

THE AEROSOL MODELING TESTBED

A Community Tool to Objectively Evaluate Aerosol Process Modules

BY JEROME D. FAST, WILLIAM I. GUSTAFSON JR., ELAINE G. CHAPMAN, RICHARD C. EASTER,
JEREMY P. RISHEL, RAHUL A. ZAVERI, GEORG A. GRELL, AND MARY C. BARTH

The test bed is a new computational framework to streamline the process of testing and evaluating aerosol process modules over a range of spatial and temporal scales.

Many of the uncertainties associated with estimates of direct (via scattering and absorption of radiation by aerosols) and indirect (via droplet nucleation influenced by aerosols) radiative forcing in climate models (Solomon et al. 2007) can be attributed to inaccurate simulations of the spatial and temporal variations of aerosol mass, number, composition, mixing state, size distribution, hygroscopicity, and optical properties. For example, the formation

and transformation of secondary organic aerosols (SOAs; e.g., Volkamer et al. 2006) and the nature of many cloud–aerosol interactions (e.g., Lohmann and Feichter 2005) are still poorly understood and consequently inadequately represented in models. The coarse horizontal and vertical grid spacings usually employed by global climate models, which cannot resolve the observed spatial variability of atmospheric aerosols as well as meteorological factors that contribute to aerosol–radiation–cloud–chemistry interactions (e.g., Haywood et al. 1997; Petch 2001), are another factor that contributes to uncertainties in predictions of aerosol radiative forcing.

Regional and global models are becoming more complex as they incorporate new representations for the size distribution of aerosol mass and number and new parameterizations of aerosol processes. Journal articles that describe new parameterizations of aerosol processes usually employ a single model along with a dataset for a specific region and/or time period to quantify the performance of the new parameterization. The models, evaluation datasets, and other factors differ from study to study. One consequence of the current modeling paradigm is that the performance and computational efficiency of multiple treatments for a specific aerosol process cannot be quantitatively compared, because many other processes among aerosol models are different

AFFILIATIONS: FAST, GUSTAFSON, CHAPMAN, EASTER, RISHEL, AND ZAVERI—Pacific Northwest National Laboratory, Richland, Washington; GRELL—NOAA/Earth System Research Laboratory/Global Systems Division, and Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado; BARTH—National Center for Atmospheric Research,* Boulder, Colorado

*The National Center for Atmospheric Research is sponsored by the National Science Foundation

CORRESPONDING AUTHOR: Jerome D. Fast, Pacific Northwest National Laboratory, P.O. Box 999, K9-30, Richland, WA 99352

E-mail: jerome.fast@pnl.gov

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/2010BAMS2868.1

In final form 16 August 2010
©2011 American Meteorological Society

as well. For example, aerosol properties from several global models have been compared (e.g., Kinne et al. 2006; Textor et al. 2006; Barrie et al. 2001), but different grid configurations, meteorology, and emission rates were employed so that variations in predicted aerosol properties among the models were not entirely due to the treatment of aerosol processes. Studies such as these do not quantify the range of uncertainty associated only with aerosol treatments. Nor does this type of uncertainty analysis provide much information on which aerosol process needs improving the most and why.

In contrast to global aerosol models, regional-scale aerosol models can resolve much of the observed spatial and temporal variations in aerosols and consequently are more compatible with measurements that are often available over sampling intervals of an hour or less. Evaluations of regional-scale model predictions of aerosol mass and composition are usually conducted for specific cases with simulation periods ranging from a few days to a season (e.g., Mathur et al. 2008; Roy et al. 2007; Yu et al. 2008), rather than multiyear simulations of global models. The level of collaboration among regional modelers has been much lower than among global modelers, and comparisons of the performance of aerosol treatments have been haphazard. Only a few model intercomparison studies have been conducted to examine uncertainties in regional-scale aerosol predictions (e.g., McKeen et al. 2007, 2009; Stern et al. 2008). These intercomparison studies have the same problems as their global counterparts in relating uncertainties to specific aerosol processes.

