007 [36].

In addition to the in-situ measurement networks, multiple satellite instruments with capability to measure the troposphere and stratosphere atmospheric chemical fields have been operating to provide real-time measurements [37,38]. For instance, Aura [39], a multi-national NASA Earth Observing System (EOS) satellite to study atmospheric chemistry following Terra (launched in 1999) and Aqua (launched in 2002), was launched in 2004. Aura carries a High Resolution Dynamics Limb Sounder (HIRDLS, which stopped operating in 2008), an Ozone Monitoring Instrument (OMI), a Tropospheric Emission Spectrometer (TES), and a Microwave Limb Sounder (MLS) [40]. The retrieval of Envisat-SCIAMACHY (Scanning Imaging Absorption Spectrometer for Atmospheric Chartography) by

European Space Agency provides various atmospheric constituents [41]. Moderate resolution imaging spectroradiometer (MODIS) aboard Terra (EOS AM) and Aqua (EOS PM) satellites provides near real time aerosol optical depth (AOD) observations with good spatial resolution and coverage [42,43] and they have been used in many aerosol assimilation applications [44–47]. One of the earliest efforts in chemical data assimilation was a OI type statistical analysis scheme to assimilate the Total Ozone Mapping Spectrometer (TOMS) total ozone data and the Solar Backscatter Ultraviolet/2 (SBUV/2) partial ozone profile observations into an off-line ozone transport model by Štajner *et al.* [48]. With a broad horizontal coverage, satellite observations are complementary to in-situ measurements that are often located at places of interest. Nassar *et al.* showed the benefit of combining the two types of observations together in chemical data assimilation systems by assimilating both satellite observations of CO_2 from TES and surface flask measurements [49].

In recent years, many field experiments have been carried out with intensive measurement activities. For instance, the International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) field campaign took place in the northeastern United States and the Maritime Provinces of Canada during summer 2004 [50]. Over 300 government-agency and university participants from U.S., Canada, UK, Germany, and France carried out eleven independent but highly coordinated field experiments with various objectives. Among them, the U.S. National Oceanic and Atmospheric Administration (NOAA) New England Air Quality Study - Intercontinental Transport and Chemical Transformation [51] 2004 experiment studied air quality along the Eastern Seaboard and transport of North American emissions into the North Atlantic [52]. The European International Transport of Ozone and Precursors (EU-ITOP) field program aimed at understanding the factors determining air quality over America and Europe and over remote regions of the North Atlantic [53,54]. As another component of ICARTT, NASA Intercontinental Chemical Transport Experiment - North America Phase A (INTEX-A), focused on the transport and transformation of gases and aerosols on transcontinental/intercontinental scales and their impact on air quality and climate. The continued NASA INTEX project, Phase B, coincided with MIRAGE-Mex (Megacities Impact on Regional and Global Environment-Mexico) in the spring of 2006 [55,56]. In the field experiments, coordinated measurements were made by multiple in-situ instruments on board aircrafts in flights, additional ozonesondes on ground or a research vessel, and airborne ozone lidars. Using ICARTT/INTEX-A data, Chai *et al.* [11] shows that the combined ozone observations provide a much better representation of the ozone distributions when assimilated simultaneously.

1.3. Chemical Data Assimilation

We now summarize some of the previous work in the field of chemical data assimilation. The field has accumulated a large body of work from contributions by many authors. Among those, many excellent papers were products of the Global and Regional Earth-System Monitoring Using Satellite and In situ Data (GEMS [57]) project funded by European Commission to develop comprehensive data analysis and modelling systems for greenhouse gases, global reactive gases, and aerosol, with a focus on Europe from March 2005 to May 2009. Since June 2009, GEMS and the Protocol Monitoring for the GMES Service Element: Atmosphere (PROMOTE [58]) merged into Monitoring Atmospheric Composition and Climate (MACC [59]) project to continue the operation and improvement of the forecasting and

assimilation developed during GEMS [60]. Some excellent previous work will inevitably end up not being included in this paper's citation list, and we apologize for this.

Two approaches to data assimilation have become widely used in applications: variational methods, rooted in control theory, and Kalman filter methods, rooted in statistical estimation theory. The base concepts of the variational approach to chemical data assimilation are discussed in [1,10,27,29,61,62]. Early work in chemical data assimilation using variational techniques has been reported in [63,64]. A growing number of applications employ the 3D-Var technique [3,65–72]. The 4D-Var approach has been used to adjust gas phase chemical tracer initial conditions [10,11,61,69,73–77], to improve estimates of pollutant emissions, *i.e.*, emission inversion, [15,78,79], and to improve aerosol fields [80–82]. Suboptimal Kalman filters have been employed successfully in chemical data assimilation for over a decade [63,83–89]. More recently, the ensemble Kalman filter [90] has been studied in the context of chemical data assimilation [91–93]. Several studies compare the relative merits and performance of different approaches [93–98].

EnKF, extended Kalman filter [99] and reduced rank square root Kalman filter [100,101] have been used in chemical data assimilation to recover ozone [99], and various ways of accurately quantifying the uncertainty in sources have been investigated.

2. Data Assimilation Methods

The true state of the system (the true distribution of tracer concentrations in the atmosphere) is a continuous vector field \mathbf{c}^t distributed across three space dimensions and one time dimension. The number of components of the vector at a given location and a given moment equals the number of chemical species present in the atmosphere. The true state is unknown and needs to be estimated from the available information.

In practice we work with a finite dimensional representation of the continuous field $\mathbf{x}^t = \mathcal{S}(\mathbf{c}^t) \in \mathbb{R}^n$, and look to estimate \mathbf{x}^t from the available information. The operator \mathcal{S} maps the physical space to the model space (for example, it can sample the continuous field at the grid points, or it can lump several chemical species into a single representative family, then average the family concentration over each grid cell, *etc.*)

In order to obtain an estimate of \mathbf{x}^t data assimilation combines three different sources of information: the prior information, the model, and the observations. The best estimate that optimally fuses all these sources of information is called the analysis, and is denoted by $\mathbf{x}^a \in \mathbb{R}^n$.

The prior information. The background (prior) probability density $\mathcal{P}^b(\mathbf{x})$ encapsulates our current knowledge of the tracer distribution. Specifically, $\mathcal{P}^b(\mathbf{x})$ describes the uncertainty with which one knows \mathbf{x}^t at the present, before any (new) measurements are taken. The mean taken with respect to this probability density is denoted by

$$\mathbb{E}^b[f] = \int f(\mathbf{x}) \mathcal{P}^b(\mathbf{x}) d\mathbf{x}$$

The current best estimate of the true state is called the apriori, or the background state $\mathbf{x}^b \in \mathbb{R}^n$. (This is often taken to be the mean of the background distribution $\mathbf{x}^b = \mathbb{E}^b[\mathbf{x}]$.) A typical assumption is that the random background errors $\varepsilon^b = \mathbf{x}^b - \mathbf{x}^t$ are unbiased and have a normal probability density, *i.e.*

$$\varepsilon^b = \mathbf{x}^b - \mathbf{x}^t \in \mathcal{N}(\mathbf{0}, \mathbf{B}) \quad (3)$$

Here $\mathbf{B} = \mathbb{E}^b [\varepsilon^b (\varepsilon^b)^T] \in \mathbb{R}^{n \times n}$ is the *background error covariance matrix*. With many nonlinear models (e.g., in the presence of nonlinear chemical kinetics) the normality assumption (3) might not always be valid. Nevertheless, it is widely used because of its convenience.

The model. The model (1) encapsulates our knowledge about physical and chemical laws that govern the evolution of the atmospheric composition. The model evolves an initial state $\mathbf{x}_0 \in \mathbb{R}^n$ at the initial time t_0 to future state values $\mathbf{x}_i \in \mathbb{R}^n$ at future times t_i

$$\mathbf{x}_i = \mathcal{M}_{t_0 \rightarrow t_i}(\mathbf{x}_0) \quad (4)$$

The size of the state space in realistic chemical transport models is very large, typically $n \in \mathcal{O}(10^7)$ variables for regional models and $n \in \mathcal{O}(10^8)$ for global models. The model is a first order Markov process, meaning that the probability distribution of the state at time t_i depends only on the probability as time t_{i-1} : $\mathcal{P}(\mathbf{x}_i | [\mathbf{x}_0, \dots, \mathbf{x}_{i-1}]) = \mathcal{P}(\mathbf{x}_i | \mathbf{x}_{i-1})$.

The observations. Observations represent snapshots of reality available at several discrete time moments. Specifically, measurements $\mathbf{y}_i \in \mathbb{R}^m$ of the physical state are taken at times $t_i, i = 1, \dots, N$

$$\mathbf{y}_i = \mathcal{H}^t(\mathbf{c}^t(t_i)) - \varepsilon_i^{\text{meas}}, \quad i = 1, \dots, N \quad (5)$$

The observation operator \mathcal{H}^t maps the physical state space onto the observation space. The measurement (instrument) errors are denoted by $\varepsilon_i^{\text{meas}}$.

In order to relate the model state to observations we also consider the relation

$$\mathbf{y}_i = \mathcal{H}(\mathbf{x}_i^t) - \varepsilon_i^{\text{obs}}, \quad i = 1, \dots, N \quad (6)$$

where the observation operator \mathcal{H} maps the model state space onto the observation space. In many practical situations \mathcal{H} is a highly nonlinear mapping (as is the case, e.g., with satellite observation operators). At present the chemical observations are sparsely distributed, and their number is small compared to the dimension of the state space, $m \ll n$.

The *observation error* term $\varepsilon_i^{\text{obs}}$ accounts for both the measurement errors $\varepsilon_i^{\text{meas}}$, as well as the representativeness errors $\varepsilon_i^{\text{repres}}$ (*i.e.*, errors in the accuracy with which the model can reproduce reality, and with which the numerical operator \mathcal{H} approximates \mathcal{H}^t)

$$\varepsilon_i^{\text{repres}} = \mathcal{H}(\mathbf{x}_i^t) - \mathcal{H}^t(\mathbf{c}^t(t_i)) = \mathcal{H}(\mathcal{S}(\mathbf{c}^t(t_i))) - \mathcal{H}^t(\mathbf{c}^t(t_i))$$

Typically observation errors are assumed to be unbiased and normally distributed

$$\varepsilon_i^{\text{obs}} \in \mathcal{N}(\mathbf{0}, \mathbf{R}_i), \quad i = 1, \dots, N \quad (7)$$

Observation errors at different times ($\varepsilon_i^{\text{obs}}$ and $\varepsilon_j^{\text{obs}}$ for $i \neq j$) are assumed to be independent. Often, the observation errors are also assumed to be spatially uncorrelated. In matrix form this is equivalent to assume that the observation error covariance matrix is diagonal. Moreover, observation errors and background errors are assumed independent of each other.

The analysis. Based on these three sources of information data assimilation computes the analysis (posterior) probability density $\mathcal{P}^a(\mathbf{x})$. Specifically, $\mathcal{P}^a(\mathbf{x})$ describes the uncertainty with which one knows \mathbf{x}^t after all the information available from measurements has been accounted for. The mean taken with respect to this probability density is denoted by $\mathbb{E}^a[f] = \int f(\mathbf{x}) \mathcal{P}^a(\mathbf{x}) d\mathbf{x}$.

The best estimate \mathbf{x}^a of the true state obtained from analysis distribution is called the *aposteriori*, or the *analysis state*. (This estimate can be the posterior mean $\mathbf{x}^a = \mathbb{E}^a[\mathbf{x}]$, but this is not necessary; in the maximum likelihood approach the refined estimate of the true state is obtained from the analysis distribution mode). The analysis estimation errors $\varepsilon^a = \mathbf{x}^a - \mathbf{x}^t$ are characterized by the *analysis mean error (bias)* $\beta^a = \mathbb{E}^a[\varepsilon^a]$ and by the *analysis error covariance matrix* $\mathbf{A} = \mathbb{E}^a[(\varepsilon^a - \beta^a)(\varepsilon^a - \beta^a)^T] \in \mathbb{R}^{n \times n}$. By design, the analysis errors are also normally distributed if the background and observation errors are assumed such.

2.1. The Bayesian Estimation Framework

The chemical data assimilation problem is formulated in a Bayesian framework. The analysis probability density is the probability density of the state *conditioned by all the available observations* $\mathbf{y} = [\mathbf{y}_1, \dots, \mathbf{y}_N]$. Bayes Theorem allows one to express the analysis probability density as follows:

$$\mathcal{P}^a(\mathbf{x}) = \mathcal{P}(\mathbf{x}|\mathbf{y}) = \frac{\mathcal{P}(\mathbf{y}|\mathbf{x}) \cdot \mathcal{P}^b(\mathbf{x})}{\mathcal{P}(\mathbf{y})} \quad (8)$$

The denominator $\mathcal{P}(\mathbf{y})$ is the marginal probability density of the observations and plays the role of a scaling factor. The probability of the observations conditioned by the states $\mathcal{P}(\mathbf{y}|\mathbf{x})$ is the probability that the observation errors in (6) assume the values $\mathcal{H}(\mathbf{x}^b) - \mathbf{y}$

$$\mathcal{P}(\mathbf{y}|\mathbf{x}) = \mathcal{P}^{\text{obs}}(\varepsilon^{\text{obs}} = \mathcal{H}(\mathbf{x}^b) - \mathbf{y})$$

Since the observation errors $\varepsilon_1^{\text{obs}}, \dots, \varepsilon_N^{\text{obs}}$ at different times t_1, \dots, t_N are (considered to be) independent, we have that:

$$\mathcal{P}(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^N \mathcal{P}^{\text{obs}}(\varepsilon_i^{\text{obs}}) = \prod_{i=1}^N \mathcal{P}^{\text{obs}}(\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i) \quad (9)$$

2.2. Bayesian Estimators

Bayes' theorem (8) completely describes the posterior error distribution. In large scale models a direct application of (8) is not possible, since it involves multidimensional probability densities defined over very large spaces (recall that $n \sim 10^7$). Approximations are needed in order to represent such densities. One approach is to approximate all probabilities involved by normal distributions, in which case closed form solutions for the posterior density are possible, see Section 2.3. Practical algorithms based on normal approximations are the suboptimal Kalman filters, discussed in Section 3.2. Another possible approximation is the Monte Carlo approach, where all the probability densities involved are represented by samples in the state space. In this case the application of Bayes' theorem (8) results in a random sample from the posterior distribution. Practical algorithms based on the Monte Carlo approach include ensemble Kalman filters (discussed in Section 3.4) and particle filters [102]. Finally, a less ambitious goal is to obtain only the first several moments of the posterior probability density based on (8).

In practice we want to use (8) to define estimators \mathbf{x}^a of the true state \mathbf{x}^t that are optimal in a certain sense. One way to define a best estimator is to minimize the expected values of the mean square error $\min \mathbb{E}^a[\|\mathbf{x}^a - \mathbf{x}^t\|^2]$. The resulting minimum mean square error (MMSE) estimator is given by the mean of the posterior distribution, $\mathbf{x}^a = \mathbb{E}^a[\mathbf{x}]$. This estimator is not practical for large scale systems, as it requires an integration in the high dimensional state space. Practical estimators are obtained by taking the mean of an approximation of the posterior distribution, see for example Section 3.4. A computationally feasible estimator is given by the mode of the posterior distribution, and is called the maximum a posteriori estimator (MAP), as discussed in Section 2.4. Of particular interest are unbiased estimators, which are characterized by a zero posterior error mean (*i.e.*, zero bias, $\beta^a = \mathbf{0}$). A minimum variance unbiased (MVUE) estimator \mathbf{x}^a has the smallest total variance ($\min \text{trace } \mathbb{E}^a[(\mathbf{x}^a - \mathbb{E}^a[\mathbf{x}^a])(\mathbf{x}^a - \mathbb{E}^a[\mathbf{x}^a])^T]$) among all unbiased estimators. MVUE estimators are not guaranteed to exist, and when they do, they are difficult to compute for practical problems.

