Adjoint Inverse Modeling of Black Carbon During ACE-Asia

A. Hakami 1, D. K. Henze 1, J. H. Seinfeld 1, T. Chai 2, Y. Tang 2, G. R. Carmichael 2, A. Sandu 3

1 Departments of Chemical Engineering and Environmental Science and Engineering, California Institute of Technology, Pasadena, California.
2 Center for Global and Regional Environmental Research, University of Iowa, Iowa City, Iowa.
3 Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Abstract: An adjoint model is used for inverse modeling of black carbon during ACE-Asia. We use the 4D-Var approach to optimally recover spatially-resolved anthropogenic and biomass burning emissions, and initial and boundary conditions of black carbon. Boundary conditions and biomass burning emissions are assigned daily scaling factors. Anthropogenic emissions are scaled by a combination of daily and monthly scaling factors. Simulation results are compared to various observations of black carbon concentrations during the campaign. Measurements at five islands and onboard the research vessel Ron Brown are used for data assimilation. Different levels of constraints are examined for data assimilation, and a case with 65% reduction in the total square errors is chosen. The assimilation is more successful at the islands of Sado and Gosan and for the observations onboard R/V Ron Brown. The assimilated results show significantly better agreement with the observations onboard the Twin Otter aircraft that were not used for assimilation. Among the scaled variables, anthropogenic emissions are the most significant, followed by the boundary conditions. Domain-wide emissions inventory does not change significantly as a result of the scaling, but sizable changes are seen on the sub-regional level. Most noticeably, anthropogenic emissions over southeastern China are reduced while those in northeast China and Japan are increased.
Introduction

Black (or elemental) carbon (BC) is the main light absorbing aerosol species; it alters the radiative properties of other aerosols with which it is mixed, and it may affect cloud formation and precipitation. In short, understanding the regional and global distributions of BC is key to predicting the effect of aerosols on global and regional climate. A number of investigators have simulated top of the atmosphere (TOA) direct radiative forcing of black carbon, values of which range from approximately +0.1 to +0.8 W/m$^2$ [Haywood and Shine, 1995; Haywood et al.; 1997, Haywood and Ramaswamy, 1998; Penner et al., 1998; Myhre et al., 1998; Cooke et al., 1999; Jacobson, 2001, 2002; Koch, 2001; Chung and Seinfeld, 2002; Wang, 2004]. The wide range of estimates is primarily a result of different assumptions about the mixing state of BC with sulfate aerosols and of the global burden of BC. A significant source of uncertainty in the estimates is the BC emissions inventory itself. Global and regional emission inventories of BC have been refined [Cooke and Wilson, 1996; Liousse et al., 1996; Cooke et al., 1999; Streets et al., 2003; Bond et al., 2004; Schaap et al., 2004]; nevertheless, emissions uncertainty remains a significant contributor to the overall uncertainty in predicted global BC distributions. These uncertainties play an even more significant role in regional/episodic simulations, where the actual emissions are more likely to deviate from the annual mean inventory. Regional studies present a particularly advantageous opportunity to evaluate our understanding of the relationship between observed BC levels and existing emissions inventories. Furthermore, owing to its relatively short atmospheric lifetime (4-10 days), BC emissions have their most pronounced effect on the regional scale.
Inverse modeling provides a powerful approach for observation-based inference about atmospheric model inputs (e.g., emissions). Methods based on the Kalman filter and its variations assume an imperfect (noisy) model, and are able to propagate the error characteristics through the formalism of the method. Kalman filtering has been widely applied to meteorological and atmospheric data assimilation problems [Menard et al., 2000; Khattatov et al., 2000; Stajner et al., 2001; Palmer et al., 2003]. Methods based on the Kalman filter tend to be computationally demanding and, therefore, have seen limited use in large-scale atmospheric chemical transport models (CTMs) [Mulholland and Seinfeld, 1995; Houtekamer and Mitchell, 1998; van Loon et al., 2000]. Alternatively, adjoint modeling (Marchuk, 1974, 1986; Caccuci, 1981) can be used to calculate gradients of an objective function with respect to model input parameters. In adjoint analysis, the model is considered as a strong constraint for the problem, i.e., the model is assumed to be perfect. Adjoint data assimilation [Talagrand, 1981a,b; Le Dimet and Talagrand, 1986; Talagrand and Coutier, 1987] has been used in meteorology and oceanography [Courtier and Talagrand, 1987; Navon, 1997; Usbeck et al., 2003].

In the context of three-dimensional (3-D) Eulerian CTMs, adjoint modeling offers an efficient method for data assimilation applications, as desired gradients of the objective function are calculated simultaneously. Elbern and Schmidt [1999] applied the adjoint method to the European air pollution dispersion chemical transport model 2 (EURAD-CTM2) for data assimilation. They present a variety of identical twin experiments to verify the adjoint implementation; most notably they were able to recover initial concentrations of NO\textsubscript{x} and VOC from ground level ozone observations. In an
identical twin experiment, the simulation results for a perturbed set of parameters are used as observations, and the adjoint analysis is used to recover the applied changes in parameters. The advantage of the identical twin experiment is that the true answer to the inverse problem is known. Elbern et al. [2000] used the same model and methodology for recovery of $\text{NO}_x$ and VOC emission rates from ozone observations. Vukicevic and Hess [2000] implemented adjoint modeling in HANK, a chemical transport model based on MM5 meteorological outputs. The sensitivity of hypothetical soluble and insoluble species concentrations at Hawaii with respect to a variety of different model parameters were calculated. Vautard et al. [2000] apply an adjoint version of the chemical transport model CHIMERE [Menut et al., 2000] using ozone observations for recovery of urban boundary ozone values. Elbern and Schmidt [2001] used EURAD-CTM2 and applied 4D-Var data assimilation to an ozone episode over central Europe to optimize various initial concentrations. Schmidt and Martin [2003] and Menut [2003] applied the adjoint technique in CHIMERE for episodic sensitivity analysis. Sandu et al. [2004] (“Adjoint sensitivity analysis of regional air quality models”, submitted to J. Comput. Physics, hereafter referred to as Sandu et al., 2004) formulate continuous and discrete adjoints for implementation in chemical kinetic systems [Sandu et al., 2003] and 3-D air quality models. They apply the method in the chemical transport model STEM-2k1 [Carmichael et al., 2003] for sensitivity analysis and recovery of various initial conditions in identical twin experiments.