Reducing the uncertainties associated with aerosols in both global and regional climate models requires that we know the advantages and disadvantages of specific aerosol treatments when the meteorology, chemistry, and other aerosol processes are identical. Penner et al. (1994) and Ghan and Schwartz (2007) describe strategies for improving estimates of aerosol radiative forcing in climate models that include three coordinated activities: 1) obtaining additional aerosol, chemical, and physical property measurements from surface, aircraft, and satellite instrumentation; 2) performing laboratory analysis and process studies that refine aerosol treatments; and 3) conducting 3D modeling studies to determine important factors contributing to aerosol radiative forcing and provide global-scale information that cannot be provided by data alone.

This study describes a new modeling paradigm that significantly advances how the third activity is conducted while also fully exploiting data and findings from the first two activities. The Aerosol Modeling Testbed (AMT) is a computational

framework for the atmospheric sciences community that streamlines the process of testing and evaluating aerosol process modules over a wide range of spatial and temporal scales. As described on the Pacific Northwest National Laboratory Web site (www.pnl.gov/atmospheric/research/aci/amt/index.stm), the AMT consists of a fully coupled meteorology–chemistry–aerosol model and a suite of tools to evaluate the performance of aerosol process modules via comparison with a wide range of field measurements. The philosophy of the AMT is to systematically and objectively evaluate aerosol process modules over local-to-regional spatial scales that are compatible with most field campaign measurement strategies. The performance of new treatments can then be quantified and compared to existing treatments before they are incorporated into regional and global climate models. Because the AMT is a community tool, it also provides a means of enhancing collaboration and coordination among aerosol modelers.

RELATIONSHIP TO OTHER TEST BEDS.

Table 1 lists two other test bed activities related to modeling aerosols: the Climate Change Prediction Program (CCPP)–Atmospheric Radiation Measurement Program (ARM) Parameterization Testbed (CAPT) and Aerosol Comparisons between Observations and Models (AeroCom). Rather than running the climate model for long periods of time, CAPT tests climate model parameterizations by performing short-term forecasts, similar to those produced by operational weather forecast models (Phillips et al. 2004). Most CAPT studies have tested cloud parameterizations within the Community Atmospheric Model (CAM), the atmospheric component of the Community Climate System Model (CCSM), using long-term data collected by the U.S. Department of Energy’s ARM program and other field campaign data (e.g., Williamson and Olson 2007; Xie et al. 2008). AeroCom (Kinne et al. 2006) assembles datasets that can be used to evaluate global aerosol models (Dentener et al. 2006) and documents the differences in aerosol sources, burden, residence time, and removal rates simulated by several global aerosol models (Textor et al. 2006). AeroCom is not a test bed per se, but its model comparison activities serve as one. It has also been used to compare global model predictions of direct radiative forcing for pre-industrial and present-day conditions (e.g., Shultz et al. 2006) and aerosol indirect effects (e.g., Penner et al. 2006; Quaas et al. 2009).

The AMT differs from CAPT and AeroCom in a number of ways. First, the AMT utilizes the Weather

TABLE 1. Comparison of the AMT capabilities to the CAPT and AeroCom test bed activities.			
	AMT	CAPT	AeroCom
Model	WRF	CAM	Multiple
Domain extent	Large-eddy simulation (LES) to mesoscale	Global and suggested control measure (SCM) mode	Global
Primary processes addressed	Aerosol properties, cloud–aerosol interactions, aerosol radiative forcing	Cloud properties	Aerosol properties, direct radiative forcing
Primary data used for evaluation	Field campaign, operational	Operational, field campaign, satellite	Satellite, AERONET

Research and Forecasting model (WRF) to simulate atmospheric conditions using grid spacings ranging from microscale to mesoscale applications ($\Delta x \sim 50\text{--}30\text{ km}$), whereas CAPT and AeroCom employ global models with grid spacings usually larger than 30–100 km. Global models also usually have fewer vertical levels (20–30) and a longer integration time step (15–20 min) than those employed by regional models. CAPT focuses on evaluating cloud parameterizations, including cloud–aerosol interactions, but does not currently evaluate aerosols. AeroCom primarily compares global model predictions of aerosol properties and direct and indirect radiative forcing but does include some evaluations using retrievals from operational satellite and Aerosol Robotic Network (AERONET) instrumentation. Thus, AMT fills a critical missing niche by evaluating aerosol process modules at scales more compatible with detailed aerosol properties and the spatial and temporal variability during field campaigns.