2.3. Analytical Solution in the Gaussian and Linear Case

Consider a time invariant ideal case where the observation operator is linear

$$\mathcal{H}(\mathbf{x}) = \mathbf{H} \cdot \mathbf{x}, \quad \mathbf{H} \in \mathbb{R}^{m \times n} \quad (10)$$

and both the background errors (3) and the observation errors (7) are normally distributed

$$\mathcal{P}^b(\mathbf{x}) = (2\pi)^{-n/2} (\det \mathbf{B})^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b)\right) \quad (11a)$$

$$\mathcal{P}^{\text{obs}}(\mathbf{y}|\mathbf{x}) = (2\pi)^{-m/2} (\det \mathbf{R})^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{y})\right) \quad (11b)$$

After inserting (11a) and (11b) in (8) a direct calculation shows that the posterior probability density is also Gaussian, $\mathcal{P}^a(\mathbf{x}) = \mathcal{N}(\mathbf{x}^a, \mathbf{A})$

$$\mathcal{P}^a(\mathbf{x}) = (2\pi)^{-n/2} (\det \mathbf{A})^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}^a)^T \mathbf{A}^{-1}(\mathbf{x} - \mathbf{x}^a)\right) \quad (11c)$$

with the analysis mean \mathbf{x}^a and covariance \mathbf{A} given by the Kalman filter [103] formulas:

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \quad (12a)$$

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K} (\mathbf{y} - \mathbf{H}\mathbf{x}^b) \quad (12b)$$

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H}) \mathbf{B} \quad (12c)$$

where \mathbf{I} is the identity matrix. The matrix $\mathbf{K} \in \mathbb{R}^{n \times m}$ is called the “Kalman gain” operator. \mathbf{A} is the covariance matrix of analysis error. Note that in the linear Gaussian case the estimate (12b) represents both the MMSE estimator and the MAP estimator. In general, however, the MMSE and the MAP estimates are distinct.

2.4. Maximum A posteriori Estimator

In the maximum likelihood approach one looks for the argument that maximizes the posterior distribution, or equivalently, minimizes its negative logarithm:

$$\mathbf{x}^a = \arg \max_{\mathbf{x}} \mathcal{P}^a(\mathbf{x}) = \arg \min_{\mathbf{x}} \mathcal{J}(\mathbf{x}), \quad \mathcal{J}(\mathbf{x}) = -\ln \mathcal{P}^a(\mathbf{x}) \quad (13)$$

Equation (13) defines the maximum a posteriori estimator (MAP). In this context the data assimilation problem is formulated as an optimization problem. Using (8) the minimization cost function can be written as

$$\mathcal{J}(\mathbf{x}) = -\ln \mathcal{P}^a(\mathbf{x}) = -\ln \mathcal{P}^b(\mathbf{x}) - \ln \mathcal{P}(\mathbf{y}|\mathbf{x}) + \text{const} \quad (14)$$

The scaling factors of the probability densities, as well as the term $-\ln \mathcal{P}(\mathbf{y})$, are constants in \mathbf{x} and do not influence the result of the minimization. Under the assumption that the background errors are normally distributed (11a) we have that

$$-\ln \mathcal{P}^b(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) + \text{const} \quad (15)$$

Similarly, under the assumption that observation errors are independent (9) and normally distributed (11b) we have that

$$-\ln \mathcal{P}(\mathbf{y}|\mathbf{x}) = -\ln \mathcal{P}^{\text{obs}}(\varepsilon^{\text{obs}}) = \frac{1}{2} (\mathcal{H}(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1} (\mathcal{H}(\mathbf{x}) - \mathbf{y}) + \text{const} \quad (16)$$

The maximum likelihood estimator is obtained as the minimizer of the cost function

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) + \frac{1}{2} (\mathcal{H}(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1} (\mathcal{H}(\mathbf{x}) - \mathbf{y}) \quad (17)$$

where the constant terms have been left out.

Note that if, in addition, the observation operator is linear (10) then the function (17) is quadratic, and the minimizer can be computed explicitly from setting the gradient to zero

$$\nabla_{\mathbf{x}} \mathcal{J}(\mathbf{x}^a) = \mathbf{B}^{-1} (\mathbf{x}^a - \mathbf{x}^b) + \mathbf{H}^T \mathbf{R}^{-1} w (\mathbf{H} \mathbf{x}^a - \mathbf{y}) = \mathbf{0} \quad (18)$$

The result is the Kalman filter estimate for the mean (12b). Moreover, the Hessian of the cost function coincides with the inverse of the Kalman filter analysis covariance matrix (12c)

$$\nabla_{\mathbf{x}, \mathbf{x}}^2 \mathcal{J} = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} = \mathbf{A}^{-1} \quad (19)$$

2.5. Time Dependent Systems

Typical data assimilation applications are concerned with time dependent systems, e.g., the evolution of the chemical composition of the atmosphere. In such applications the interest is not focused on one analysis at one time, but on a series of analyses for times t_1, \dots, t_N when observations are available.

There are two approaches to obtain the analysis probability densities $\mathcal{P}^a(\mathbf{x}_i)$. In the *smoothing* (*simultaneous*) data assimilation approach all observations at all times t_1, \dots, t_N are considered at once. Corrections of the concentration state vectors at all times are determined in the same analysis step. The result is a sequence of posterior probabilities of states, $\mathcal{P}(\mathbf{x}_i | [\mathbf{y}_1, \dots, \mathbf{y}_N])$, $i = 1, \dots, N$, each conditioned by all available observations. The application of (12) in the simultaneous setting leads to the Kalman smoother approach, while the maximum likelihood estimator obtained from (17) leads to the four dimensional variational (4D-Var) assimilation method.

In the *filtering* (*sequential*) data assimilation approach [63] the observations (6) are considered successively at times t_1, \dots, t_N . Corrections of the concentration state vector are computed and applied

at each t_i as soon as observations become available. The result is a sequence of posterior probabilities of states $\mathcal{P}(\mathbf{x}_i | [\mathbf{y}_1, \dots, \mathbf{y}_i])$, $i = 1, \dots, N$, each conditioned by all past and current observations (but not by the future observations). The application of (12) in the sequential setting leads to the Kalman filter approach, while the maximum likelihood estimator obtained from (17) leads to the three dimensional variational (3D-Var) assimilation method.

We now discuss the Kalman filter approach in the ideal case where the observation operator is linear (10), and, in addition, the model dynamics (4) is also linear, $\mathcal{M}_{t_{i-1} \rightarrow t_i}(\mathbf{x}) = \mathbf{M}_{t_{i-1} \rightarrow t_i} \cdot \mathbf{x}$.

The background state (*i.e.*, the best state estimate) at time t_i is given by the model forecast, starting from the analysis (*i.e.*, the best estimate at the previous time t_{i-1}):

$$\mathbf{x}_i^b \equiv \mathbf{x}_i^f = \mathbf{M}_{t_{i-1} \rightarrow t_i} \cdot \mathbf{x}_{i-1}^a \quad (20a)$$

Note that a model forecast starting from the true state at t_{i-1} does not reproduce the true state at t_i since the model only approximates the dynamics of the physical system. Specifically, we have that

$$\mathbf{x}_i^t = \mathbf{M}_{t_{i-1} \rightarrow t_i} \cdot \mathbf{x}_{i-1}^t - \eta_i$$

where η_i is the model error. Typically the model error is assumed to be a normal random variable $\eta_i \in \mathcal{N}(\mathbf{0}, \mathbf{Q}_i)$, where the zero mean represents the unbiased model assumption.

The background error at t_i has two components: the analysis error at t_{i-1} , transported through the model equations, and the model error

$$\varepsilon_i^b = \mathbf{M}_{t_{i-1} \rightarrow t_i} \cdot \varepsilon_{i-1}^a - \eta_i$$

The model error η_i and the solution error ε_{i-1}^a are typically assumed to be independent. Consequently, the background error covariance at t_i (the forecast error covariance matrix \mathbf{P}_i^f) is obtained by transporting the analysis covariance at t_{i-1} to t_i through the linearized dynamics, and adding the model error covariance

$$\mathbf{B}_i \equiv \mathbf{P}_i^f = \mathbf{M}_{t_{i-1} \rightarrow t_i} \mathbf{P}_{i-1}^a \mathbf{M}_{t_{i-1} \rightarrow t_i}^T + \mathbf{Q}_i \quad (20b)$$

For every observation time t_i , the filter starts with the model forecast state (\mathbf{x}_i^f) and provides an analysis state (\mathbf{x}_i^a) that reduces the discrepancy between the model forecast and the observations \mathbf{y}_i . The analysis state vector is obtained from (12b)

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}_i (\mathbf{y}_i - \mathbf{H} \mathbf{x}_i^f) \quad (20c)$$

with the Kalman gain matrix given by (12a)

$$\mathbf{K}_i = \mathbf{P}_i^f \mathbf{H}_i^T (\mathbf{H}_i \mathbf{P}_i^f \mathbf{H}_i^T + \mathbf{R}_i)^{-1} \quad (20d)$$

where \mathbf{R}_i is the observation error covariance matrix at time t_i . At each observation time, along with the analysis state, the analysis error covariance matrix \mathbf{P}_i^a is also calculated via (12c)

$$\mathbf{A}_i \equiv \mathbf{P}_i^a = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i) \mathbf{P}_i^f \quad (20e)$$

3. Practical Algorithms for Chemical Data Assimilation

Practical data assimilation algorithms use the estimation approaches presented in Section 2.2, together with various approximations most often related to Gaussian assumptions and to the structure of the underlying physical model.

3.1. The Extended Kalman Filter

The extended Kalman filter (EKF) generalizes the original Equations (20) to nonlinear systems (4) and nonlinear observations (6) by linearization about the forecast state. Consider the linearized model and observation operators

$$\mathbf{M}_{t_{i-1} \rightarrow t_i} = \mathcal{M}'(\mathbf{x}^f), \quad \mathbf{H}_i = \mathcal{H}'(\mathbf{x}_i^f)$$

The EKF approach modifies (20) as follows. The forecast state equation uses the nonlinear model \mathcal{M} , while the forecast covariance equation uses the linearized dynamics \mathbf{M} . Similarly, the analysis equation uses the nonlinear observation operator \mathcal{H} , but both the gain equation and the analysis covariance equation use the linearized operator \mathbf{H} . The resulting EKF equations are:

$$\mathbf{x}_i^f = \mathcal{M}_{t_{i-1} \rightarrow t_i}(\mathbf{x}_{i-1}^a) \quad (21a)$$

$$\mathbf{P}_i^f = \mathbf{M}_{t_{i-1} \rightarrow t_i} \mathbf{P}_{i-1}^a \mathbf{M}_{t_{i-1} \rightarrow t_i}^T + \mathbf{Q}_i \quad (21b)$$

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}_i (\mathbf{y}_i - \mathcal{H}(\mathbf{x}_i^f)) \quad (21c)$$

$$\mathbf{K}_i = \mathbf{P}_i^f \mathbf{H}_i^T (\mathbf{H}_i \mathbf{P}_i^f \mathbf{H}_i^T + \mathbf{R}_i)^{-1} \quad (21d)$$

$$\mathbf{P}_i^a = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i) \mathbf{P}_i^f \quad (21e)$$

3.2. Suboptimal Kalman Filters

The extended Kalman filter is not practical for large systems because of the $\mathcal{O}(n^2)$ memory size needed to store full covariance matrices, and the prohibitive computational costs associated with inverting large matrices in (21c)–(21d), and with propagating the covariance matrices in time via (21e). *Suboptimal Kalman filters* designate a wide class of assimilation algorithms which are based on EKF formulas (21), but approximate the covariance matrices as well as the covariance propagation Equation (21e) in order to obtain computationally feasible algorithms. The approximations lead to suboptimal solutions, even in the case of linear Gaussian systems. There are multiple ways in which this analysis covariance matrix is made available to the next observation window, and different approximation strategies lead to different suboptimal filters.

A low memory approximation of a covariance matrix \mathbf{B} can store only the diagonal terms (the variances $\mathbf{B}_{(\ell),(\ell)} = \sigma_{(\ell)}^2$ of the error in state variables $\mathbf{x}_{(\ell)}$ for $\ell = 1, \dots, n$), and use a model to represent the error correlation structure. For example, the correlation between the errors in $\mathbf{x}_{(\ell)}$ and $\mathbf{x}_{(k)}$ can be modeled as decreasing with the distance between the gridpoints of ℓ and k . When a Gaussian de-correlation formula is used, with a correlation distance of L (space units), the $\{(\ell), (k)\}$ entry of the approximate covariance matrix is

$$\mathbf{B}_{(\ell), (k)} = \sigma_{(\ell)} \sigma_{(k)} \exp(-\text{distance}\{\text{gridpoint}(\ell), \text{gridpoint}(k)\}^2 / L^2) \quad (22)$$

Polynomial models of spatial correlations [104] are also widely used.

The simplest approach to avoid the cost of (21e) is to keep the forecast covariance equal to the background covariance for the entire assimilation period, $\mathbf{P}_i^f = \mathbf{B}_0$ for $i = 1, \dots, N$ [89]. A more complex approach is to build diagonal approximations to \mathbf{P}_i^f by transporting the standard deviations $\sigma_{(\ell)}$ as passive tracers from t_{i-1} to t_i [83]. The propagated variances can be used together with a model of the correlation structure to reconstruct \mathbf{P}_i^f .