Here, we adapt an adjoint model of the chemical transport model STEM-2k1 [Carmichael et al., 2003; Sandu et al., 2004] for data assimilation and recovery of BC
emissions (anthropogenic and biomass burning), boundary conditions, and initial
conditions in eastern Asia during the Asian-Pacific Regional Aerosol Characterization
Experiment (ACE-Asia) field study [Huebert et al., 2003; Seinfeld et al., 2004]. East and
Southeast Asia are major contributors to the total global BC burden [Bond et al., 2004],
and the ACE-Asia campaign provides a unique opportunity to better constrain these
emissions. Our goal is to employ BC measurements carried out during ACE-Asia (April
2001) to optimally estimate east Asian BC emissions (and initial and boundary
conditions). The long assimilation window is essential for regional inverse modeling,
where long-range transport of pollutants becomes important. Data assimilation is based
on direct observations of BC aerosol mass concentrations in four Japanese islands along
~140°E longitude, Gosan, South Korea, and onboard R/V Ron Brown.

**Adjoint Formulation**

The equation governing the dynamics of a chemically non-reactive species (e.g., BC) in
atmospheric CTMs is [Seinfeld and Pandis, 1998; Jacobson, 1998],

\[
\frac{\partial C}{\partial t} = -\nabla \cdot (u C) + \nabla \cdot (K \nabla C) + E - k_w C
\]

where \( C \) is the concentration of the species, \( u \) is the vector wind field, and \( K \) is the
diffusivity tensor. \( E \) represents elevated emissions, and \( k_w \) is the first-order rate constant
for wet removal of the species. Equation (1), which in inverse modeling is referred to as
the forward model, is solved subject to specific initial and boundary conditions; while the
formulation of these conditions may vary slightly from one CTM to another, they
invariably contain similar key input parameters. For STEM, the initial conditions, surface boundary condition, and lateral inflow boundary conditions are \cite{Sandu et al., 2004},

\begin{equation}
C(\omega, t_0) = C^0(\omega)
\end{equation}

\begin{equation}
C(\omega, t)\bigg|_{\omega \in \Gamma^{in}} = C^b(\omega, t)
\end{equation}

\begin{equation}
K \frac{\partial C}{\partial \omega} \bigg|_{\omega \in \Gamma^{out}} = 0
\end{equation}

\begin{equation}
\left( K \frac{\partial C}{\partial \omega} - \gamma_d C + E^0 \right) \bigg|_{\omega \in \Gamma^{out}} = 0
\end{equation}

where $\omega$ is a generalized spatial coordinate vector, $C^0$ and $C^b$ are the initial and boundary concentrations, respectively, and $\Gamma$ represents the boundary cells (inflow, outflow, or ground level); $\gamma_d$ is the species dry deposition velocity, and $E^0$ is its surface-level emission rate.

In adjoint sensitivity analysis, the gradients of a scalar function $J$ with respect to a set of input parameters are calculated. For data assimilation applications, this scalar takes the form of an objective (cost) function, and is generally defined as,

\begin{equation}
J(C, \alpha) = \frac{1}{2} \left[ \frac{1}{\mu} \int_{t_0}^{t_f} (\dot{\alpha}^b - \alpha)^T N^{-1} (\dot{\alpha}^b - \alpha) + \int_{t_0}^{t_f} (\hat{C} - C)^T R^{-1} (\hat{C} - C) \right]
\end{equation}

where $\hat{C}$ and $C$ are vectors of the observed and simulated concentrations (temporal and spatial dependences are omitted for simplicity), respectively. $t_0$ and $t_f$ are the initial and final times of the simulation; $\alpha$ denotes the set of input parameters that are to be optimally estimated in the data assimilation, and $\alpha_0$ is the vector of initial (background)
estimates for those parameters. Uncertainty in the data and inputs is incorporated into the objective function through the error covariance matrices for the observations and inputs, N and R, respectively. In so doing, observations or a-priori estimates with high uncertainty have a lower contribution to the overall objective function. \( \mu \) is a conversion factor with appropriate units (depending on the type and units of the input parameters) that can also be used as a global weighting factor for assigning relative emphases on either the observations or background values. The objective function is evaluated for grid cells and times where observations are available. For uncorrelated error and discrete temporal and spatial observations, the scalar objective function can be written as,

\[
J = \frac{1}{2} \left[ \frac{1}{\mu} \sum_k N_k^{-1}(\alpha^b_k - \alpha_k)^2 \right] + \frac{1}{2} \left[ \sum_i \sum_j R_{ij}^{-1}(\hat{C}_{ij} - C_{ij})^2 \right] \tag{4}
\]

where indices \( i \) and \( j \) represent the time and location of the observations, and index \( k \) represents the particular parameter to be estimated.

A perturbation in the input parameters, \( \delta a \), translates into perturbations in the simulated concentrations and the cost function, \( \delta C \) and \( \delta J \). Applying a Lagrange multiplier \( \lambda \) to the perturbed form of Equation (3) (and assuming uncorrelated observation errors) results in \cite{Marchuk, 1986; Vukicevic and Hess, 2000; Elbern et al., 2000; Sandu et al., 2004],

\[
\delta J = \frac{1}{\mu} \int \left( a - a^b \right)^T N^{-1} \delta a \, dt + \int \int_{\Omega} \phi(\omega, t) \delta C \, d\omega \, dt
\]

\[
- \int \int_{\Omega} \lambda \left( \frac{\partial C}{\partial t} + \nabla \cdot (u \delta C) - \nabla \cdot (K \nabla \delta C) - \delta E + k \delta C \right) \, d\omega \, dt \tag{5}
\]
where \( \phi(\omega, t) = \left[ (C - \hat{C})/R(\omega, t) \right] \) is the observation-driven forcing term for the adjoint system. After integration by parts and rearrangement, the following equation can be used to calculate the gradients of the cost function,

\[
\delta J = \frac{1}{\mu} \int_t^t \left( \mathbf{a} - \mathbf{a}_b \right)^T N^{-1} \delta \mathbf{a} \, dt + \int_t^t \left( \mathbf{u} \lambda \bigg|_{\partial \Omega} + K \frac{\partial \lambda}{\partial \omega} \bigg|_{\partial \Omega} \right) \delta \mathbf{a}_b \, dt \\
+ \int_t^t \lambda \delta \mathbf{a}_E \, dt + \int_\Omega \int_\Omega \lambda \delta \mathbf{a}_E \, d\omega \, dt + \int_\Omega \left( \lambda \bigg|_{t=0} \right) \delta \mathbf{a}_S \, d\omega
\]

subject to the following conditions,

\[
\begin{align*}
\lambda(t_f) &= 0 \\
\lambda \bigg|_{\partial \Omega} &= 0 \\
K \frac{\delta \lambda}{\delta \omega} \bigg|_{\partial \Omega} &= 0 \\
\left( K \frac{\partial \lambda}{\partial \omega} - \nu \lambda \right) \bigg|_{\partial \Omega} &= 0
\end{align*}
\]