COMPONENTS OF THE AMT. *Host model.* WRF (Skamarock et al. 2005) was selected for the AMT to test and evaluate new aerosol process modules because its online coupling of meteorology and chemistry permits the simulation of aerosol radiative forcing (Grell et al. 2005; Fast et al. 2006). Because WRF is a freely available community model, it has also attracted thousands of users and developers worldwide.

Evaluating the performance and computational efficiency of new treatments for aerosol processes requires a sufficiently modular model so that all other atmospheric processes are treated similarly. Isolation of aerosol processes provides a better means of assessing the performance of specific aerosol treatments than traditional model intercomparison studies, where different treatments for various aerosol processes and other atmospheric processes confound interpretations of the findings. Consequently, we altered the structure of WRF to provide consistent

aerosol process modules for both modal and sectional aerosol models and made it easier for users to add new modules or modify existing modules.

Documenting the performance of aerosol process modules over time requires managing changes in WRF originating from both the WRF and AMT communities. WRF evolves as a result of new capabilities, revised meteorological parameterizations, and bug fixes, whereas changes in the AMT version originate from users adding or modifying aerosol process modules. Version control system software, called SubVersion (SVN; <http://subversion.tigris.org/>), is used to document changes in the model. The AMT will be a “branch” of the standard WRF release, and each user will have his own branch when making his modifications, as shown in Fig. 1. Some of our improved interoperability approaches for aerosol optical properties, aerosol–radiation–cloud interactions, and dry deposition modules have already been distributed back to the scientific community via new public releases of WRF.

Test bed cases. The amount and type of aerosol property measurements are limited compared

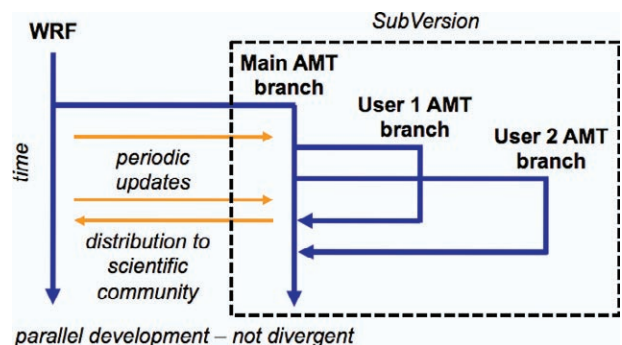


FIG. 1. Schematic diagram depicting how the SVN software is used to manage changes in the WRF code distributed by the National Center for Atmospheric Research (NCAR) and NOAA and code employed by the AMT.

to the quantity of meteorological and trace gas measurements. Although column-integrated quantities (i.e., aerosol optical depth) are routinely made by satellite instruments, in situ monitoring measurements of particulate mass and composition are usually available only at a limited number of sites. Aerosol size distributions and optical properties, key quantities for model evaluation, are not routinely obtained at such sites. Nor are measurements of aerosol properties aloft routinely collected. Measurements in the boundary layer and free troposphere are critical when evaluating models, because surface measurements are often not representative of the conditions very far above the ground. Vertical gradients in aerosol mass, composition, and size distribution also affect atmospheric heating rates and the amount and type of aerosols entrained into clouds; therefore, the semidirect and indirect effects are also dependent on the vertical distribution of aerosols.