The *reduced rank Kalman filter* approach [105] is based on the observation that the symmetric positive definite matrix \mathbf{B} can be completely described in terms of its eigenvalues λ_i and its orthonormal eigenvectors v_i . A rank r approximation of the matrix can be constructed from the dominant eigenvalue-eigenvector pairs as follows:

$$\mathbf{B} = \sum_{i=1}^n \lambda_i v_i v_i^T \approx \sum_{i=1}^r \lambda_i v_i v_i^T = \mathbf{V} \mathbf{V}^T, \quad \mathbf{V} = [\sqrt{\lambda_1} v_1, \dots, \sqrt{\lambda_r} v_r] \in \mathbb{R}^{n \times r} \quad (23)$$

Using a rank r approximation for the analysis covariance matrix at t_{i-1}

$$\mathbf{P}_{i-1}^a = \mathbf{V}_{i-1}^a (\mathbf{V}_{i-1}^a)^T$$

leads to the following forecast covariance (21b)

$$\mathbf{P}_i^f = (\mathbf{M}_{t_{i-1} \rightarrow t_i} \mathbf{V}_{i-1}^a) (\mathbf{M}_{t_{i-1} \rightarrow t_i} \mathbf{V}_{i-1}^a)^T + \mathbf{Q}_i \approx \mathbf{V}_i^f (\mathbf{V}_i^f)^T \quad (24)$$

The terms $\mathbf{M}_{t_{i-1} \rightarrow t_i} \mathbf{V}_{i-1}^a$ are evaluated by propagating the r vectors through the linearized model dynamics. A rank r approximation of the forecast covariance is obtained via (23). Using this approximation, the Kalman gain matrix (21d) becomes

$$\mathbf{W}_i^f = \mathbf{H} \cdot \mathbf{V}_i^f, \quad \mathbf{K}_i = \mathbf{V}_i^f (\mathbf{W}_i^f)^T (\mathbf{W}_i^f \cdot (\mathbf{W}_i^f)^T + \mathbf{R}_i)^{-1}$$

and the analysis covariance (21e) reads

$$\mathbf{P}_i^a = (\mathbf{V}_i^f - \mathbf{K}_i \mathbf{W}_i^f) (\mathbf{V}_i^f)^T = \mathbf{V}_i^a (\mathbf{V}_i^a)^T, \quad \mathbf{V}_i^a = \mathbf{V}_i^f \left(\mathbf{I} - (\mathbf{W}_i^f)^T (\mathbf{W}_i^f \cdot (\mathbf{W}_i^f)^T + \mathbf{R}_i)^{-1} \mathbf{W}_i^f \right)^{1/2}$$

It is immediate that the analysis increments $\mathbf{x}_i^a - \mathbf{x}_i^f$ in (21c) are restricted to the r -dimensional subspace spanned by the columns of \mathbf{V}_i^f (the so-called “rank problem”). In particular, the r degrees of freedom available to the analysis may be insufficient to produce a good fit to observations. One way to overcome this problem is to perform local analyses, as discussed in Section 3.3. Another way is through covariance localization [106], where an assumed correlation structure is superimposed to the low rank approximation. For example, using (22), the $\{(\ell), (k)\}$ entry of the forecast covariance matrix (24) becomes

$$(\mathbf{P}_i^f)_{(\ell), (k)} = (\mathbf{V}_i^f)_{(\ell)} (\mathbf{V}_i^f)_{(k)} \exp(-\text{distance}\{\text{gridpoint}(\ell), \text{gridpoint}(k)\}^2 / L^2)$$

Localization improves the accuracy of the approximation by removing spurious long-distance correlations, and results in a full rank forecast covariance matrix.

3.3. Optimal Interpolation

Optimal interpolation [107] simplifies the extended Kalman filter formulation (21) by assuming that, during the analysis process, each model variable is influenced by only a subset of observations. Consider, without loss of generality, that only $\mu \ll m$ observations have an impact on the model variable $\mathbf{x}_{(\ell)}$. For example, these can be observations located sufficiently close to the grid point where $\mathbf{x}_{(\ell)}$ is defined. Let $\mathbf{I}_\ell \in \mathbb{R}^{\mu \times m}$ be the operator that selects the important μ components out of the m -dimensional vector of observations, $\mathbf{y}_\ell = \mathbf{I}_\ell \mathbf{y} \in \mathbb{R}^\mu$. Then $\mathbf{H}_\ell = \mathbf{I}_\ell \mathbf{H} \in \mathbb{R}^{\mu \times n}$ is the observation operator associated with the locally important observations, and $\mathbf{R}_\ell = \mathbf{I}_\ell \mathbf{R} \mathbf{I}_\ell^T \in \mathbb{R}^{\mu \times \mu}$ is the corresponding observation error covariance matrix.

Let $\mathbf{e}_\ell \in \mathbb{R}^n$ be the ℓ -th column of the identity matrix. From (12a) and (12b) the analysis of variable $\mathbf{x}_{(\ell)}$ is given by

$$\begin{aligned} \mathbf{x}_{(\ell)}^a &= \mathbf{x}_{(\ell)}^b + \mathbf{e}_\ell^T \mathbf{B} \mathbf{H}^T (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{y} - \mathbf{H} \mathbf{x}^b) \\ &\approx \mathbf{x}_{(\ell)}^b + \mathbf{e}_\ell^T \mathbf{B} \mathbf{H}_\ell^T (\mathbf{H}_\ell \mathbf{B} \mathbf{H}_\ell^T + \mathbf{R}_\ell)^{-1} (\mathbf{y}_\ell - \mathbf{H}_\ell \mathbf{x}^b) \end{aligned}$$

The cost of forming and solving the matrix $\mathbf{H}_\ell \mathbf{B} \mathbf{H}_\ell^T + \mathbf{R}_\ell$ is $\mathcal{O}(\mu^3)$, instead of $\mathcal{O}(m^3)$ for the complete matrix $\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R}$. Only the increments $\mathbf{y}_\ell - \mathbf{H}_\ell \mathbf{x}^b$ of the important observations are used. The weight $\mathbf{e}_\ell^T \mathbf{B} \mathbf{H}_\ell^T$ is a row vector obtained by applying the relevant part of the observation operator to the ℓ -th column of the background covariance $\mathbf{H}_\ell (\mathbf{B} \mathbf{e}_\ell)$, and transposing the result. The analyses for different components ℓ can be computed in parallel.

When approximations of \mathbf{B} are employed this is easy to compute. For example, using the approximation (22), the weight vector reads

$$(\mathbf{H}_\ell \mathbf{B} \mathbf{e}_\ell)_{(j)} = \sigma_{(\ell)} \sum_k (\mathbf{H}_\ell)_{(j),(k)} \sigma_{(k)} \exp(-\text{distance}\{\text{gridpoint}(\ell), \text{gridpoint}(k)\}^2 / L^2)$$

3.4. Ensemble Kalman Filters

The ensemble Kalman filter (EnKF) [90,105,108] uses a Monte-Carlo approach to propagate covariances. An ensemble of E states (labeled $e = 1, \dots, E$) is used to sample the probability distribution of the error. The analysis probability density at time t_{i-1} is represented by the sample points $\mathbf{x}_{i-1}^a[e]$, $e = 1, \dots, E$, in the state space. Each member of the ensemble is propagated to t_i using the model (4) to obtain the “forecast” ensemble

$$\mathbf{x}_i^f[e] = \mathcal{M}_{t_{i-1} \rightarrow t_i} \mathbf{x}_{i-1}^a[e] + \eta_i[e], \quad e = 1, \dots, E \quad (25)$$

where the random variable η_i represents the *model error*, and is typically assumed to be Gaussian and unbiased, $\eta_i \in \mathcal{N}(\mathbf{0}, \mathbf{Q}_i)$. The forecast error covariance \mathbf{P}_i^f is estimated from the statistical samples

$$\langle \mathbf{x}_i^f \rangle = \frac{1}{E} \sum_{e=1}^E \mathbf{x}_i^f[e], \quad \mathbf{P}_i^f \approx \frac{1}{E-1} \sum_{e=1}^E (\mathbf{x}_i^f[e] - \langle \mathbf{x}_i^f \rangle) (\mathbf{x}_i^f[e] - \langle \mathbf{x}_i^f \rangle)^T \quad (26)$$

and the Kalman gain matrix is computed using Equation (20d).

Each member forecast ensemble is processed separately using (20c) to obtain the “analysis” ensemble

$$\mathbf{x}_i^a[e] = \mathbf{x}_i^f[e] + \mathbf{K}_i (\mathbf{y}_i[e] - \mathcal{H}_i(\mathbf{x}_i^f[e])), \quad e = 1, \dots, E \quad (27)$$

To obtain the correct posterior statistics, a different set of perturbed observations is used for each ensemble member, $\mathbf{y}_i[e] = \mathbf{y}_i + \theta_i[e]$, with perturbations drawn from the real observation error statistics $\theta_i[e] \in \mathcal{N}(\mathbf{0}, \mathbf{R}_i)$ [90,108]. The analysis covariance is estimated from the statistical samples $\mathbf{x}_i^a[e]$, $e = 1, \dots, E$, using the formula (26).

The ensemble Kalman filter raises several issues. First the rank of the estimated covariance matrix (26) is typically several orders of magnitude smaller than the dimension of the matrix, and additional approximations are needed to fix the rank-deficiency problem [106]. Next, the random errors in the statistically estimated covariance decrease slowly, with only the square-root of the ensemble size E . Furthermore, the subspace spanned by random vectors for expressing the forecast error is not optimal.

In spite of the problems, ensemble Kalman filter has many attractive features. The effects of non-linear dynamics are captured by the use of the forward model (25). This model is used as is, and there is no need for the tangent linear or adjoint models. EnKF allows one to easily account for model errors, and the calculations are almost ideally parallelizable.

Numerous improvements of the original EnKF [90,109] have been proposed in the literature to alleviate inbreeding [110], to increase computational efficiency [105,106,111], to relax the normal error distribution assumptions [112,113], and to allow observations to occur at times different than assimilation times [114,115]. The square-root implementations of EnKF [116,117] update the ensemble by applying linear transformations to the prior ensemble, and avoid adding perturbations to observations (e.g., the ensemble adjustment [118], the variance reduced [119], and the ensemble transform [120] Kalman filters).

The use of EnKF [90] in chemical data assimilation has been studied in [91–93,121–124]. Three techniques have proved essential for the practical performance of the EnKF. Due to the small ensemble size many entries in the forecast covariance matrix are poorly approximated; such sampling errors are referred to as spurious correlations. Covariance localization scales each entry $\mathbf{P}_{(k,\ell)}^f$ by a function that decreases with the physical distance between the gridpoints where $\mathbf{x}_{(\ell)}$ and $\mathbf{x}_{(k)}$ are defined in Equation (22). Covariance localization alleviates the effect of spurious correlations, and improves the rank of \mathbf{P}^f . It has been observed in practice that, after a number of assimilation cycles, all ensemble members tend to be close to one another in the state space. In this case the estimated forecast covariance (26) is small, and the filter trusts the model too much and starts rejecting the observations. This situation is referred to as filter divergence. Covariance inflation scales \mathbf{P}^f by a factor $\alpha > 1$ at each cycle. The scaling has the net effect of accounting for larger model errors, and helps prevent filter divergence. It has also been observed in practice that the inflation goes uncorrected in data-sparse regions, and the ensemble spread continues to grow to unreasonable values. To alleviate this, the third important technique is adaptive inflation (inflation is localized to data-rich areas) [93].

3.5. Three Dimensional Variational Data Assimilation (3D-Var)

Variational methods solve the data assimilation problem in an optimal control framework [125–127]. Specifically, one finds the control variable values (e.g., initial conditions) which minimize the discrepancy between model forecast and observations; the minimization is subject to the governing equations, which are imposed as strong constraints in most practical applications. Similar as OI, 3D-Var does not consider evolution of the model in the assimilation. Thus, it is possible

to have a dual formulation of OI/3D-Var [128]. In OI applications, analysis is often solved in blocks due to the computation difficulties of the large size matrix inversion problems. Complicated observation operators are often obstacles to use OI in practice. In this discussion, for simplicity of presentation, we focus on discrete models where the initial conditions are the control variables.

In the 3D-Var data assimilation the observations (6) are considered successively at times t_1, \dots, t_N . The background state (*i.e.*, the best state estimate at time t_i) is given by the model forecast, starting from the previous analysis (*i.e.*, best estimate at time t_{i-1})

$$\mathbf{x}_i^b = \mathcal{M}_{t_{i-1} \rightarrow t_i} (\mathbf{x}_{i-1}^a)$$

The discrepancy between the model state \mathbf{x}_i and observations at time t_i , together with the departure of the state from the model forecast \mathbf{x}_i^b , are measured by the 3D-Var cost function (17):

$$\mathcal{J}(\mathbf{x}_i) = \frac{1}{2} (\mathbf{x}_i - \mathbf{x}_i^b)^T \mathbf{B}_i^{-1} (\mathbf{x}_i - \mathbf{x}_i^b) + \frac{1}{2} (\mathcal{H}(\mathbf{x}_i) - \mathbf{x}_i^{\text{obs}})^T \mathbf{R}_i^{-1} (\mathcal{H}(\mathbf{x}_i) - \mathbf{x}_i^{\text{obs}}) \quad (28)$$

While in principle a different background covariance matrix should be used at each time, in practice the same matrix is re-used throughout the assimilation window, $\mathbf{B}_i = \mathbf{B}$, $i = 1, \dots, N$. The 3D-Var analysis is the MAP estimator, and is computed as the state which minimizes (28)

$$\mathbf{x}_i^a = \arg \min \mathcal{J}(\mathbf{x}_i) \quad (29)$$

Typically a gradient-based numerical optimization procedure is employed to solve (29). The gradient $\nabla \mathcal{J}$ of the cost function (28) is

$$\nabla \mathcal{J}(\mathbf{x}_i) = \mathbf{B}_i^{-1} (\mathbf{x}_i - \mathbf{x}_i^b) + \mathbf{H}_i^T \mathbf{R}_i^{-1} (\mathcal{H}(\mathbf{x}_i) - \mathbf{x}_i^{\text{obs}}) \quad (30)$$

Note that the gradient requires to computation of the adjoint \mathbf{H}_i^T of the linearized observation operator $\mathbf{H}_i = \mathcal{H}'(\mathbf{x}_i)$ about the current state.

Preconditioning is often used to improve convergence of the numerical optimization problem (29). A change of variables is performed by shifting the state and scaling it with the square root of covariance:

$$\hat{\mathbf{x}}_i = \mathbf{B}_i^{1/2} (\mathbf{x}_i - \mathbf{x}_i^b) \quad (31)$$

and carrying out the optimization with the new variables $\hat{\mathbf{x}}_i$.

3.6. Four Dimensional Variational Data Assimilation (4D-Var)

In strongly-constrained 4D-Var data assimilation all observations (6) at all times t_1, \dots, t_N are considered simultaneously over the assimilation window. The control parameters are the initial conditions \mathbf{x}_0 ; they uniquely determine the state of the system at all future times via the model Equation (4). The background state is the prior value of the initial conditions \mathbf{x}_0^b .

Given the background value of the initial state \mathbf{x}_0^b , the covariance of the initial background errors \mathbf{B}_0 , the observations \mathbf{y}_i at t_i and the corresponding observation error covariances \mathbf{R}_i , $i = 1, \dots, N$, the 4D-Var problem looks for the MAP estimate \mathbf{x}_0^a of the true initial conditions by solving the optimization problem (13). Combining (14), (15), and (16) leads to the 4D-var cost function:

$$\mathcal{J}(\mathbf{x}_0) = \frac{1}{2} (\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2} \sum_{i=1}^N (\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i)^T \mathbf{R}_i^{-1} (\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i) \quad (32)$$

Note that the departure of the initial conditions from the background is weighted by the inverse background error covariance matrix, while the differences between the model predictions $\mathbf{H}(\mathbf{x}_i)$ and observations \mathbf{y}_i are weighted by the inverse observation error covariance matrices. The 4D-Var analysis is computed as the initial condition which minimizes (32) subject to the model equation constraints (4)

$$\mathbf{x}_0^a = \arg \min \mathcal{J}(\mathbf{x}_0) \quad \text{subject to: } \mathbf{x}_i = \mathcal{M}_{t_0 \rightarrow t_i}(\mathbf{x}_0), \quad i = 1, \dots, N \quad (33)$$

The model (4) propagates the optimal initial condition (32) forward in time to provide the analysis at future times, $\mathbf{x}_i^a = \mathcal{M}_{t_0 \rightarrow t_i} \mathbf{x}_0^a$.

The large scale optimization problem (33) is solved numerically using a gradient-based technique. The gradient of (32) reads

$$\nabla \mathcal{J}(\mathbf{x}_0) = \mathbf{B}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \sum_{i=1}^N \left(\frac{\partial \mathbf{x}_i}{\partial \mathbf{x}_0} \right)^T \mathbf{H}_i^T \mathbf{R}_i^{-1} (\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i) \quad (34)$$

The 4D-Var gradient requires not only the linearized observation operator $\mathbf{H}_i = \mathcal{H}'(\mathbf{x}_i)$, but also the transposed derivatives of future states with respect to the initial conditions $(\partial \mathbf{x}_i / \partial \mathbf{x}_0)^T = M_{t_0 \rightarrow t_i}^T$. It can be demonstrated that the solution of the adjoint equations at the initial time provides the gradient of the cost function with respect to the initial condition in a computationally efficient way. The 4D-Var gradient can be obtained effectively by forcing the adjoint model with observation increments, and running it backwards in time. The construction of an adjoint model is a nontrivial task.