\( \mathbf{a}_1, \mathbf{a}_b, \mathbf{a}_E, \) and \( \mathbf{a}_E^* \) in Equation (6) represent the initial conditions, boundary conditions, elevated emissions, and surface emissions, respectively. Note that the adjoint equation is driven by differences between observations and simulations (the forcing term
The negative sign for the time derivative term in Equation (7) indicates that the adjoint equation is integrated backward in time from \( t_F \) to \( t_0 \). Also note that a reverse wind field is applied to the transport of the adjoints, and therefore, an inflow boundary for the concentration corresponds to an outflow boundary for the adjoint.

Equation (7) and the initial and boundary conditions in Equation (8) uniquely define the solution for the adjoint system. Once the adjoint is integrated backward in time, the gradients of the cost function can be calculated based on Equation (6). These equations are derived for the continuous Equation (1) and must be solved numerically. Alternatively, one can derive the adjoint equation for the numerical (discretized) solution to Equation (1). The two alternatives are not exactly equivalent, as the adjoint and discretization operations are not generally commutable [Sirkes and Tziperman, 1997; Sandu et al., 2003]. The discrete gradient equation based on numerical algorithms used in STEM [Sandu et al., 2004] can be written as,

\[
\delta J = \frac{1}{\mu} \sum_t \sum_{k} f_k \left( f_k \frac{1}{N_k} \right) \delta E_k(t)
+ \sum_{\Omega} \left[ C^0(\omega) \lambda(\omega, t_0) \right] \delta E_i(t)
+ \sum_{t} \sum_{i} \left[ E^0(\omega_i, t) \lambda(\omega_i, t) \Delta t \right] \delta E_{e^0}(\omega_i, t)
+ \sum_{t} \sum_{\Omega} \left[ E(\omega, t) \lambda(\omega, t) \Delta t \right] \delta E_{E}(\omega, t)
+ \sum_{t} \sum_{i} \left[ C^b(\omega, t) \left[ \left( \frac{u(\omega, t)}{\Delta \omega} + \frac{K(\omega, t)}{\Delta \omega^2} \right) \lambda(\omega, t) - \frac{u(\omega, t)}{6 \Delta \omega} \lambda(\omega, t) \right] \Delta t \right] \delta E_{b}(\omega, t)
\]
where $\varepsilon$ represents an individual dimensionless scaling factor applied to each of the input parameters (distinguished by the subscripts), i.e., $\alpha_k = \varepsilon_k \alpha_k^b$. The function $f_k$ is defined as $f_k = \max(\varepsilon_k, 1/\varepsilon_k)$ and $f_k'$ is its derivative with respect to the individual scaling factor. This form of background-driven cost is used to assign equivalent penalties to scaling up or down from the background values. Superscripts $b_1$, $b_1$, and $b_2$ show the outside, first, and second stripe of interior boundary cells, respectively. For horizontal advection, a third-order upwind numerical scheme is used, hence the term for the second interior stripe in the boundary condition gradient. During the assimilation process each of the scaling factors (all of which have an initial value of one) is optimally estimated. By employing the scaling factors, the optimization process focuses on the input parameters with larger magnitudes. An important advantage of using scaling factors is that input parameters with a magnitude of zero will not play a role in the optimization, e.g. a grid cell with zero emissions (over the ocean) cannot have any emissions assigned to it as a result of the optimization.

**Inverse Modeling of Black Carbon over East Asia**

The objective of data assimilation is to estimate the input parameters in order to achieve optimized model performance, i.e., to minimize the cost function. In the larger scheme, it is desirable to identify any existing biases and systematic errors in the input estimates. CTMs and GCMs require a multitude of inputs, many of which are highly uncertain. Furthermore, most of such inputs vary in space and/or time. As a result, a large number of input parameters can be estimated in the assimilation process. As the available
observations are usually scarce, atmospheric data assimilation is often an extremely under-determined problem.

Equation (9) [or (6) for continuous form] allows for calculation of the gradient of the cost function with respect to any input parameter at any location and/or time. The 4-D adjoint analysis results in spatial and temporal distributions of the gradients of the objective function. These gradients directly link the objective function to the input parameters at different times and locations, taking into account all physical processes that are included in the model. In other words, adjoint results can be easily used to objectively identify locations and times that most affect model performance.

The CTM used in this study is the adjoint version of STEM-2k1 [Sandu et al., 2004; Carmichael et al., 2003], a parallel implementation of STEM using the communication and parallelization library of PAQMSG [Miehe et al., 2002]. The model utilizes an operator-splitting scheme, and in the parallel mode breaks the computational domain into horizontal and vertical slices to be sent to available processors. For each time step, the required data (emissions, wind field, etc.) are gathered and processed by the master node, while computations are distributed among the worker nodes. The model uses efficient two-level checkpointing storage, as the state vector (concentrations) is required for the integration of the adjoint in nonlinear processes. However, as BC is chemically non-reactive, chemistry is not active and no checkpointing is required.
The modeling domain for this study extends approximately between 10° N, 50° N, 90° E, and 155° E, as shown in Figure 1. The computational domain consists of 90 x 60 x 18 grid cells. Horizontal grid resolution is 80 km, and variable vertical spacing follows the topography of the terrain. The simulation period is the month of April 2001. Adjoint analysis and scaling is conducted for the month-long assimilation window. The meteorological fields (winds, turbulence, precipitation, etc.) for STEM are produced by the mesoscale meteorological model RAMS (Pielke et al., 1992). Base case boundary conditions are taken to be time-invariant and for each altitude are set to the 5th percentile of all the observations during the TRACE-P measurements (Carmichael et al., 2003). Boundary conditions for all lateral boundaries in each vertical layer are set to the same value, and the model does not account for time- and location-specific variabilities in lateral boundary conditions. Initial conditions are calculated from the simulations of the previous month (March 2001).