Field campaigns address these issues by obtaining detailed measurements of particulate mass, composition, size distribution and aerosol optical properties, cloud condensation nuclei, and cloud properties both at the surface and in the troposphere at regional spatial scales. They are often conducted to coincide with satellite overpasses. These datasets have not been fully utilized to help reduce the uncertainties in the prediction of aerosol radiative forcing. The relatively short sampling periods (approximately 1 month) is a disadvantage; therefore, the AMT utilizes several datasets so that the performance of new and updated aerosol process modules can be evaluated over a wide range of atmospheric conditions.

Field campaign data are usually located at multiple Web and FTP sites or must be obtained directly from the principal investigator associated with each instrument. The time required to access and consolidate this disjointed mass of data is excessive and unfortunately limits their usefulness for the modeling community. Nor are many data archives maintained indefinitely, although archival processes have improved somewhat in recent years with the establishment of the National Aeronautics and Space Administration (NASA) and U.S. Department of Energy (DOE) Web sites that make their field campaign archives accessible to the public. Many different types of file formats are also employed, including American Standard Code for Information Interchange (ASCII), binary, network Common Data Format (netCDF), Hierarchical Data Format (HDF), and Microsoft Excel, depending on the archive. One of the tasks of the AMT is to collect field campaign data from multiple sources into a single archive

utilizing a common format for the data files, called a “test bed case.”

The Megacity Initiative Local and Global Research Observations (MILAGRO) field campaign was selected as the first test bed case. The focus of MILAGRO was to better understand how anthropogenic trace gases and particulates evolve over multiple spatial scales downwind of Mexico City (Molina et al. 2008; Fast et al. 2007). Instrumentation was deployed at three surface “supersites” and several other sites around Mexico City. Six research aircraft collected information up to ~11 km above sea level and up to several hundred kilometers downwind of Mexico City, and many aircraft flights were coordinated with satellite overpasses. The challenges of representing dust, biomass burning, and volcanic sources in addition to anthropogenic sources makes central Mexico a useful location to test and evaluate aerosol process modules.

As listed in Table 2, we are currently incorporating data on cloud–aerosol interactions from three recent field campaigns into the AMT, including the Cumulus Humilis Aerosol Plume Study (CHAPS), the Variability of the American Monsoon System (VAMOS) Ocean–Cloud–Atmosphere–Land Study (VOCALS), and the Indirect and Semi-Direct Aerosol Campaign (ISDAC). CHAPS was conducted in Oklahoma to better understand how fair-weather cumulus clouds affect the properties of anthropogenic aerosols (Berg et al. 2009). Although the goals of VOCALS are broad, the measurements collected off the coast of Chile were designed to provide information on how aerosols from natural and anthropogenic sources affect marine stratocumulus (Wood et al. 2006). ISDAC collected aerosol and cloud property measurements in the vicinity of Barrow, Alaska, to better understand how aerosols affect radiative forcing associated with mixed-phase clouds in the Arctic (<http://acrf-campaign.arm.gov/isdac/>). The extensive measurements from the NASA Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS; www.espo.nasa.gov/arctas/) and the National Oceanic and Atmospheric Administration (NOAA) Aerosol, Radiation, and Cloud Processes affecting Arctic Climate (ARCPAC; www.esrl.noaa.gov/csd/arcpac/) field campaigns are also included in this test bed case. To complement MILAGRO, measurements of urban organic matter and black carbon and how their optical and hygroscopic properties evolve from the Carbonaceous Aerosols and Radiative Effects Study (CARES; <http://acrf-campaign.arm.gov/cares/>) and the Research at the Nexus of Air Quality and Climate

TABLE 2. Field campaigns currently being developed as test bed cases for the AMT. The MILAGRO field campaign is completed, whereas the other campaigns are at various stages of development.