In the incremental formulation of 4D-Var [129,130], the estimation problem is linearized around the background trajectory. By expressing the state as $\mathbf{x}_i = \mathbf{x}_i^b + \delta \mathbf{x}_i$, $i = 1, \dots, N$, we have

$$\begin{aligned} \mathcal{J}'(\mathbf{x}_0) \delta \mathbf{x}_0 &= \frac{1}{2} \delta \mathbf{x}_0^T \mathbf{B}_0^{-1} \delta \mathbf{x}_0 + \frac{1}{2} \sum_{i=0}^N (\mathbf{H}_i \delta \mathbf{x}_i + d_i^b)^T \mathbf{R}_i^{-1} (\mathbf{H}_i \delta \mathbf{x}_i + d_i^b) \\ d_i^b &= \mathcal{H}(\mathbf{x}_i^b) - \mathbf{y}_i \end{aligned} \quad (35)$$

where $\delta \mathbf{x}_i = \mathbf{M}_{t_0 \rightarrow t_i} \cdot \delta \mathbf{x}_0$, and \mathbf{H}_i is the linearized observation operator. The incremental 4D-Var problem (35) uses linearized operators and leads to a quadratic cost function \mathcal{J}' whose minimizer is $\delta \mathbf{x}_0^a$. The incremental 4D-Var estimate is $\mathbf{x}_0^a = \mathbf{x}_0^b + \delta \mathbf{x}_0^a$. A new linearization can be performed about this estimate and the incremental problem (35) can be solved again to improve the resulting analysis. The iterated incremental 4D-Var is nothing but a sequential quadratic programming approach [131] to solve the constrained optimization problem (33).

Weakly constrained 4D-Var avoids the assumption of a perfect model, implicit in the formulation (33), at the expense of solving a larger optimization problem. The state \mathbf{x}_i at t_i is allowed to differ from the model prediction; the difference is the model error, considered to be a random variable. With the assumption that the model is not biased, and the model error is normally distributed, we have that

$$\mathbf{x}_i = \mathcal{M}_{t_{i-1} \rightarrow t_i}(\mathbf{x}_{i-1}) + \eta_i, \quad \eta_i \in \mathcal{N}(\mathbf{0}, \mathbf{Q}_i), \quad i = 1, \dots, N$$

The weakly constrained 4D-Var estimate of $\mathbf{x} = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N]$ is the unconstrained minimizer of the following cost function:

$$\begin{aligned} \mathcal{J}^{\text{weak}}(\mathbf{x}) = & \frac{1}{2} (\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2} \sum_{i=1}^N (\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i)^T \mathbf{R}_i^{-1} (\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i) \quad (36) \\ & + \frac{1}{2} \sum_{i=1}^N (\mathbf{x}_i - \mathcal{M}_{t_{i-1} \rightarrow t_i}(\mathbf{x}_{i-1}))^T \mathbf{Q}_i^{-1} (\mathbf{x}_i - \mathcal{M}_{t_{i-1} \rightarrow t_i}(\mathbf{x}_{i-1})) \end{aligned}$$

The optimization variables are the model states at all times $\mathbf{x} \in \mathbb{R}^{n(N+1)}$, and therefore the resulting optimization problem is of larger dimension than that for strongly-constrained 4D-Var.

3.7. A Comparison of Various Data Assimilation Approaches

Insightful comparisons of the relative merits of EnKF and 4D-Var [132–134], and of EnKF and 3D-Var [95] have been reported in the context of numerical weather prediction. Similar arguments hold in the context of CTMs. A comprehensive comparison of the performance of several methods applied to the assimilation of ozone satellite measurements in a global chemistry and transport framework has recently been carried out [17].

EnKF is simple to implement, while 4D-Var requires the construction of adjoint models, a non-trivial task in the presence of stiff chemistry [61]. EnKF allows for a simple integration of model errors, whereas strong-constrained 4D-Var assumes a perfect model. The ensemble propagates the forecast covariance and an estimate of the background covariance is readily available at the beginning of the next assimilation cycle.

On the other hand the 4D-Var optimal solution is consistent with model dynamics throughout the assimilation window. 4D-Var naturally incorporates asynchronous observations while for EnKF asynchronous observations require a more involved framework [114]. A consistent derivation of the initial ensemble in EnKF is difficult. Moreover, in the presence of stiff chemistry, each application of the filter throws the model state off balance; consequently, after each assimilation cycle a new stiff transient will be introduced, and this may considerably impact the computational time needed to advance the model state for each ensemble member.

Very recent work has focused on the development of hybrid data assimilation methods, that attempt to combine the advantages of both variational and ensemble techniques [135,136].

4. Challenges to Chemical Data Assimilation

4.1. Data Assimilation Inputs

Running chemical transport models requires several essential components. Firstly, model-ready emission files have to be processed using emission inventories. Secondly, meteorological states are needed for commonly-used off-line CTMs. Lastly, the realistic initial concentrations for various constituents are required. A spin-up period is often chosen to generate such initial fields when no previous run results are available. Chemical data assimilation adds two more components to these, *i.e.*, the observational inputs and model background error statistics.

Obtaining and utilizing atmospheric chemical observations remains a challenge. Currently atmospheric chemical observations come from many different sources. They vary greatly in their

dissemination methods, availability, data reliability due to different validation and quality control methods, instrument descriptions and measurement uncertainties, temporal and spatial resolutions, and data formats. “Integrated Global Atmospheric Chemistry Observations” (IGACO) is an ongoing effort as a component of the Integrated Global Observing Strategy (IGOS) partnership [137]. To manage and utilize the observational data from various sources, preprocessing is often required. In the preprocessing, the observations with higher spatial and temporal resolutions can be re-gridded into the model grid and model representative errors can be approximated in such steps [11,15].

4.2. Construction of Adjoint Chemical Transport Models for 4D-Var

The most important challenge posed by 4D-Var data assimilation is the need to construct and maintain an adjoint of the chemical transport model. The construction of adjoint models is a labor intensive and error prone task. Moreover, the adjoint is specific to the chemical transport model version at hand; any new release of an improved version of the code requires changes in the adjoint model to reflect the changes in the forward model. The construction of the adjoint model is a continuous process that follows closely the development of the forward chemical transport model.

The adjoint of a chemical transport model consists of adjoints of all the individual science processes [61,138,139]. Two routes can be taken toward building science process adjoints. In the *continuous adjoint approach* the mathematical equations governing the science model are differentiated analytically, in an appropriate framework, to obtain a new set of “adjoint” mathematical equations. The latter system is discretized with the numerical methods of choice. In the *discrete adjoint approach* one starts with the numerical implementation of the science process, as available in the CTM, and differentiates it in the discrete setting. The resulting computational process yields the sensitivities of the numerical solution. Discrete adjoints can be obtained with the help of automatic differentiation [140,141].

The two approaches lead to different results, since taking the adjoint and discretization operations do not commute. Considerable work has been done to understand the theoretical properties of different types of adjoint models, and the implications they have on sensitivity analysis and chemical data assimilation [142–149]. A good choice is to use continuous adjoints for advection, and discrete adjoints for other processes like chemistry and particles [16]. Recent work has proposed the use of simplified adjoint models for 4D-Var chemical data assimilation [150].

Specialized tools have been developed to assist the construction of chemical transport adjoint models. The chemical kinetic preprocessor KPP produces efficient code for the simulation of stiff chemistry, together with efficient tangent linear and discrete adjoint chemical kinetic models [151–153]. Sustained effort from several research groups in the past few years has lead to the construction of complete adjoints for the widely used chemical transport models STEM [1,61], CMAQ [27], and GEOS-Chem [62,154].

4.3. Correct Models of the Background and Observation Error Covariances

The quality of the assimilation depends on the accuracy with which the background and observation error covariances are known; misspecification of these covariances directly impacts the accuracy of the analysis [155]. Models of observation errors include information about the measuring instrument noise

and bias (measurement error), and about the resolution with which the model reproduces the pointwise variability of the physical system and the quality of the observation operator (representativeness error).

Background error covariances determine the relative weighting between observations and a priori data, and dictate how the information is spread in space and among variables. Background error covariances are based on models of the error at the current time (or at initial time in 4D-Var). In case of cyclic data assimilation the analysis error covariance from the previous cycle, transported to the current time, may be used as the new background error covariance. Background error covariance matrices need to:

- capture the spatial error correlations created by the flow (transport and diffusion),
- capture the inter-species error correlations created by the chemical interactions,
- have full rank, such that terms of the form $\mathbf{x}^T \mathbf{B}^{-1} \mathbf{x}$ make sense, and
- allow for computationally efficient evaluations of matrix vector operations of the form $\mathbf{B} \mathbf{x}$, $\mathbf{B}^{1/2} \mathbf{x}$, and $\mathbf{B}^{-1} \mathbf{x}$.

Reasonable approximations and representations of the background error are crucial to data assimilation applications. Chai [11] has estimated the CTM error statistics through both the NMC (National Meteorological Center) and the Hollingsworth-Lönnberg methods. The statistics were successfully implemented through a truncated singular vector decomposition regularization method in 4D-Var data assimilation applications with the STEM model.

An autoregressive (AR) model approach to represent background error covariance matrices has been proposed in [156]. The background error field is assumed to have zero mean $\langle \varepsilon^b \rangle = 0$, and background covariance \mathbf{B} . The background state error field is modeled as a multilateral autoregressive (AR) process [157] of the form

$$\delta \mathbf{x}_{i,j,k}^b = \alpha_{i\pm 1,j,k} \delta \mathbf{x}_{i\pm 1,j,k}^b + \beta_{i,j\pm 1,k} \delta \mathbf{x}_{i,j\pm 1,k}^b + \gamma_{i,j,k\pm 1} \delta \mathbf{x}_{i,j,k\pm 1}^b + \sigma_{i,j,k} \xi_{i,j,k} \quad (37)$$

Here (i, j, k) are gridpoint indices on a three dimensional structured grid. The model (37) captures the correlations among neighboring grid points, with α, β, γ representing the correlation coefficients in the x, y and z directions respectively. The last term represents the additional uncertainty at each grid point, with $\xi \in \mathcal{N}(0, 1)$ normal random variables and σ local error variances. The AR model coefficients α, β, γ depend on the wind field vector at each point and are obtained from a monotonic discretization of the linearized dynamics on the structured grid. Relation (37), with proper coefficients, is nothing but a finite difference approximation of the advection-diffusion equation. This approach accurately captures the flow dependent correlations, does not need any prior assumptions regarding correlation lengths, can be extended to include chemical correlations, is computationally inexpensive, and results in well conditioned covariance matrices.

A simplified approach proposed in [158] constructs multidimensional correlation matrices as tensor products of one-dimensional correlations. This method has resulted in improved chemical data assimilation results with GEOS-Chem.

In the context of 4D-Var chemical data assimilation the hybrid approach discussed in [159] estimates the analysis covariance at the end of one assimilation window (*i.e.*, the background covariance at the beginning of the next window). An ensemble drawn from the background distribution is run side by

side with the optimization process, the subspace of errors corrected by 4D-Var is identified, and this information is used to transform the background ensemble into one that samples the analysis distribution.

4.4. Estimating the Quality of the Analysis

At the end of any data assimilation calculation one would like to estimate the quality of the analysis, *i.e.*, the magnitude of the posterior estimate error, and its impact on given aspects of the subsequent forecast. The most robust way is to use an independent data set (not used directly in assimilation, and not correlated with the assimilated observations). The discrepancy between the model results and the independent data set, before and after data assimilation, gives a good indication of the error reduction through assimilation.

In operational data assimilation the goal is to improve forecasts. The model is initialized with the analysis that incorporates information from all past observations; the model is run, and the forecast is compared against the new observations that become available in the subsequent time window. Well established metrics for model-observation discrepancies in forecast mode are the forecast skill scores [107]. To estimate the quality of the analysis in hindcast (reanalysis) mode one can withhold part of the data from the assimilation system, and use it to assess the accuracy of the result.

The data assimilation system itself has the ability to provide estimates of the posterior error magnitude. If an ensemble Kalman filter is used, estimates of the analysis covariance matrices \mathbf{P}_i^a are readily available at each assimilation time t_i . For variational methods additional calculations are necessary. The second order adjoint (SOA) of the chemical transport model [160,161] computes matrix vector products between the Hessian of the 3D/4D-Var cost function $\nabla_{\mathbf{x}_0, \mathbf{x}_0}^2 \mathcal{J}$ and user-supplied vectors. The SOA model provides information about the a posteriori error via the observation that the Hessian inverse approximates the posterior error covariance [162]

$$\mathbf{A}_0 \approx (\nabla_{\mathbf{x}_0, \mathbf{x}_0}^2 \mathcal{J}(\mathbf{x}_0^a))^{-1}$$

In [160] the smallest Hessian eigenvalues, and the associated eigenvectors, were computed using a Lanczos approach for an ozone data assimilation problem. (The Lanczos approach uses only matrix-vector products, provided by the SOA). The inverses of the smallest eigenvalues, and their eigenvectors, approximate the principal components of the 4D-Var analysis error.

5. Chemical Data Assimilation Results with CMAQ

5.1. CMAQ Model Error Statistics

As described in Section 4.3, model background error statistics are crucial in data assimilation applications. It is important to gain knowledge of model uncertainties for a CTM with its specific setups, including the gas phase chemistry mechanism and aerosol module, model resolution, emission inventories, *etc.* In the following vertical ozone error statistics estimation and ozone OI data assimilation test runs, the CMAQ model is from the released version 4.6 with the Carbon Bond IV (CBIV) gas-phase chemical mechanism and aerosol module version 4 (AERO-4) [163,164]. In the aerosol optical depth assimilation test cases presented in Section 5.3, an updated Carbon Bond version (CB05) is used with the

same AERO-4 aerosol module [165]. The 2001 National Emission Inventory (NEI) with recent updates is used.

A computational grid with a 12-km resolution covering the contiguous United States (CONUS, shown in Figure 1) used in the United States National Air Quality Forecast Capability (NAQFC) is adopted here [166]. A sub-domain covering the Mid-Atlantic region (see [29] for detail) is used in ozone data assimilation tests and the horizontal error statistics estimation. The aerosol optical depth (AOD) assimilation tests in Section 5.3 and vertical error statistics estimation using the ozonesondes are carried out over the CONUS domain. The grid has a 22 sigma pressure hybrid vertical layers spanning from surface to 100 hPa.

Repeating the steps described in [11], the CMAQ error statistics were estimated using the Hollingsworth-Lönnberg method. AIRNow hourly ozone observations in the sub-domain were used to calculate the horizontal error statistics.

Model error correlation coefficients are shown in Figure 2 (left) as a function of horizontal distance between pairs of two surface stations. Pair density is also shown to indicate the number of station pairs used in the calculation. The CMAQ background model error for ozone is about 14 ppbv and its horizontal correlation length is around 50 km. Ozonesonde profiles from the measurements sites shown in Figure 1 were used to calculate the vertical model error statistics shown in Figure 2 (right) as a correlation coefficient contour plot.