BC lifetime and concentrations can be strongly affected by precipitation if the BC occurs internally mixed with soluble species. The first-order wet removal constant is assumed to depend on the precipitation rate via the following empirical relation [Uno et al., 2003], \( k_w = 10^{-5} h^{0.88} \), where \( k_w \) is the first-order removal rate constant (s\(^{-1}\)), and \( h \) is the precipitation rate, in mmh\(^{-1}\). The constant (10\(^{-5}\)) is chosen based on forward sensitivity analysis with respect to the magnitude of the constant such that the model over-predictions at VMAP stations during the precipitation episodes are minimized, and it corresponds to \(~20\%\) of the value used for the wet removal of sulfate aerosols [Uno et al., 2003].
The basic gridded BC emission inventory is based on the regional gaseous and aerosol primary emissions inventory for the year 2000 of Streets et al. [2003a,b], specifically developed for support of the intensive field studies in the region, TRACE-P [Jacob et al., 2003] and ACE-Asia. The inventory was prepared using a bottom-up approach and is based on energy consumption and activity information for various emission sectors: industrial, residential, transportation, power generation, agricultural, biomass burning (BB), and others. The biomass burning inventory (assumed to be solely anthropogenic) was prepared using vegetation cover maps, satellite fire count data, and a variety of other open fire information. The overall uncertainty in BC emissions (BB and other anthropogenic sources) is considered to be one of the largest among different species represented in the inventory. In different regions, total anthropogenic emissions and BB have a range of uncertainty (95% confidence interval) of 80-490 %, and 200-700 %, respectively. Over all of Asia, the uncertainties in anthropogenic and BB emissions are estimated at about 360 and 450 %, respectively.

From an implementation point of view, BC emissions are separated into BB and (other) anthropogenic emissions. For biomass burning, the files contain the spatial distribution of the daily emissions. It is assumed that the rate of emission for each grid cell remains constant throughout the day. These daily values are then distributed into different vertical layers. The anthropogenic emissions are aggregated from different emission sectors (except BB) into one file. This input file contains gridded BC diurnal emissions (18 vertical layers) for a typical day in April 2001. Therefore, the available
emissions do not provide information on day-specific variability for anthropogenic BC, or on time-specific variability for BB emissions of BC. Furthermore, BB emission inputs do not account for location-specific variability in vertical profile of BC emissions. These limitations (in addition to that pertaining to boundary conditions) are relaxed in the time- and location-dependent scaling of input parameters through adjoint data assimilation.

Four different types of input parameters are scaled for data assimilation: anthropogenic BC emissions (first 4 vertical layers, time-dependent), biomass burning BC emissions (first 12 vertical layers, time-dependent), lateral boundary conditions (18 layers, time-dependent), and initial conditions (18 layers). The scaling factors for biomass BC emissions and boundary BC concentrations are considered as daily averages, and their corresponding gradients are integrated over each day. Scaling factors for initial BC concentration are time independent. For anthropogenic BC emissions, the scaling factor is assumed to include a monthly portion (the gradient of which is integrated over the entire month), and a daily portion that is added to the monthly scaling factor. The monthly and daily portions are assigned initial values of one and zero, respectively. The daily portion allows for limited day-to-day variability in the emissions for each grid cell. Data assimilation and parameter optimization is an iterative procedure. A quasi-Newton limited memory optimization routine, L-BFGS [Byrd et al., 1995], is used for optimization after each iteration. 4-D (time and space) data assimilation results in a field of scaling parameters (3-D in case of initial conditions, and monthly anthropogenic BC emissions) following each iteration.
Black carbon observations during ACE-Asia

During the ACE-Asia campaign, an extensive network of aircraft, shipboard, and surface instruments was utilized to measure different characteristics of the regional aerosols and their gaseous precursors. Measurements included mass concentrations of BC, as well as frequent measurements of aerosol optical (absorption) properties were carried out, from which BC mass concentrations can be estimated.

The success of any data assimilation depends critically on the amount of data available for analysis. Since the three-dimensional adjoint equations are driven by the discrepancy between observations and simulation, greater spatial coverage (i.e., more observation sites) will significantly enhance the ability of the adjoints to capture the areas of influence on the objective function. In other words, comprehensive temporal and spatial coverage of observations provides the best opportunity for high quality data assimilation. Whereas it was deemed important to include all the applicable BC observations in the analysis, it is also necessary to assign some measure of reliability to the available observations. Three main factors can contribute to the degree of importance of the observational data sets for 4-D variational data assimilation:

1. Low uncertainty: In addition to the obvious case of measurement/sampling uncertainty, particular attention should be paid to the representativeness uncertainty, i.e., the uncertainty in representing a grid cell or computational node by a single point measurement.

2. High time resolution: short-term observations produce more of an impulse perturbation in the adjoint variables than those that are averaged over long time.
In the latter case, under- and over-predictions may cancel each other and result in loss of information in the adjoints. It should be noted, however, that very low sampling frequency (i.e., grab sampling) usually results in increased uncertainty.

3. Added spatial coverage: As mentioned above, more complete spatial coverage results in better representation of the adjoints, and in more information about the areas of influence. Therefore, a low quality data set in an otherwise under-sampled location may be more important than higher quality, but spatially redundant, observations. In addition, proximity of the observations to the more important parameters (e.g., emissions in a particular region) can be a significant advantage.

For the current study, short-term BC mass measurements are the most suitable type of observations for inverse modeling. Mass concentrations deduced from optical measurements are inherently more uncertain, due to variability in the specific mass absorption efficiency [Chuang et al., 2003; Clarke et al., 2004], and are therefore considered less reliable and are not used in this study. The available BC measurements that are considered for this study are summarized in Table 1.

The Variation of Marine Aerosol Properties (VMAP) network was designed for the estimation of latitudinal gradients of aerosols along the ~140° E longitude, and provides the largest number of BC mass observations for the ACE-Asia. Another important set of observations are daily measurements at Gosan, as they were cross-examined with independent measurements at the station. The observations onboard the research vessel
(R/V) Ron Brown are also of particular interest, as they are sampled over a short period of time and are unaffected by direct emission sources.