Field campaign	Time period	Location	Processes relevant to the AMT
MILAGRO	March 2006	Central Mexico and western Gulf of Mexico	Aerosol aging, organic aerosols, optical properties, oxidant chemistry, megacity emissions, biomass burning
CHAPS	June 2007	Vicinity of Oklahoma City	Effect of shallow cumulus on aerosol properties, urban emissions
VOCALS	Mid-October–mid-November 2008	Southeastern Pacific Ocean near Chile	Cloud–aerosol interactions, anthropogenic point source, biogenic emissions
ISDAC/ARCTAS/ARCPAC	March–April 2008	North slope of Alaska	Effect of aerosols on mixed-phase stratus clouds, long-range transport of aerosols
CARES/CalNex	May–June 2010	Vicinity of Sacramento, CA for CARES; primarily Los Angeles area and San Joaquin Valley for CalNex	Organic aerosols, aerosol mixing state, optical properties, urban and biogenic emissions

Change (CalNex; www.esrl.noaa.gov/csd/calnex/) will be added to the AMT in the near future.

The number of test bed cases will grow over time based on new field campaigns that are conducted, new instruments that are developed and deployed, and the scientific objectives of the users. Users are also free to develop additional test bed cases for the community, by following the same data format guidelines for the AMT described on our Web site.

Analysis Toolkit. Although software exists to evaluate meteorological predictions from WRF using operational monitoring data, there are no community tools that handle nonstandardized formats of the wide range of trace gas and aerosol property measurements associated with field campaigns. To address this issue, we developed a suite of programs that eliminates redundant tasks and streamlines the process of evaluating models for real-world cases. Thus, scientists can focus a larger fraction of their time on the relevant scientific questions rather than tedious file processing.

The Analysis Toolkit automatically extracts and derives simulated variables compatible with the available measurements using “instrument simulators” (e.g., Fan et al. 2009) as shown in Fig. 2. To facilitate a direct comparison of field campaign observed and simulated quan-

ties, they are organized into similar data file structures in five measurement categories: surface, aircraft, profile, radar, and satellite. The surface category contains in situ measurements at fixed sites, including supersites (extensive instrumentation), “secondary sites” (limited instrumentation), and operational monitoring network sites. The aircraft category contains in situ and remote sensing measurements that vary in space and time along flight paths. The radar category contains cloud and precipitation data from vertically pointing or scanning radars. Measurements at fixed sites from instrumentation such as vertically pointing lidars, radiosondes, and radar wind profilers are included in the profile category. The satellite category contains aerosol and cloud property

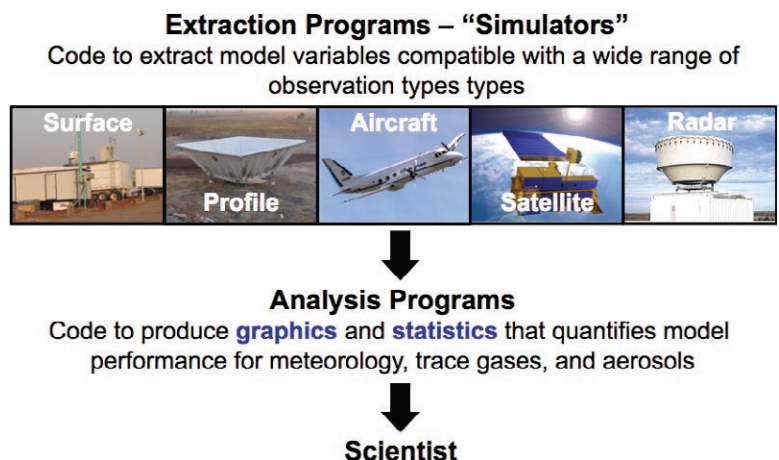


FIG. 2. Types of instrument simulators and data categories included in the Analysis Toolkit.

measurements that vary in space and time.

By default, information is extracted and derived from WRF needed to compare with all the available field measurements. Utilizing as much data as possible is important when assessing the impact of revised aerosol process modules on meteorological, chemical, or other aerosol quantities. Optionally, users can select specific variables or categories.