Figure 1. CMAQ CONUS computational domain and ozonesonde locations. Red circles indicate ozonesonde locations where observations are used to calculate vertical model error statistics. Unit of longitude and latitude: degree.

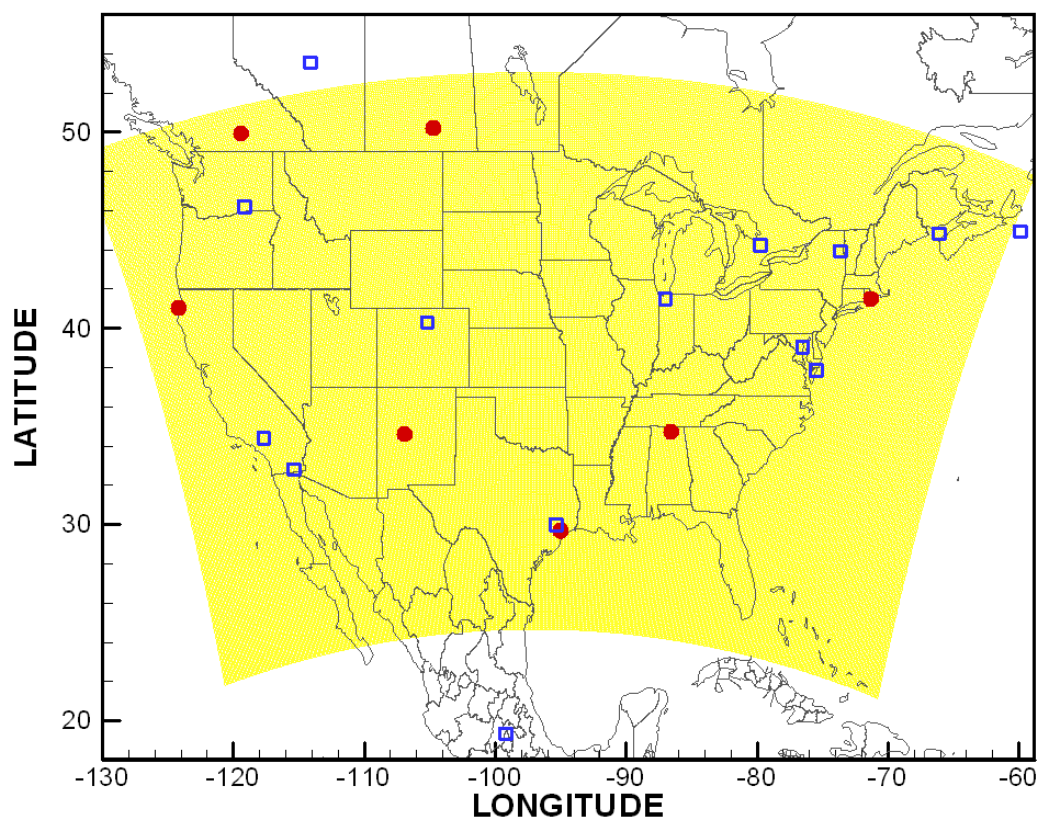
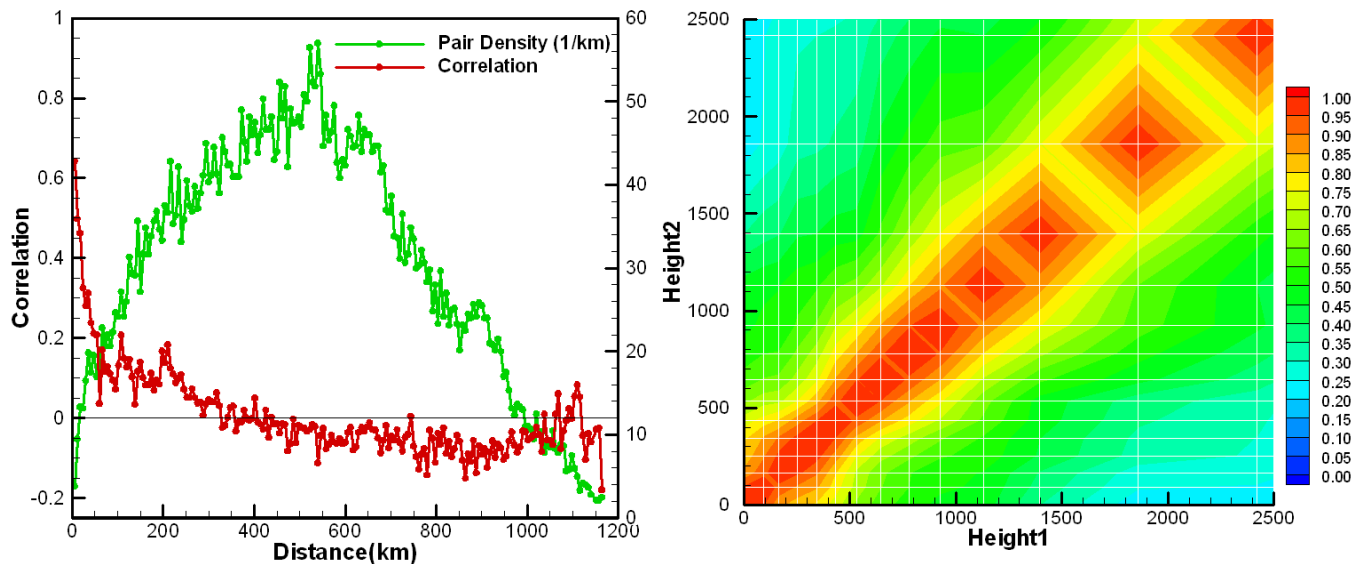


Figure 2. Ozone error statistics results through Hollingsworth-Lönnerberg approach. AIRNow observations are used to get horizontal error statistics (left). Ozonesonde observations are used in calculating vertical model error statistics (right). Unit of height: meter.



5.2. AIRNow Ozone Assimilation

Two CMAQ data assimilation systems are built with 4D-Var and OI approaches separately. The data assimilation time window is set to start from 1200Z on August 5, 2007 until 1200Z on August 6, 2007. In this 24-h period, the AIRNow hourly-averaged observations are assimilated and the observations are assumed to be un-correlated with each other and have a uniform root-mean-square error set as 3.3 ppbv. To check the effect of the data assimilation tests, an additional “forecast” day, starting from 1200Z on August 6, 2007 until 1200Z on August 7, 2007 is continuously run and will be evaluated against the AIRNow observations that are not assimilated in any of the assimilation tests.

In the 4D-Var data assimilation, the initial ozone concentrations are chosen as the only control parameters to be adjusted. Currently, the ozone background error covariance matrix B is assumed to be diagonal, with the root-mean-square errors set as 14.3 ppbv at every grid point. A quasi-Newton limited memory L-BFGS [167,168] is used in the cost functional minimization. The maximum number of iterations is set to be 15.

For the OI data assimilation runs, the assimilation happens every hour by combining the model results with the observations. To illustrate the effect of the background error covariance, we designed a case that eliminates the spatial correlation usage, both horizontally and vertically. It is listed in Table 1 as Case 3. In the other OI case, *i.e.*, Case 4 in Table 1, the horizontal background error covariance is approximated as

$$B = H \otimes V \otimes C \quad (38)$$

where H and V are matrices that represent the error correlation in horizontal and vertical directions respectively. C is the error covariance matrix at a single grid point that represents the error variances. \otimes denotes the Kronecker product [169]. The horizontal correlation between two grid points are calculated using a simple function $e^{-\frac{\Delta}{l_h}}$, where Δ is the horizontal distance between the two grid points

and l_h is set as 48 km. The background error variances are 14.3^2 ppbv². Instead of using a constant vertical correlation structure obtained in Section 5.1, we use the boundary layer depth information available from the meteorological inputs. In Case 4, the vertical correlation coefficients are set as 1.0 for any two model grid layers inside the boundary layers. Otherwise, it is assumed there is no correlation for the background error.

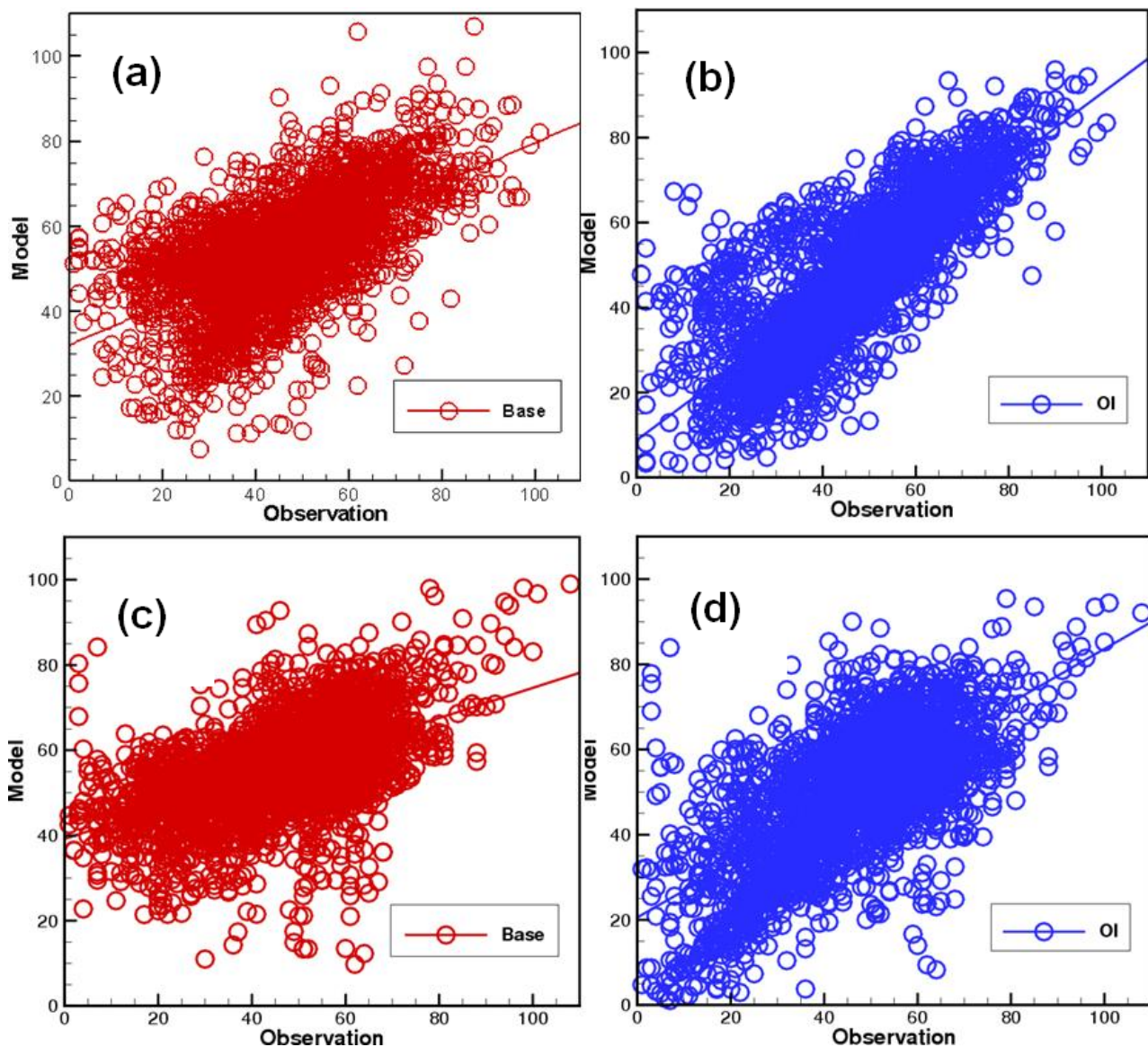
Table 1. Model ozone biases and root-mean-square errors (RMSE) against AIRNow observations during 8:00 am–8:00 pm local time on Day 1 (August 5, 2007) and Day 2 (August 6, 2007). Case 1 is the base case, *i.e.*, without data assimilation. B: background error covariance matrix. Unit: ppbv.

| | Assimilation | B | Day 1 Bias | Day 1 RMSE | Day 2 Bias | Day 2 RMSE |
|---|--------------|-------------------------|------------|------------|------------|------------|
| 1 | N/A | N/A | 8.3 | 15.9 | 8.7 | 16.3 |
| 2 | 4D-Var | Diagonal | −0.8 | 11.0 | 7.6 | 15.6 |
| 3 | OI | Diagonal | 2.6 | 12.7 | 7.5 | 15.8 |
| 4 | OI | $H \otimes V \otimes C$ | −1.3 | 13.2 | 3.1 | 12.8 |

Figure 3 shows the comparisons between the model predictions and observations of ozone during the assimilation and forecast periods for the base case and the OI case with spatial correlation accounted, *i.e.*, Cases 1 and 4 in Table 1 respectively. After assimilation, the model has a much better agreement with AIRNow ozone measurements. The correlation coefficient improved from 0.59 to 0.81 during the daytime, 1300–2400Z on August 5, 2007. For the next day “forecast” run, the improvement of model ozone predictions is also apparent, with the daytime correlation coefficient between model and observations changed from 0.56 to 0.68.

Table 1 lists the comparisons between the different assimilation cases and the base case run. All three assimilation cases prove to be able to generate better results not only in the assimilation day, but also in the next day “forecast”. Without fully accounting for the background error covariance, the 4D-Var case still generates the best results during the first day in terms of the model biases and root-mean-square errors (RMSEs) against the AIRNow observations. By utilizing the error statistics obtained from Section 5.1, Case 4 with the simple OI method provides the best “forecast” for the second day, where the model bias and RMSE are reduced from 8.7 ppbv to 3.1 ppbv and from 16.3 ppbv to 12.8 ppbv respectively. Without using the model background error spatial correlations, Case 3 is only slightly better than the base case for the “forecast” day. From Table 1, we can see that the 4D-Var case has comparable results as Case 3, which implements the simple OI method. As indicated by the comparison between Case 3 and Case 4, replacing the diagonal background error covariance used in Case 2 with one accounting for the spatial correlation is expected to improve next day forecast for the 4D-Var case. It cannot be generalized to conclude the 4D-Var system has the same performance as OI approach. It has to be noted that the 4D-Var system is based upon CMAQ version 4.5 and the other cases implement CMAQ version 4.6.

Figure 3. Scatter plots of AIRNow ozone observations and CMAQ predictions for the assimilation (upper, (a) and (b)) and hindcast (lower, (c) and (d)) period of the base (left, (a) and (c)) and OI assimilation (right, (b) and (d)) runs. (a) and (b): 1300–2400Z on August 5, 2007; (c) and (d): 1300–2400Z August 6, 2007. Correlation coefficients are 0.59, 0.81, 0.56, and 0.68 for (a), (b), (c), and (d) plots, respectively.



5.3. MODIS Aerosol Optical Depth Assimilation

Compared to ozone predictions, CMAQ PM_{2.5} predictions are much worse for the NAQFC experimental runs [170]. MODIS AOD observations can be used to constrain the model input parameters such as emissions or initial concentrations. As a test case here, we assimilate the MODIS AOD using OI approach.

In the test, the MODIS AOD fine mode products are used. The model counterpart can be reconstructed by integrating the hourly extinction coefficients over the whole vertical columns. The extinction coefficients calculated from two visibility methods, Mie theory approximation and mass reconstruction

method [171], are quite similar and we chose to use the results from the mass reconstruction method. Both Terra and Aqua fine mode AOD data are used during the assimilation time period (August 14–20, 2009). Before the data assimilation tests, the AOD background error statistics is first estimated using Hollingsworth-Lönnberg approach. As an integrated quantity, only horizontal correlation is needed in constructing the error statistics. The horizontal correlation between two grid points are modeled as a function $e^{-\frac{\Delta}{l_h}}$, where l_h is set as 84 km. The AOD background error is assumed to be $0.6 \times \text{AOD}_{\text{MODIS}}$. In the OI assimilation, the analysis process takes place once a day, at 1700Z, which is close to the midpoint of the Terra and Aqua observation time. The adjust factor of AOD at each grid point is then uniformly applied to mass concentrations of all the aerosol species.