We compare all the observations listed in Table 1 with the model predictions. BC measurements at Gosan, VMAP network, and onboard R/V Ron Brown are used for data assimilation. Data collected at Kwangju and Yulin seem to be too closely affected by the local sources to be properly represented by the coarse grid size employed in this study. For airborne measurements that span over multiple grid cells during the course of one sampling period, the distribution of the forcing term in the adjoint equation among those grid cells becomes an additional source of uncertainty. Observations made onboard the C-130 and Twin Otter aircraft are, therefore, used for independent evaluation of the inverse modeling and optimization.

Results and Discussion

STEM-2k1 is applied to the ACE-Asia domain as the forward model for simulation of BC concentrations and adjoint values during the month of April 2001. Figure 2 shows an example of the predicted spatial distribution of ground-level BC at the time of the maximum simulated concentration during the month of April 2001. Ground-level BC concentrations have a range of 0-5.5 $\mu$g/m$^3$ during the month, and highest concentrations are simulated in the proximity of the major emission areas of western India and eastern China. Highest BC concentrations coincide with episodes of significant BB; downwind of major emissions more moderate concentrations are simulated.
The time series of simulated BC concentrations based on the starting emission inventory are compared to the VMAP observations at the four Japanese islands of Chichi-jima, Hachijo, Sado, and Rishiri. In general, the simulations capture the overall behavior and variability of the observations; however, significant differences exist between the simulated and observed concentrations. In particular, there are significant under-predictions at Rishiri and Sado. Similar under-predictions are reported by Uno et al. [2003], who suggest under-estimates of Japanese biomass burning emissions before the rice-planting season as a possible explanation. In their work, the same underlying meteorological model (RAMS) [Pielke et al., 1992] and emission inventory [Streets et al., 2003] was used, but the model did not include wet removal and assumed zero inflow boundary conditions. The effect of including wet removal in the simulations is also shown in Figure 3. Without wet removal the model significantly over-estimates BC concentrations at Chichijima and Hachijo. Note that the northern islands of Rishiri and Sado rarely experience heavy precipitation episodes, and therefore, inclusion of wet removal has little effect on the BC concentration (and over-estimation) at those stations.

VMAP BC measurements provide important information with high temporal resolution and consistency across the stations, and therefore, form the basis for the current inverse modeling. Figure 4 shows simulated and observed BC concentrations at a few other stations during the ACE-Asia campaign, namely Gosan [Chuang et al., 2003; Schauer et al., 2003], Kwangju [Kim et al., 2004], and Yulin [Xu et al., 2004], as well as aboard R/V Ron Brown [Lim et al., 2003]. At Gosan, the simulated BC concentrations are generally lower than the observations. The most significant differences are observed
during the periods of yellow sand dust storms (April 10-13 and 24-27). Unexpectedly, during these dust events a large fraction of BC is present in the coarse mode [Chuang et al., 2003]. Chuang et al. [2003] associate the coarse mode BC with the coagulation of fine aerosols onto the coarse dust during long-range transport. The simulations presented here, however, do not consider the coagulation processes and should represent all BC concentrations regardless of aerosol size. The measurements at Gosan are deemed among the most reliable BC observations during ACE-Asia, as they were verified by intercomparison with independent side-by-side sampling and analysis. Therefore, the underpredictions at Gosan may be a result of under-estimated or uninventoried emissions.

Observations at Yulin and Kwangju are grossly underpredicted (Figure 4b and 4c). These two sites are likely close to the local emission sources; the 80 km grid resolution implemented in this study is not capable of resolving concentrations in the vicinity of strong local sources. Because of the large grid size, the same problem, but to a lesser extent, can even exist for VMAP stations. Onboard the R/V Ron Brown the simulated and observed BC concentrations agree more closely (Figure 4d), although the simulation does not consistently reproduce the observed values. The closer agreement at sea is likely a result of the absence of local sources near the ship.

The adjoint equation is forced (driven) by the discrepancy between the simulated and observed concentrations, at the time and location of the observations. The temporal and spatial distributions of the adjoint variables characterize the areas of influence on the overall objective function, i.e. the concentrations and emissions at the locations and times
with significant adjoint values (positive or negative) play a more important role in assimilating the observations. The adjoint variable is used to calculate the gradient of the objective function with respect to different input parameters in Equation (9). These gradients are the sensitivities of the objective function with respect to various scaling factors in Equation (6) at different locations and times. For data assimilation of BC, the gradients with respect to initial conditions, boundary conditions, anthropogenic emissions, and biomass burning emissions are calculated (examples in Figure 5). A negative gradient in Figure 5 indicates that an increased scaling factor is required for the cost function to decrease, and vice versa. Boundary conditions and biomass burning emissions are represented by their daily gradients, i.e., their values for each day and grid cell are scaled by one scaling factor. For anthropogenic emissions, the scaling factor (for each location and time) is assumed to be comprised of a monthly and a daily contribution.

An iterative optimization procedure based on the calculated gradients results in reduced cost function and newly estimated (optimized) scaling factors for different parameters. Unlike data assimilation with simulated observations (the so-called identical twin experiment) in which the cost function can theoretically be reduced to zero, that with actual data generally does not result in dramatic reductions in the cost function. Daily averaging of the gradients and applying a single monthly factor for the anthropogenic emissions further constrain the optimization problem and contribute to the modest reduction in the cost function.
The objective function in Equation (9) is comprised of two parts: one accounting for the distance between the observations and simulations, the other for the deviation of the optimized parameters from their background (initial) values. The cost associated with each individual observation or scaled parameter is, in part, weighted by the uncertainty assigned to it. In other words, the observations with lower uncertainty account for a larger fraction of the total cost and are more aggressively assimilated. Likewise, input parameters with larger uncertainties are scaled more aggressively, and vice versa. Here, we assign generic uncertainties of 30, 10, and 20 percent to the observations made at the VMAP stations, Gosan, and onboard R/V Rob Brown, respectively. These uncertainties should be regarded as ensemble measures of the importance ascribed to each data set based on reliability of the observations and representativeness of the station. For background values we assign uncertainties of 300, 2000, 500, and 500 percent to the initial conditions, boundary conditions, biomass burning, and anthropogenic emissions of BC, respectively. The uncertainties assigned for the emissions are in line with those reported in Streets et al. [2003], noting that day- and location-specific emissions uncertainties are likely to be higher than those reported for annual and regional estimates. Boundary conditions are treated as highly uncertain, as for each specific day and location they may be significantly different from the time- and location-independent background value.

Apart from the uncertainties, the observation- and the background-driven parts of the objective function can be weighted by a global factor, i.e. $\mu$ in Equation (9). Such weight factor, albeit subjective in nature, can be interpreted as a measure of the strength
of the constraints that are applied to the input parameters. The effect of this global factor on the overall assimilation is shown in Table 2, where the reductions in total square errors (observation-driven portion of the objective function) are given for different levels of constraints in the background values. As expected, the unconstrained case shows the largest reduction, and the reductions become more modest for stronger constraints. On the other hand, the unconstrained assimilation also leads to unrealistically aggressive scaling of the input parameters. An example can be seen in Figure 6, where the monthly scaling factors for anthropogenic emissions of BC (for the four cases presented in Table 2) are shown. For this study we adapt the results of the case dubbed as “moderate constraint” in Table 2, as it provides reasonable reduction in the simulation errors and realistic estimates of the input parameters.

Spatial distributions of the optimized scaling factors for the initial conditions (as well as the base case initial concentrations) are shown in Figure 7. Scaling initial conditions is likely to have a small effect on the overall cost function, as after the first few days the effect of the initial values disappears. However, the initial concentrations can be important in assimilating the first few days of observations. As expected, the initial concentrations are mostly changed within a few days distance (travel time) from the observation sites, although slight scaling is seen as far as Myanmar.

The most significant change among the optimized parameters is the monthly scaling factors for anthropogenic emissions (Figure 7d), where the magnitude of those emissions are scaled as far as western India. The scaling is more significant in areas of
heavy emissions (Figure 7c), as these areas most significantly influence the objective function. The main feature of the estimated emission field is the reduction in southeastern China, and to lesser extent increased emissions in Japan, northeastern and eastern China.

As the adjoints are integrated backward in time and space from the observation sites, one can consider the adjoints as an ensemble of all back trajectories from the observation sites. Therefore, significant adjoint values, gradients, and scaling are more likely to occur in areas in the vicinity of observation sites.

Scales shown in Figure 7b and 7d are for the first vertical layer (surface-level). Table 3 gives a summary of the number of cells with significant scaling factors for different parameters and different vertical layers. Monthly anthropogenic emissions and initial conditions scaling factors are assigned a single value for each grid cell over the month, but scaling factors for boundary conditions and biomass burning emissions are adjusted daily. Overall, only a small fraction of the grid cells experience significant (larger than 10%) change from the starting point of unit scale factors; that fraction for initial conditions, boundary conditions, anthropogenic emissions, and biomass burning emissions amounts to 0.27, 2.32, 0.25, and 0.03 %, respectively. The small fraction of the changed values is due to the constraint that is applied to the optimization. Anthropogenic emissions and boundary conditions undergo the most significant scaling. Biomass burning emissions are sporadic in nature, and consequently are scaled for the smallest fraction of grid cells. First layer boundary inflow is less likely to affect the cells deep inside the domain, and, as expected, boundary conditions are adjusted more often and more significantly in the higher layers. The largest scaling for boundary conditions
occurs in layers 4-12, corresponding to approximately 1-3 km altitude. For the anthropogenic emissions (injected into layers 1-4 only), the first layer is the most significant. From a domain-wide point of view and on an average basis, anthropogenic and biomass burning emissions tend to be reduced as a result of the optimization, while the initial and boundary conditions are increased.

Apart from the domain-wide averages, a closer look at the smaller regions in the modeling domain provides better insight into the behavior of estimated scaling factors. Figures 8 and 9 show time series of the scaling factors for the average regional emissions (anthropogenic and biomass burning) and lateral boundary conditions (layers 1, 4, 8, and 12). For the boundary conditions, the effect of the southern boundary is insignificant, as it is predominantly one of outflow. Even though the western boundary is very far from the observation sites, it is scaled at higher altitudes. The eastern and northern boundaries exert larger effect on the observation sites and have more significant scaling factors, particularly in the higher layers. The period of large scaling factors in the northern boundary (days 21-25) coincides with a reported large wildfire in Siberia.

On the country level, anthropogenic scaling factors (Figure 9) are generally larger than those associated with biomass burning emissions. The Philippines and Vietnam/Thailand/Cambodia/Laos region do not have significantly modified emissions. The anthropogenic emissions in southeastern China are consistently and significantly reduced in the optimization process, while those for northeastern China generally exceed unity. For Japan the overall regional scaling factors for anthropogenic emissions indicate
an increase; however, some areas on the northern island of Hokkaido show decreased emissions (see Figure 7d). Emissions in Western India/Bangladesh and Korean peninsula remain relatively unchanged. Table 3 shows the total monthly emissions from these sub-regions. Despite some significant daily scaling factors, the biomass burning emissions do not change significantly as a result of assimilation. The total domain-wide anthropogenic emissions do not change significantly either; however, some notable changes at the country/sub-regional levels are seen. It should be noted that the anthropogenic scaling factors are specific to the month of April 2001, and they may reflect an inaccuracy in the predicted seasonality for the emissions rather than one in the inventoried annual emissions.