Figure 4. MODIS AOD (fine mode) and CMAQ reconstructed AOD. AOD-Recona and AOD-Reconb are calculated before and after assimilation. The differences (AOD-Recona-AOD-Reconb) are also shown.

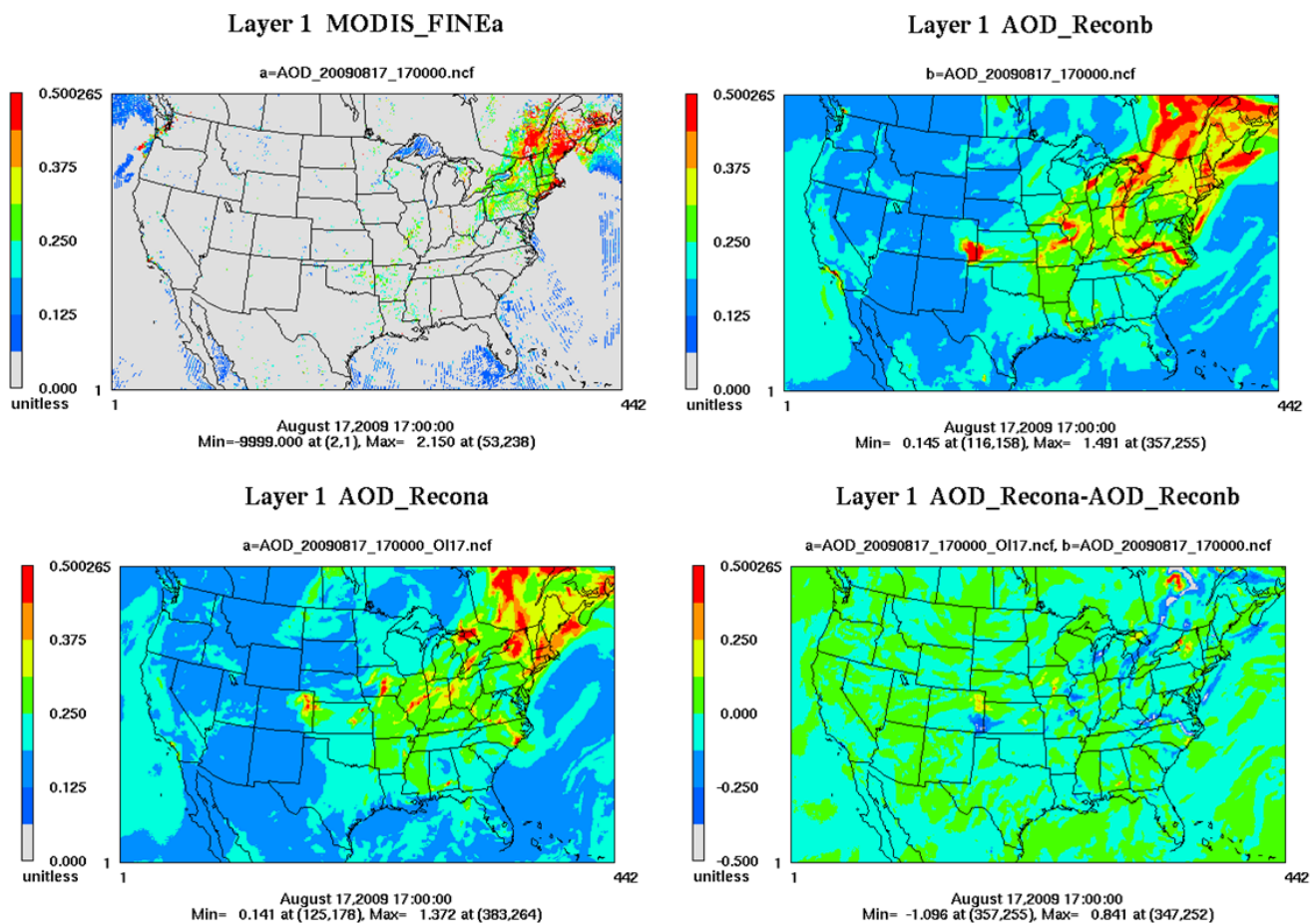


Figure 4 shows the AOD distributions from MODIS and CMAQ simulation with and without data assimilation. The differences after assimilation are also shown. Note that the MODIS AOD data are quite sparse, but the OI assimilation spreads the information using the obtained horizontal correlations between AOD background errors. The CMAQ PM_{2.5} predictions before and after AOD assimilations are evaluated using the AIRNow PM_{2.5} observations for each day. Table 2 shows the correlations between the MODIS observed and CMAQ predicted AOD before and after OI in the upper Midwest and Northeast of the U.S. (see [172] for region definition), where most of data reside. It is seen that the R^2 improve over

four out of six days in both regions. It is encouraging as the correlation between the column quantity of AOD and the surface PM_{2.5} is not linear. A better reconstructed AOD cannot guarantee better predictions of surface aerosol. The current simplification of placing the observations at a single time each day and adjusting all the aerosol species using a single factor will be modified in the future. In addition, switching OI approach to 3D-Var or 4D-Var method is expected to generate better assimilation results.

Table 2. Correlation between CMAQ PM_{2.5} predictions and AIRNow hourly observations in Upper Midwest (UM) and Northeast (NE) US before and after (OI) MODIS AOD assimilation.

| R^2 | 8/15/09 | 8/16/09 | 8/17/09 | 8/18/09 | 8/19/09 | 8/20/09 |
|-------|---------|---------|---------|---------|---------|---------|
| UM | 0.420 | 0.138 | 0.355 | 0.154 | 0.234 | 0.021 |
| UM-OI | 0.399 | 0.178 | 0.311 | 0.180 | 0.270 | 0.041 |
| NE | 0.253 | 0.416 | 0.097 | 0.070 | 0.156 | 0.217 |
| NE-OI | 0.306 | 0.367 | 0.110 | 0.207 | 0.171 | 0.206 |

6. Conclusions and Future Directions in Chemical Data Assimilation

New developments in chemical data assimilation techniques and algorithms, and the increased volume and diversity of available chemical measurements, have opened exciting opportunities for better science through the integration of chemical transport models and observations. Chemical data assimilation has begun to play an essential role in air quality assessments for environmental management. Widely used chemical transport models such as STEM, CMAQ, and GEOS-Chem, have been endowed with adjoint sensitivity analysis and data assimilation capabilities, and are now being used by the community to answer important scientific questions. The availability of these tools, and the growing importance of chemical weather forecasting to society, should help stimulate significant advances in chemical data assimilation in the foreseeable future.

Future advances will require a sustained development of new chemical data assimilation algorithms. While there is much to build upon from the assimilation experience in weather prediction, there are significant differences and challenges that are specific to chemical weather. Promising possibilities are opened up by combining the strengths of 4D-Var and EnKF techniques in hybrid data assimilation methods. Feedbacks between the meteorological and air quality components, which have mostly been studied as separate systems, are critical to improving the understanding of air quality. Future work needs to build the infrastructure required to couple meteorological and air quality forecasting and data assimilation systems. Finally, current chemical data assimilation system capabilities should be extended to enable the optimal design of the observing systems, and to rigorously quantify the informational value added by each instrument in heterogeneous sensor networks.

Acknowledgements

The work of A. Sandu has been supported in part by NSF through awards NSF OCI-0904397, NSF CCF-0916493, NSF DMS-0915047.

References

1. Carmichael, G.; Chai, T.; Sandu, A.; Constantinescu, E.; Daescu, D. Predicting air quality: Improvements through advanced methods to integrate models and measurements. *J. Comput. Phys.* **2008**, *227*, 3540–3571.
2. Daley, R. *Atmospheric Data Analysis*; Cambridge University Press: Cambridge, UK, 1991.
3. Courtier, P.; Andersson, E.; Heckley, W.; Pailleux, J.; Vasiljevic, D.; Hamrud, M.; Hollingsworth, A.; Rabier, F.; Fisher, M. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Q. J. R. Meteorol. Soc.* **1998**, *124*, 1783–1807.
4. Rabier, F.; Jarvinen, H.; Klinker, E.; Mahfouf, J.; Simmons, A. The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.* **2000**, *126*, 1148–1170.
5. Kalnay, E. *Atmospheric Modeling, Data Assimilation and Predictability*; Cambridge University Press: Cambridge, UK, 2002.
6. Navon, I. *Data assimilation for numerical weather prediction: A Review*; Springer: Berlin, Germany, 2009.
7. Evensen, G. *Data Assimilation: The Ensemble Kalman Filter*; Springer: Berlin, Germany, 2007.
8. Chai, T.; Carmichael, G.; Sandu, A.; Tang, Y.; Daescu, D. Chemical data assimilation with TRACE-P aircraft measurements. *J. Geophys. Res.* **2006**, *111*, D02301.
9. Hakami, A.; Seinfeld, J.; Chai, T.; Tang, Y.; Carmichael, G.; Sandu, A. Adjoint sensitivity analysis of ozone non-attainment over the continental United States. *Environ. Sci. Technol.* **2006**, *40*, 3855–3864.
10. Zhang, L.; Sandu, A. Data assimilation in multiscale chemical transport models. *LNCS* **2007**, *4487*, 1026–1033.
11. Chai, T.; Carmichael, G.; Tang, Y.; Sandu, A.; Hardesty, M.; Pilewskie, P.; Whitlow, S.; Browell, E.; Avery, M.; Thouret, V.; *et al.* Four dimensional data assimilation experiments with ICARTT (international consortium for atmospheric transport and transformation) ozone measurements. *J. Geophys. Res.* **2007**, *112*, D12S15.
12. Zhang, L.; Constantinescu, E.; Sandu, A.; Tang, Y.; Chai, T.; Carmichael, G.; Byun, D.; Olaguier, E. An adjoint sensitivity analysis and 4D-Var data assimilation study of Texas air quality. *Atmos. Environ.* **2008**, *42*, 5787–5804.
13. Singh, K.; Eller, P.; Sandu, A.; Bowman, K.; Jones, D.; Lee, M. Improving GEOS-chem model forecasts through profile retrievals from tropospheric emission spectrometer. *LNCS* **2009**, *5545*, 302–311.
14. Gou, T.; Singh, K.; Sandu, A. Chemical data assimilation with CMAQ: Continuous vs. discrete advection adjoints. *LNCS* **2009**, *5545*, 312–321.
15. Chai, T.; Carmichael, G.; Tang, Y.; Sandu, A. Regional NO_x emission inversion through a four-dimensional variational approach using SCIAMACHY tropospheric NO_2 column observations. *Atmos. Environ.* **2009**, *43*, 5046–5055.
16. Gou, T.; Sandu, A. Continuous versus discrete advection adjoints in chemical data assimilation with CMAQ. *Atmos. Environ.* **2011**, *45*, 4868–4881.

17. Singh, K.; Sandu, A.; Parrington, M.; Jones, D.; Bowman, K.; Lee, M. Ozone data assimilation with GEOS-Chem: A comparison between 3D-Var, 4D-Var, and suboptimal Kalman filter approaches. Unpublished work, **2011**.
18. Alexe, M.; Sandu, A. Adaptive Solution of Time-Dependent Inverse Problems with the Discrete Adjoint Method. In *Proceedings of the International Conference on Computational Science (ICCS-2011)*, Bali, Indonesia, 26–28 October 2011, in press.
19. Stewart, R. Multiple steady states in atmospheric chemistry. *J. Geophys. Res.* **1993**, *98*, 20601–20612.
20. Menut, L. Adjoint modeling for atmospheric pollution process sensitivity at regional scale. *J. Geophys. Res.* **2003**, *108*.
21. Dee, D.; da Silva, A. Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.* **1998**, *124*, 269–295.
22. Van der A, R.J.; Allaart, M.A.F.; Eskes, H.J. Multi sensor reanalysis of total ozone. *Atmos. Chem. Phys.* **2010**, *10*, 11277–11294.
23. Evans, L. *Partial Differential Equations*; American Mathematical Society: Providence, RI, USA, 1998.
24. Community Modeling and Analysis System (CMAS). Available online: <http://www.cmascenter.org> (accessed on 2 May 2011).
25. Byun, D. Dynamically consistent formulations in meteorological and air quality models for multiscale atmospheric studies part I: Governing equations in a generalized coordinate system. *J. Atmos. Sci.* **1999**, *56*, 3789–3807.
26. Byun, D. Dynamically consistent formulations in meteorological and air quality models for multiscale atmospheric studies part II: Mass conservation issues. *J. Atmos. Sci.* **1999**, *56*, 3808–3820.
27. Hakami, A.; Henze, D.; Seinfeld, J.; Singh, K.; Sandu, A.; Kim, S.; Byun, D.; Li, Q. The adjoint of CMAQ. *Environ. Sci. Technol.* **2007**, *41*, 7807–7817.
28. The AIRNow network. Available online: <http://airnow.gov> (accessed on 2 May 2011).
29. Sandu, A.; Chai, T.; Carmichael, G. Integration of models and observations: A modern paradigm for air quality. In *Modelling of Pollutants in Complex Environments*; Hanrahan, G., Ed.; ILM Publications: St Albans, UK, 2010; Volume 2, Chapter 15, pp. 419–434.
30. Zubrow, A.; Chen, L.; Kotamarthi, V. EAKF-CMAQ: Introduction and evaluation of a data assimilation for CMAQ based on the ensemble adjustment Kalman filter. *J. Geophys. Res.* **2008**, *113*, D09302.
31. Jeuken, A.; Eskes, H.; van Velthoven, P.; Kelder, H.; Holm, E. Assimilation of total ozone satellite measurements in a three-dimensional tracer transport model. *J. Geophys. Res.* **1999**, *104*, 5551–5563.
32. Thompson, A.; Witte, J.; McPeters, R.; Oltmans, S.; Schmidlin, F.; Logan, J.; Fujiwara, M.; Kirchhoff, V.; Posny, F.; Coetzee, G.; *et al.* Southern hemisphere additional ozonesondes (SHADOZ) 1998–2000 tropical ozone climatology—1. Comparison with total ozone mapping spectrometer (TOMS) and ground-based measurements. *J. Geophys. Res.* **2003**, *108*, 8238; doi:10.1029/2001JD000967.

33. Thompson, A.; Witte, J.; Schmidlin, F.; Logan, J.; Fujiwara, M.; Kirchhoff, V.; Posny, F.; Coetzee, G.; Hoegger, B.; Kawakami, S.; *et al.* Southern hemisphere additional ozonesondes (SHADOZ) 1998–2000 tropical ozone climatology—2. Tropospheric variability and the zonal wave-one. *J. Geophys. Res.* **2003**, *108*, 8241; doi:10.1029/2002JD002241.
34. Thompson, A.; Witte, J.; Smit, H.; Oltmans, S.; Johnson, B.; Kirchhoff, V.; Schmidlin, F. Southern hemisphere additional ozonesondes (SHADOZ) 1998–2000 tropical ozone climatology—3. Instrumentation, station-to-station variability, and evaluation with simulated flight profiles. *J. Geophys. Res.* **2007**, *112*, D03304; doi:10.1029/2005JD007042.
35. Japan National Institute for Environmental Studies (NIES) lidars. Available online: <http://www-lidar.nies.go.jp> (accessed on 2 May 2011).
36. Yumimoto, K.; Uno, I.; Sugimoto, N.; Shimizu, A.; Liu, Z.; Winker, D. Adjoint inversion modeling of Asian dust emission using lidar observations. *Atmos. Chem. Phys.* **2008**, *8*, 2869–2884.
37. Fishman, J.; Bowman, K.W.; Burrows, J.P.; Richter, A.; Chance, K.V.; Edwards, D.P.; Martin, R.V.; Morris, G.A.; Pierce, R.B.; Ziemke, J.R.; *et al.* Remote sensing of tropospheric pollution from space. *Bull. Am. Meteorol. Soc.* **2008**, *89*, 805–821.
38. Martin, R.V. Satellite remote sensing of surface air quality. *Atmos. Environ.* **2008**, *42*, 7823–7843.
39. Aura satellite. Available online: <http://aura.gsfc.nasa.gov/index.html> (accessed on 2 May 2011).
40. Schoeberl, M.; Douglass, A.; Hilsenrath, E.; Bhartia, P.; Beer, R.; Waters, J.; Gunson, M.; Froidevaux, L.; Gille, J.; Barnett, J.; *et al.* Overview of the EOS aura mission. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1066–1074.
41. Bovensmann, H.; Burrows, J.P.; Buchwitz, M.; Frerick, J.; Noël, S.; Rozanov, V.V. SCIAMACHY: Mission objectives and measurement modes. *J. Atmos. Sci.* **1999**, *56*, 127–150.
42. Kaufman, Y.; Tanre, D.; Remer, L.; Vermote, E.; Chu, A.; Holben, B. Applications of the quasi-inverse method to data assimilation. *J. Geophys. Res.* **1997**, *102*, 17051–17067.
43. Remer, L.; Kaufman, Y.; Tanre, D.; Mattoo, S.; Chu, D.; Martins, J.; Li, R.R.; Ichoku, C.; Levy, R.; Kleidman, R.; *et al.* The MODIS aerosol algorithm, products, and validation. *J. Atmos. Sci.* **2005**, *62*, 947–973.
44. Matsui, T.; Kreidenweis, S.; Pielke, R.; Schichtel, B.; Yu, H.; Chin, M.; Chu, D.; Niyogi, D. Regional comparison and assimilation of GOCART and MODIS aerosol optical depth across the eastern US. *Geophys. Res. Lett.* **2004**, *31*, L21101.
45. Adhikary, B.; Kulkarni, S.; Dallura, A.; Tang, Y.; Chai, T.; Leung, L.; Qian, Y.; Chung, C.; Ramanathan, V.; Carmichael, G. A regional scale chemical transport modeling of Asian aerosols with data assimilation of AOD observations using optimal interpolation technique. *Atmos. Environ.* **2008**, *42*, 8600–8615.
46. Zhang, J.; Reid, J.S.; Westphal, D.L.; Baker, N.L.; Hyer, E.J. A system for operational aerosol optical depth data assimilation over global oceans. *J. Geophys. Res.* **2008**, *113*, D10208.
47. Benedetti, A.; Morcrette, J.J.; Boucher, O.; Dethof, A.; Engelen, R.J.; Fisher, M.; Flentje, H.; Huneeus, N.; Jones, L.; Kaiser, J.W.; *et al.* Aerosol analysis and forecast in the European Centre for medium-range weather forecasts integrated forecast system: 2. Data assimilation. *J. Geophys. Res.* **2009**, *114*, D13205.

48. Štajner, I.; Riishøjgaard, L.P.; Rood, R.B. The GEOS ozone data assimilation system: Specification of error statistics. *Q. J. R. Meteorol. Soc.* **2001**, *127*, 1069–1094.
49. Nassar, R.; Jones, D.; Kulawik, S.; Worden, J.; Bowman, K.; Andres, R.; Suntharalingam, P.; Chen, J.; Brenninkmeijer, C.; Schuck, T.; *et al.* Inverse modeling of CO_2 sources and sinks using satellite observations of CO_2 from TES and surface flask measurements. *Atmos. Chem. Phys.* **2011**, *11*, 6029–6047.
50. Fehsenfeld, F.C.; Ancellet, G.; Bates, T.S.; Goldstein, A.H.; Hardesty, R.M.; Honrath, R.; Law, K.S.; Lewis, A.C.; Leaitch, R.; McKeen, S.; *et al.* International consortium for atmospheric research on transport and transformation (ICARTT): North America to Europe-overview of the 2004 summer field study. *J. Geophys. Res.* **2006**, *111*, D23S01.
51. New England Air Quality Study - Intercontinental Transport and Chemical Transformation (NEAQS-ITCT). Available online: <http://saga.pmel.noaa.gov/Field/NEAQS-ITCT> (accessed on 2 May 2011).
52. Gustafson, W.; Chapman, E.; Ghan, S.; Easter, R.C.; Fast, J. Impact on modeled cloud characteristics due to simplified treatment of uniform cloud condensation nuclei during NEAQS 2004. *Geophys. Res. Lett.* **2007**, *34*, 1–5.
53. Ravetta, F.; Ancellet, G.; Colette, A.; Schlager, H. Long-range transport and tropospheric ozone variability in the western mediterranean region during the intercontinental transport of ozone and precursors (ITOP-2004) campaign. *J. Geophys. Res.* **2007**, *112*, D10S46.
54. Petzold, A.; Weinzierl, B.; Huntrieser, H.; Stohl, A.; Real, E.; Cozic, J.; Fiebig, M.; Hendricks, J.; Lauer, A.; Law, K.; *et al.* Perturbation of the European free troposphere aerosol by North American forest fire plumes during the ICARTT-ITOP experiment in summer 2004. *Atmos. Chem. Phys.* **2007**, *7*, 5105–5127.
55. Molina, L.T.; Gaffney, J.S.; Singh, H.B. *Overview of MILAGRO/INTEX-B Campaign*; International Global Atmospheric Chemistry (Newsletter): Seattle, WA, USA, 2008.
56. Singh, H.B.; Brune, W.H.; Crawford, J.H.; Flocke, F.; Jacob, D.J. Chemistry and transport of pollution over the Gulf of Mexico and the Pacific: Spring 2006 INTEX-B campaign overview and first results. *Atmos. Chem. Phys.* **2009**, *9*, 2301–2318.
57. Global and Regional Earth-System Monitoring Using Satellite and In situ Data (GEMS). Available online: <http://gems.ecmwf.int> (accessed on 2 May 2011).
58. Protocol Monitoring for the GMES Service Element: Atmosphere (PROMOTE). Available online: <http://www.gse-promote.org> (accessed on 2 May 2011).
59. Monitoring Atmospheric Composition and Climate (MACC). Available online: <http://www.gmes-atmosphere.eu> (accessed on 2 May 2011).
60. Hollingsworth, A.; Engelen, R.; Textor, C.; Benedetti, A.; Boucher, O.; Chevallier, F.; Dethof, A.; Elbern, H.; Eskes, H.; Flemming, J.; *et al.* Toward a monitoring and forecasting system for atmospheric composition: The GEMS project. *Bull. Am. Meteorol. Soc.* **2008**, *89*, 1147–1164.
61. Sandu, A.; Daescu, D.; Carmichael, G.; Chai, T. Adjoint sensitivity analysis of regional air quality models. *J. Comput. Phys.* **2005**, *204*, 222–252.
62. Henze, D.; Hakami, A.; Seinfeld, J. Development of the adjoint of GEOS-chem. *Atmos. Chem. Phys.* **2007**, *7*, 2413–2433.

63. Khattatov, B.; Lamarque, J.F.; Lyjak, L.; Menard, R.; Levelt, P.; Tie, X.; Brasseur, G.; Gille, J. Assimilation of satellite observations of long-lived chemical species in global chemistry transport models. *J. Geophys. Res.* **2000**, *105*, 29–135.
64. Elbern, H.; Schmidt, H. Ozone episode analysis by four dimensional variational chemistry data assimilation. *J. Geophys. Res.* **2001**, *106*, 3569–3590.
65. Derber, J.; Parrish, D.; Lord, S. The new global operational analysis system at the national meteorological center. *Weather Forecast.* **1991**, *6*, 538–547.
66. Parrish, D.F.; Derber, J.C. The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Weather Rev.* **1992**, *120*, 1747–1763.
67. Cohn, S.; da Silva, A.; Guo, J.; Sienkiewicz, M.; Lamich, D. Assessing the effects of data selection with DAO physical space statistical analysis system. *Mon. Weather Rev.* **1998**, *126*, 2913–2926.
68. Gauthier, P.; Charette, C.; Fillion, L.; Koclas, P.; Laroche, S. Implementation of a 3D variational data assimilation system at the canadian meteorological centre. Part I: The global analysis. *Atmos. Ocean* **1999**, *37*, 103–156.
69. Dethof, A.; Hólm, E.V. Ozone assimilation in the ERA-40 reanalysis project. *Q. J. R. Meteorol. Soc.* **2004**, *130*, 2851–2872.
70. Polavarapu, S.; Ren, S.; Rochon, Y.; Sankey, D.; Ek, N.; Koshyk, J.; Tarasick, D. Data assimilation with the Canadian middle atmosphere model. *Atmos. Ocean* **2005**, *43*, 77–100.
71. Jackson, D.R. Assimilation of EOS MLS ozone observations in the Meteorological and oceanic data-assimilation system. *Q. J. R. Meteorol. Soc.* **2007**, *133*, 1771–1788.
72. Bei, N.; de Foy, B.; Lei, W.; Zavala, M.; Molina, L. Using 3DVAR data assimilation system to improve ozone simulations in the Mexico City basin. *J. Atmos. Chem. Phys.* **2008**, *8*, 7353–7366.
73. Dragani, R.; Dee, D. Progress in ozone monitoring and assimilation. *ECMWF Newsl.* **2008**, *116*, 35–42.
74. Inness, A.; Flemming, J.; Suttie, M.; Jones, L. *GEMS Data Assimilation System for Chemically Reactive Gases*; Technical Report RD Tech Memo 587; European Centre for Medium-Range Weather Forecasts: Reading, UK, 2009.
75. Flemming, J.; Inness, A.; Jones, L.; Eskes, H.J.; Huijnen, V.; Schultz, M.G.; Stein, O.; Cariolle, D.; Kinnison, D.; Brasseur, G. Forecasts and assimilation experiments of the Antarctic ozone hole 2008. *Atmos. Chem. Phys.* **2011**, *11*, 1961–1977.
76. Tang, Y.; Carmichael, G.; Horowitz, L.; Uno, I.; Woo, J.; Streets, D.; Dabdub, D.; Kurata, G.; Sandu, A.; Allan, J.; *et al.* Multi-scale simulations of tropospheric chemistry in the Eastern Pacific and U.S. West coast during spring 2002. *J. Geophys. Res. Atmos.* **2004**, *109*, D23S11.
77. Tang, Y.; Carmichael, G.; Seinfeld, J.; Dabdub, D.; Weber, R.; Huebert, B.; Clarke, A.; Guazzotti, S.; Sodeman, D.; Prather, K.; *et al.* Three-dimensional studies of aerosol ions and their size distribution in East Asia during spring 2001. *J. Geophys. Res.* **2004**, *109*, D19S23.
78. Maki, T.; Tanaka, T.Y.; Sekiyama, T.T.; Mikami, M. The impact of ground-based observations on the inverse technique of aeolian dust aerosol. *Sci. Online Lett. Atmos.* **2011**, *7A*, 21–24.
79. Yumimoto, K.; Uno, I. Adjoint inverse modeling of CO emissions over Eastern Asia using four-dimensional variational data assimilation. *Atmos. Environ.* **2006**, *40*, 6836–6845.

80. Hakami, A.; Henze, D.; Seinfeld, J.; Chai, T.; Tang, Y.; Carmichael, G.; Sandu, A. Adjoint inverse modeling of black carbon during ACE-Asia. *J. Geophys. Res.* **2005**, *110*, D14301.
81. Sandu, A.; Liao, W.; Carmichael, G.; Henze, D.; Seinfeld, J. Inverse modeling of aerosol dynamics using adjoints—Theoretical and numerical considerations. *Aerosol Sci. Technol.* **2005**, *39*, 1–18.
82. Henze, D.K.; Seinfeld, J.H.; Shindell, D.T. Inverse modeling and mapping US air quality influences of inorganic PM_{2.5} precursor emissions using the adjoint of GEOS-Chem. *Atmos. Chem. Phys. Discuss.* **2008**, *8*, 15031–15099.
83. Menard, R.; Cohn, S.; Chang, L.P.; Lyster, P. Assimilation of stratospheric chemical tracer observations using a Kalman filter I: Formulation. *Mon. Weather Rev.* **2000**, *128*, 2654–2671.
84. Lamarque, J.F.; Khattatov, B.; Gille, J. Constraining tropospheric ozone column through data assimilation. *J. Geophys. Res.* **2002**, *107*, 4651; doi:10.1029/2001JD001249.
85. Liao, W.; Sandu, A.; Chai, T.; Carmichael, G. Total energy singular vector analysis for atmospheric chemical transport models. *Mon. Weather Rev.* **2006**, *134*, 2443–2465.
86. Segers, A.; Eskes, H.; van der A, R.; van Oss, R.; van Velthoven, P. Assimilation of GOME ozone profiles and a global chemistry-transport model, using a Kalman filter with anisotropic covariance. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 477–502.
87. Clark, H.; Cathala, M.L.; Teyssedre, H.; Cammas, J.P.; Peuch, V.H. Cross-tropopause fluxes of ozone using assimilation of MOZAIC observations in a global CTM. *Tellus B* **2007**, *59*, 39–49.
88. Pierce, R.B.; Schaack, T.; Al-Saadi, J.A.; Fairlie, T.D.; Kittaka, C.; Lingenfelser, G.; Natarajan, M.; Olson, J.; Soja, A.; Zapotocny, T.; *et al.* Chemical data assimilation estimates of continental U. S. ozone and nitrogen budgets during the intercontinental chemical transport experiment-North America. *J. Geophys. Res.* **2007**, *112*, D12S21:1–D12S21:30.
89. Parrington, M.; Jones, D.; Bowman, K.; Thompson, A.; Tarasick, D.; Merrill, J.; Oltmans, S.; Leblanc, T.; Witte, J.; Millet, D. Impact of the assimilation of ozone from the tropospheric emission spectrometer on surface ozone across North America. *Geophys. Res. Lett.* **2009**, *36*, L04802.
90. Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **1994**, *99*, 10143–10162.
91. Constantinescu, E.; Sandu, A.; Chai, T.; Carmichael, G. Assessment of ensemble-based chemical data assimilation in an idealized setting. *Atmos. Environ.* **2007**, *41*, 18–36.
92. Constantinescu, E.; Sandu, A.; Chai, T.; Carmichael, G. Ensemble-based chemical data assimilation. I: General approach. *Q. J. R. Meteorol. Soc.* **2007**, *133*, 1229–1243.
93. Constantinescu, E.; Sandu, A.; Chai, T.; Carmichael, G. Ensemble-based chemical data assimilation. II: Covariance localization. *Q. J. R. Meteorol. Soc.* **2007**, *133*, 1245–1256.
94. Li, Z.; Navon, I. Optimality of 4D-Var and its relationship with the Kalman filter and Kalman smoother. *Q. J. R. Meteorol. Soc.* **2001**, *127*, 661–684.
95. Houtekamer, P.; Mitchell, H.; Pellerin, G.; Buehner, M.; Charron, M.; Spacek, L.; Hansen, B. Atmospheric data assimilation with the ensemble Kalman filter: Results with real observations. *Mon. Weather Rev.* **2005**, *133*, 604–620.