The scaling and optimization reduces the cost function by pushing the simulations closer to the observations, as shown in Figure 10. At different stations and over the course of assimilation, some of the features of the observational time series are successfully reproduced in the simulations. Most notably, model under-predictions at Sado are well compensated. However, some of the discrepancies between the observations and modeled concentrations (particularly at Hachijo and Chichijima) remain relatively unaffected by the optimization and scaling of the parameters. Note that compensating for the errors at one location may drive the simulations further from the observations at another, as seen in some instances in Figure 10. Such cases are more likely to occur at the VMAP stations, as those observations are assigned higher uncertainties. As expected the observations at Gosan and onboard R/V Ron Brown are assimilated fairly well, as they are assigned lower uncertainties.
It is desirable to evaluate the assimilation results with observations that were not used in the assimilation. We use the BC measurements onboard the C-130 and Twin Otter aircraft in this regard (Figure 11). In case of the C-130, both base case and assimilated simulations grossly under-predict the observations. Moreover, C-130 BC observations are significantly higher than other comparable observations made during the campaign [Huebert et al., 2004]. Highest concentrations are observed during very short sampling periods (10-30 minutes), likely indicative of passage through polluted air masses. Once again, due to the grid size (80 km) such concentrated plumes are not properly resolved in the simulations. Even though the simulations significantly under-predict the observed values onboard the C-130, the assimilated results are in general higher (and, therefore, closer to the observations), particularly during the periods of highest observed BC concentrations. In other words, the assimilation seems to move the simulation results in the right direction. The BC measurements onboard the Twin Otter are significantly lower than the C-130 observations and show better agreement with the simulations. Note that Twin Otter had three modes for BC sampling, and therefore, overlapping samples were taken during each flight [Mader et al., 2002]. Here, we use one sample/measurement for each flight, trying to avoid short sampling periods and pollution plumes (usually the same). Also, some of the reported BC concentrations are below the method’s detection limit; in case of two such observations, the lower value is used. The assimilated results are generally in better agreement with the observations. Out of 15 flights, the base case simulation is within the range of uncertainty for 2 observations, as compared to 10 for assimilated results.
It is important to note that the discrepancies between the observations and simulations may be caused by inaccuracies in parameters that are not included in the inverse modeling, e.g., different removal processes or unaccounted for emissions. However, the single most plausible source of inaccuracy (apart from those addressed in the data assimilation) is the meteorological input (in particular the wind fields) into the simulation. For instance, if the wind field incorrectly predicts that an air mass reaching a station does not cross the boundary or pass over major emission areas, the optimization will not properly assimilate that specific observation. For these reasons, the assimilation and the resulting scaling factors should be regarded as only trend indicators rather than strictly as revised values in the base case emissions inventory (or other input parameters). Furthermore, no uncertainty value or confidence interval has been assigned to the estimated scaling factors; in the absence of such consideration for uncertainties, the results are intended to be treated semi-qualitatively.

Summary

In this paper, we apply adjoint data assimilation for the recovery of BC emissions and initial and boundary conditions from observations during ACE-Asia. Measurements at four stations in the VMAP network, at Gosan, South Korea, onboard R/V Ron Brown and C-130 and Twin Otter aircraft, are chosen as the basis for the evaluation of the forward model, inverse modeling, and verification of the assimilated results. Optimizing location-dependent scaling factors to different input parameters through adjoint data assimilation significantly reduces the discrepancy between the model predictions and observations.
After assimilation, main temporal features at different stations are successfully reproduced by the model. Among different stations, data assimilation is more successful for Sado, Gosan, and R/V Ron Brown. The assimilated concentrations show markedly better agreement with the measurements made onboard the Twin Otter aircraft, and are generally in the direction of improved agreement with C-130 observations.

The assimilation results show that the northern and eastern boundary concentrations, particularly those in mid-altitudes, affect the simulated BC concentrations at the observation sites. Among the different types of parameters estimated, BC anthropogenic emissions have the most significant effect on the overall objective function and, therefore, change most during the course of the optimization. These emissions are also changed more significantly in the course of assimilation. Total anthropogenic or biomass burning BC emissions does not change noticeably as a result of assimilation. However, important changes to the anthropogenic BC emissions are seen at the sub-regional level; notably emissions increase for Japan and northeastern China, and significantly decrease for southeastern China. One can conclude that China in general and its southeastern provinces in particular, are the main areas where further work on emission inventory is required.

Adjoint inverse modeling is a powerful tool for providing insight into constraining various underlying inputs for CTMs. However, reliable and representative data and wide temporal and spatial coverage of measurements are essential for conclusiveness of data assimilation. Despite the range of observations during the ACE-Asia field campaign, and
because of the size of the eastern Asia region, the scarcity of available and applicable BC observations during ACE-Asia limits the range of conclusions that can be drawn from inverse modeling of this region. Nevertheless, the results from inverse modeling contain important information (in particular about the emissions inventory) that can help better understand the BC distributions in the region.
References:


**Acknowledgements**

This work was supported by the National Science Foundation award NSF ITR AP&IM 0205198.
Table 1: Summary of the observations considered for this study.

<table>
<thead>
<tr>
<th>Site/Platform</th>
<th>Location</th>
<th>Sampling duration</th>
<th>Number of observations</th>
<th>Cut-point (µm)</th>
<th>Use in this study</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMAP-Chichijima</td>
<td>27.0°N, 142.2°E</td>
<td>4 hours</td>
<td>165</td>
<td>2.5</td>
<td>Data assimilation</td>
<td>Matsumoto et al. [2003]</td>
</tr>
<tr>
<td>VMAP-Hachijo</td>
<td>33.2°N, 139.8°E</td>
<td>4 hours</td>
<td>164</td>
<td>2.5</td>
<td>Data assimilation</td>
<td>Matsumoto et al. [2003]</td>
</tr>
<tr>
<td>VMAP-Sado</td>
<td>38.3°N, 138.4°E</td>
<td>4 hours</td>
<td>178</td>
<td>2.5</td>
<td>Data assimilation</td>
<td>Matsumoto et al. [2003]</td>
</tr>
<tr>
<td>VMAP-Rishiri</td>
<td>45.1°N, 141.2°E</td>
<td>4 hours</td>
<td>172</td>
<td>2.5</td>
<td>Data assimilation</td>
<td>Matsumoto et al. [2003]</td>
</tr>
<tr>
<td>Gosan</td>
<td>33.3°N, 126.2°E</td>
<td>15-30 hours</td>
<td>30</td>
<td>2.5</td>
<td>Data assimilation</td>
<td>Schauer et al. [2003]</td>
</tr>
<tr>
<td>Kwangju</td>
<td>35.1°N, 126.5°E</td>
<td>~24 hours</td>
<td>28</td>
<td>2.5</td>
<td>None</td>
<td>Kim et al. [2004]</td>
</tr>
<tr>
<td>Yulin</td>
<td>38.3°N, 109.7°E</td>
<td>~24 hours</td>
<td>29</td>
<td>2.5</td>
<td>None</td>
<td>Xu et al. [2004]</td>
</tr>
<tr>
<td>Ron Brown</td>
<td>Moving</td>
<td>1-5 hours</td>
<td>118</td>
<td>1.0</td>
<td>Data assimilation</td>
<td>Lim et al. [2003]</td>
</tr>
<tr>
<td>C-130</td>
<td>Moving</td>
<td>0.4-3 hours</td>
<td>64</td>
<td>1.0</td>
<td>Verification</td>
<td>Huebert et al. [2004]</td>
</tr>
<tr>
<td>Twin-Otter</td>
<td>Moving</td>
<td>1-5 hours</td>
<td>15</td>
<td>1.0</td>
<td>Verification</td>
<td>Mader et al. [2002]</td>
</tr>
</tbody>
</table>

Table 2: Reduction in the total square error (observation-driven portion of the objective function) for varying levels of constraint in assimilation. The reduction in the errors for each case are given after 30 iterations.