96. Laroche, S.; Dorval, E.; Canada, Q.; Gauthier, P.; Tanguay, M.; Pellerin, S.; Morneau, J. Evaluation of the operational 4D-Var at the meteorological service of Canada. *Geophys. Res. Abstr.* **2005**, *7*, 05567.
97. Wu, L.; Mallet, V.; Bocquet, M.; Sportisse, B. A comparison study of data assimilation algorithms for ozone forecasts. *J. Geophys. Res.* **2008**, *113*, D20310.
98. Geer, A.; et.al. The ASSET intercomparison of ozone analyses: Method and first results. *J. Atmos. Chem. Phys.* **2006**, *6*, 5445–5474.
99. Van Loon, M.; Builtjes, P.; Segers, A. Data assimilation of ozone in the atmospheric transport chemistry model LOTOS. *Environ. Model. Softw.* **2000**, *15*, 603–609.
100. Segers, A.; Heemink, A.; Verlaan, B.; van Loon, M. Modified RRSQRT-filter for assimilating data in atmospheric chemistry models. *Environ. Model. Softw.* **2000**, *15*, 663–671.
101. Hanea, R.; Velders, G.; Heemink, A. Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport model. *J. Geophys. Res.* **2004**, *109*, D10302.
102. Arulampalam, M.S.; Maskell, S.; Gordon, N.J.; Clapp, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **2002**, *150*, 174–188.
103. Kalman, R. A new approach to linear filtering and prediction problems. *Trans. ASME-J. Basic Eng.* **1960**, *82*, 35–45.
104. Gaspari, G.; Cohn, S. Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.* **1999**, *125*, 723–757.
105. Fisher, M. *Assimilation Techniques (5): Approximate Kalman Filters and Singular Vectors*; Meteorological Training Course Lecture Series, 2002. Available online: http://www.ecmwf.int/newsevents/training/lecture_notes/pdf_files/ASSIM/Kalman.pdf (accessed on 2 May 2011).
106. Houtekamer, P.; Mitchell, H. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Weather Rev.* **2001**, *129*, 123–137.
107. Bouttier, F.; Courtier, P. *Data Assimilation Concepts and Methods*; Meteorological Training Course Lecture Series, 1999. Available online: http://www.msi.ttu.ee/elken/Assim_concepts.pdf (accessed on 2 May 2011).
108. Burgers, G.; van Leeuwen, P.J.; Evensen, G. Analysis scheme in the ensemble Kalman filter. *Mon. Weather Rev.* **1998**, *126*, 1719–1724.
109. Evensen, G. The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dyn.* **2003**, *53*, 343–367.
110. Houtekamer, P.; Mitchell, H. Data assimilation using an ensemble Kalman filter technique. *Mon. Weather Rev.* **1998**, *126*, 796–811.
111. Barkmeijer, J.; van Gijzen, M.; Bouttier, F. Singular vectors and estimates of the analysis error covariance metric. *Q. J. R. Meteor. Soc.* **1998**, *124*, 1695–1713.
112. Anderson, J.L.; Anderson, S.L. A monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.* **1999**, *127*, 2741–2785.
113. Pham, D. Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Weather Rev.* **2001**, *129*, 1194–1207.
114. Harlim, J.; Hunt, B. Four-dimensional local ensemble transform Kalman filter: Numerical experiments with a global circulation mode. *Tellus* **2004**, *59A*, 731–748.

115. Evensen, G.; van Leeuwen, P. An ensemble Kalman smoother for nonlinear dynamics. *Mon. Weather Rev.* **2000**, *128*, 1852–1867.
116. Evensen, G. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dyn.* **2004**, *54*, 539–560.
117. Whitaker, J.; Hamill, T.M. Ensemble data assimilation without perturbed observations. *Mon. Weather Rev.* **2002**, *130*, 1913–1924.
118. Anderson, J.L. An ensemble adjustment kalman filter for data assimilation. *Mon. Wea. Rev.* **2001**, *129*, 2884–2903.
119. Heemink, A.; Verlaan, M.; Segers, A.J. Variance reduced ensemble Kalman filtering. *Mon. Weather Rev.* **2001**, *129*, 1718–1728.
120. Bishop, C.; Etherton, B.; Majumdar, S. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Weather Rev.* **2001**, *129*, 420–436.
121. Sekiyama, T.T.; Tanaka, T.Y.; Shimizu, A.; Miyoshi, T. Data assimilation of CALIPSO aerosol observations. *Atmos. Chem. Phys.* **2010**, *10*, 39–49.
122. Sekiyama, T.T.; Tanaka, T.Y.; Maki, T.; Mikami, M. The effects of snow cover and soil moisture on asian dust: II. Emission estimation by lidar data assimilation. *Sci. Online Lett. Atmos.* **2011**, *7A*, 40–43.
123. Sekiyama, T.T.; Deushi, M.; Miyoshi, T. Operation-oriented ensemble data assimilation of total column ozone. *Sci. Online Lett. Atmos.* **2011**, *7*, 41–44.
124. Schutgens, N.A.J.; Miyoshi, T.; Takemura, T.; Nakajima, T. Applying an ensemble Kalman filter to the assimilation of AERONET observations in a global aerosol transport model. *Atmos. Chem. Phys.* **2010**, *10*, 2561–2576.
125. Courtier, P.; Talagrand, O. Variational assimilation of meteorological observations with the adjoint equations. Part 2: Numerical results. *Q. J. R. Meteorol. Soc.* **1987**, *113*, 1329–1347.
126. Le-Dimet, F.; Talagrand, O. Variational algorithms for analysis and assimilation of meteorological observations. *Tellus A* **1986**, *38*, 97–110.
127. Lions, J.L. *Optimal Control of Systems Governed by Partial Differential Equations*; Springer Verlag: Berlin, Germany, 1971.
128. Courtier, P. Dual formulation of four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.* **1997**, *123*, 2449–2462.
129. Bennett, A. *Inverse Modeling of the Ocean and Atmosphere*; Cambridge University Press: Cambridge, UK, 2002.
130. J, M.L.; Lakshmivarahan, S.; Dhall, S. *Dynamic Data Assimilation-A Least Squares Approach*; Cambridge University Press: Cambridge, UK, 2005.
131. Fisher, M.; Nocedal, J.; Tremolet, Y.; Wright, S.J. Data assimilation in weather forecasting: A case study in pde-constrained optimization. *Optim. Eng.* **2009**, *10*, 409–426.
132. Lorenc, A.C. The potential of the ensemble Kalman filter for NWP-A comparison with the 4D-VAR. *Q. J. R. Meteorol. Soc.* **2003**, *129*, 3183–3203.
133. Kalnay, E.; Li, H.; Miyoshi, T.; Yang, S.; Ballabrera-Poy, J. 4D-Var or ensemble Kalman filter. *Tellus A* **2007**, *59*, 758–773.

134. Hamill, T. *Ensemble-Based Atmospheric Data Assimilation*; University of Colorado and NOAA-CIRES Climate Diagnostics Center: Boulder, Colorado, USA, 2004.
135. Sandu, A.; Cheng, H. A subspace approach to data assimilation and new opportunities for hybridization. *Phys. D* **2011**, submitted for publication.
136. Fletcher, S.; Zupanski, M. A study of ensemble size and shallow water dynamics with the maximum likelihood ensemble filter. *Tellus A* **2008**, *60*, 348–360.
137. *The Integrated Global Atmospheric Chemistry Observations (IGACO)-WMO*; Technical Report: WMO TD No. 1235; World Meteorological Organization: Geneva, Switzerland, 2004.
138. Miehe, P.; Sandu, A. Forward, tangent linear, and adjoint Runge Kutta methods in KPP-2.2. *LNCS* **2006**, *3991*, 120–127.
139. Alexe, M.; Sandu, A. An Investigation of Discrete Adjoints for Flux-Limited Numerical Schemes. In *Proceedings of the 45th Annual Southeast Regional Conference*; Winston-Salem, NC, USA, 23–24 March 2007; pp. 373–378.
140. Giering, R. *Tangent Linear and Adjoint Model Compiler, Users Manual 1.4*; Massachusetts Institute of Technology: Cambridge, MA, USA, 1999.
141. Hascöet, L.; Pascual, V. *TAPENADE 2.1 User's Guide*. Technical Report 0300; INRIA: Sophia Antipolis, France, 2004.
142. Sandu, A. *On Consistency Properties of Discrete Adjoint Linear Multistep Methods*. Technical Report CS-TR-07-40; Computer Science Department, Virginia Tech: Blacksburg, VA, USA, 2007.
143. Liu, Z.; Sandu, A. Analysis of discrete adjoints of numerical methods for the advection equation. *Int. J. Numer. Methods Fluids* **2008**, *56*, 769–803.
144. Sandu, A., Reverse automatic differentiation of linear multistep methods. In *Advances in Automatic Differentiation*; Bischof, C., Bucker, H., Hovland, P., Naumann, U., Utke, J., Eds.; Springer: Berlin, Germany, 2008; pp. 1–12.
145. Sandu, A. Chemical Data Assimilation: Computational Tools and Applications. In *Proceedings of the 8th International Conference on HydroScience and Engineering (ICHE 2008)*, Nagoya, Japan, 9–12 September 2008.
146. Alexe, M.; Sandu, A. Forward and adjoint sensitivity analysis with continuous explicit Runge-Kutta schemes. *Appl. Math. Comput.* **2009**, *208*, 328–334.
147. Alexe, M.; Sandu, A. On the discrete adjoints of variable step time integrators. *J. Comput. Appl. Math.* **2009**, *233*, 1005–1020.
148. Sandu, A. Solution of inverse ODE problems using discrete adjoints. In *Large Scale Inverse Problems and Quantification of Uncertainty*; Biegler, L., Biros, G., Ghattas, O., Heinkenschloss, M., Keyes, D., Mallick, B., Tenorio, L., van Bloemen Waanders, B., Willcox, K., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2010; Chapter 12, pp. 345–364.
149. Alexe, M.; Sandu, A. Space-time adaptive solution of inverse problems with the discrete adjoint method. *Inverse Probl.* **2011**, submitted for publication.
150. Singh, K.; Sandu, A. Variational chemical data assimilation with approximate adjoints. *Comput. Geosci.* **2011**, in press.

151. Sandu, A.; Miehe, P. Forward, tangent linear, and adjoint Runge Kutta methods in KPP—2.2 for efficient chemical kinetic simulations. *Int. J. Comput. Math.* **2010**, *87*, 2458–2479.
152. Zhang, H.; Sandu, A. FATODE: A Library for Forward, Adjoint and Tangent Linear Integration of Stiff Systems. In *Proceedings of the 2011 Spring Simulation Multiconference (SpringSim'11), High Performance Computing Symposium (HPC-2011)*, Boston, MA, USA 3–7 April 2011; pp. 152–159.
153. Eller, P.; Singh, K.; Sandu, A.; Bowman, K.; Henze, D.; Lee, M. Implementation and evaluation of an array of chemical solvers in a global chemical transport model. *Geophys. Model Dev.* **2009**, *2*, 1–7.
154. Singh, K.; Eller, P.; Sandu, A.; Henze, D.; Bowman, K.; Kopacz, M.; Lee, M. Towards the Construction of a Standard Adjoint GEOS-Chem Model. In *Proceedings of the 2009 Spring Simulation Multiconference (SpringSim'09), High Performance Computing Symposium (HPC-2009)*, Ribbens, C., Sandu, A., Thacker, W., Eds.; San Diego, CA, USA, 22–27 March 2009; p. 8.
155. Daescu, D. On the sensitivity equations of four-dimensional variational (4D-Var) data assimilation. *Mon. Weather Rev.* **2008**, *136*, 3050–3065.
156. Constantinescu, E.; Sandu, A.; Chai, T.; Carmichael, G. Autoregressive models of background errors for chemical data assimilation. *J. Geophys. Res.* **2007**, *112*, D12309.
157. Hasselmann, K. Stochastic climate models. Part I. Theory. *Tellus* **1976**, *28*, 473–484.
158. Singh, K.; Jardak, M.; Sandu, A.; Bowman, K.; Lee, M.; Jones, D. Construction of non-diagonal background error covariance matrices for global chemical data assimilation. *Geosci. Model Dev.* **2011**, *4*, 299–314.
159. Cheng, H.; Jardak, M.; Alexe, M.; Sandu, A. A hybrid approach to estimating error covariances in variational data assimilation. *Tellus A* **2010**, *62*, 288–297.
160. Sandu, A.; Zhang, L. Discrete second order adjoints in atmospheric chemical transport modeling. *J. Comput. Phys.* **2008**, *227*, 5949–5983.
161. Cioaca, A.; Alexe, M.; Sandu, A. Second order adjoints for solving PDE-constrained optimization problems. *Optim. Methods Softw.* **2011**, submitted for publication.
162. Le Dimet, F.X.; Shutyaev, V.; Gejadze, J. Analysis error via hessian in variational data assimilation. *ARIMA J.* **2006**. Available online: <http://www.cari-info.org/actes2006/58.pdf> (accessed on 2 May 2011).
163. Gery, M.W.; Whitten, G.Z.; Killus, J.P.; Dodge, M.C. A photochemical kinetics mechanism for urban and regional scale computer modeling. *J. Geophys. Res.* **1989**, *94*, 12925–12956.
164. CMAQ v4.6 Operational Guidance Document; 2007. Available online: <http://www.cmaq-model.org> (accessed on 2 May 2011).
165. Sarwar, G.; Luecken, D.; Yarwood, G.; Whitten, G.; Carter, W.P. Impact of an updated carbon bond mechanism on predictions from the CMAQ modeling system: Preliminary assessment. *J. Appl. Meteor. Climatol.* **2008**, *47*, 3–14.
166. Otte, T.L.; Pouliot, G.; Pleim, J.E.; Young, J.O.; Schere, K.L.; Wong, D.C.; Lee, P.C.S.; Tsidulko, M.; McQueen, J.T.; Davidson, P.; *et al.* Linking the Eta Model with the community

- multiscale air quality (CMAQ) modeling system to build a national air quality forecasting system. *Weather Forecast.* **2005**, *20*, 367–384.
167. Zhu, C.; Byrd, R.H.; Nocedal, J. L-BFGS-B–FORTRAN routines for large scale bound constrained optimization. *ACM Trans. Math. Softw.* **1997**, *23*, 550–560.
 168. Byrd, R.; Lu, P.; Nocedal, J. A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* **1995**, *16*, 1190–1208.
 169. Horn, R.A.; Johnson, C.R. *Topics in Matrix Analysis, Chapter 4*; Cambridge University Press: Cambridge, UK, 1991.
 170. Gorline, J.L.; Lee, P. Performance evaluation of NOAA-EPA developmental aerosol forecasts. *Environ. Fluid Mech.* **2009**, *9*, 109–120.
 171. Mebust, M.; Eder, B.; Binkowski, F.; Roselle, S. Models-3 community multiscale air quality (CMAQ) model aerosol component-2. Model evaluation. *J. Geophys. Res.* **2003**, *108*, 4184; doi:10.1029/2001JD001410.
 172. Eder, B.; Yu, S. A performance evaluation of the national air quality forecast capability for the summer of 2007. *Atmos. Environ.* **2009**, *43*, 2312–2320.

© 2011 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>.)