<table>
<thead>
<tr>
<th>Assimilation case</th>
<th>µ in Equation 9 (µg²m⁶)</th>
<th>Reduction in total square error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconstrained</td>
<td>N/A</td>
<td>87.5</td>
</tr>
<tr>
<td>Weak constraint</td>
<td>100</td>
<td>78.7</td>
</tr>
<tr>
<td>Moderate constraint</td>
<td>20</td>
<td>64.5</td>
</tr>
<tr>
<td>Strong constraint</td>
<td>5</td>
<td>51.8</td>
</tr>
</tbody>
</table>
Table 3: Distribution of significant scaling factors in different vertical layers.

<table>
<thead>
<tr>
<th></th>
<th>&lt; 0.5</th>
<th>0.5 – 0.75</th>
<th>0.75 – 0.9</th>
<th>1.1 – 1.5</th>
<th>1.5 – 2.5</th>
<th>&gt; 2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anthropogenic emissions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 1</td>
<td>133</td>
<td>107</td>
<td>145</td>
<td>203</td>
<td>74</td>
<td>28</td>
</tr>
<tr>
<td>Layers 2-4</td>
<td>73</td>
<td>130</td>
<td>326</td>
<td>350</td>
<td>54</td>
<td>12</td>
</tr>
<tr>
<td>Layers 5-8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Layers 9-12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Biomass burning emissions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 1</td>
<td>66</td>
<td>62</td>
<td>153</td>
<td>66</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Layers 2-4</td>
<td>8</td>
<td>13</td>
<td>30</td>
<td>37</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Layers 5-8</td>
<td>7</td>
<td>21</td>
<td>18</td>
<td>18</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Layers 9-12</td>
<td>1</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Boundary conditions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 1</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>76</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Layers 2-4</td>
<td>0</td>
<td>17</td>
<td>78</td>
<td>291</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>Layers 5-8</td>
<td>15</td>
<td>68</td>
<td>194</td>
<td>668</td>
<td>241</td>
<td>53</td>
</tr>
<tr>
<td>Layers 9-18</td>
<td>57</td>
<td>172</td>
<td>598</td>
<td>845</td>
<td>242</td>
<td>72</td>
</tr>
<tr>
<td><strong>Initial conditions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layer 1</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>9</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Layers 2-4</td>
<td>0</td>
<td>1</td>
<td>35</td>
<td>54</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Layers 5-8</td>
<td>0</td>
<td>2</td>
<td>33</td>
<td>99</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Layers 9-18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4: Total anthropogenic and biomass burning emissions (Gg) of BC for the month of April 2001, in base case and optimized inventories.

<table>
<thead>
<tr>
<th>Region</th>
<th>Anthropogenic emissions</th>
<th>Biomass burning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base case</td>
<td>Assimilated</td>
</tr>
<tr>
<td>Japan</td>
<td>3.16</td>
<td>7.19</td>
</tr>
<tr>
<td>Vietnam/Laos/Cambodia/Thailand</td>
<td>7.18</td>
<td>6.93</td>
</tr>
<tr>
<td>North Korea/South Korea</td>
<td>3.19</td>
<td>2.86</td>
</tr>
<tr>
<td>Philippines</td>
<td>1.10</td>
<td>1.08</td>
</tr>
<tr>
<td>Southeastern China</td>
<td>28.51</td>
<td>18.75</td>
</tr>
<tr>
<td>Northeastern China</td>
<td>7.48</td>
<td>11.69</td>
</tr>
<tr>
<td>Eastern China/Beijing</td>
<td>18.57</td>
<td>19.76</td>
</tr>
<tr>
<td>Western India/Bangladesh</td>
<td>16.14</td>
<td>13.38</td>
</tr>
<tr>
<td>Others</td>
<td>26.54</td>
<td>31.06</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>111.87</td>
<td>112.70</td>
</tr>
</tbody>
</table>

37
Figure 1: Modeling domain for the adjoint data assimilation. Observation sites used for assimilation are also shown.

Figure 2: Spatial distributions of BC concentrations at its monthly peak, on April 7, 1:00 UTC.
Figure 3: Time series of BC observed and simulated concentrations at VMAP stations. Gray line indicates simulation results without wet removal processes.
Figure 3: (continued).
Figure 4: Observed and simulated daily BC concentrations at Gosan and Kwangju, Korea, and Yulin, China, and onboard the R/V Ron Brown.
Figure 4: (continued).
Figure 5: Spatial distributions of the gradient of the objective function with respect to a) monthly scaling factor for BC anthropogenic emissions, b) BC initial concentrations. The gradients shown are calculated from the base simulation results, as they change after each iteration.
Figure 6: Spatial distributions of the monthly scaling factor for the anthropogenic emissions of BC for different constraint levels in the optimization process: a) Unconstrained, b) weak constraint, c) moderate constraint, and d) strong constraint (see Table 2).
Figure 7: Spatial distribution of the (surface-level) scaling factors for b) BC initial concentrations, and d) monthly BC anthropogenic emissions. Also shown are the starting (base case) initial conditions (a) and ground-level anthropogenic emissions (c).
Figure 8: Time series of the scaling factors for BC lateral boundary concentrations (different vertical layers). Only grid cells with scaling factors that are significantly (>10%) different than one are used in the averaging.
Figure 9: Time series of the scaling factors for BC daily emissions over different sub-regions.
Figure 9: (continued).
Figure 10: Assimilated BC concentrations at different observation sites.
Figure 10: (continued).
Figure 10: (continued).
Figure 11: Comparison of the base case and assimilated BC concentrations with airborne measurements.