Standard statistical measures are generated (e.g., bias, index of agreement, root-mean-square error, correlation coefficient, percentiles), employing protocols that determine how to best compare the observed and simulated values that do not match spatially or temporally. For example, aircraft instrument sampling intervals typically range from 1 s to a few minutes and ground speeds vary among aircraft. For a ground speed of 100 m s⁻¹, instruments sampling at 1-, 10-, and 60-s intervals will collect data over distances of 0.1, 1, and 6 km, respectively. The Analysis Toolkit currently interpolates model output in space and time to the aircraft data points, although other protocols could easily be employed to evaluate the model results. Although WRF outputs instantaneous quantities, simulated quantities are unlikely to change significantly in time over one aircraft sampling period, so that averaging simulated quantities in time over 1-, 10-, or 60-s intervals will not likely affect the statistics. Differences between the observed sampling distance and the fixed model gridcell dimensions are likely to be more important. Interpolating model grid cells in space may be adequate for smoothly varying observed quantities, but it may be more appropriate to average several data points together when the observations exhibit significant random variability within a model grid cell. Similarly, when data sampling intervals span multiple model grid cells, the simulated quantities for those grid cells could be averaged. These types of issues need to be addressed further when quantifying model performance, which is relevant to the entire atmospheric modeling community. Interactions among the AMT user community will be one way of obtaining a con-

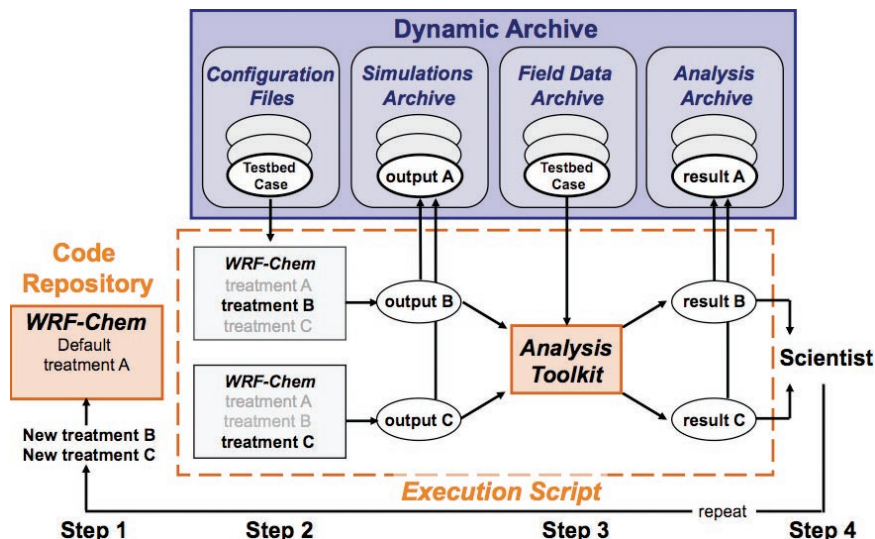


FIG. 3. Conceptual diagram depicting how the AMT would operate for the user community.

sensus on protocols for statistical evaluations using field campaign data.

Graphical tools, based on Gnuplot freeware for platform compatibility, provide users with “quick look” plots of observed and simulated quantities, because visually scrutinizing the results in space or time provides insights into model performance that are not necessarily obtained from statistics alone. The standard format of the files generated by the Analysis Toolkit permits them to be imported into other graphical software.

Operational aspects. The AMT design permits users to spend more time on science issues related to aerosol processes. It reduces tedious tasks, enabling new findings to be disseminated in less time and systematically documents improvements for specific aerosol process treatments.

A conceptual diagram of the four primary steps associated with the AMT is depicted in Fig. 3. For step 1, the user first checks out a copy of WRF from the code repository and implements a new treatment for a specific aerosol process. In this case, the user wishes to compare treatments “B” and “C” with the existing default treatment “A.” After preliminary tests are performed, a script executes steps 2–4. Currently, researchers independently spend inordinate amounts of time configuring model simulations, massaging the varied formats of measurement data, extracting model output into a form compatible with measurements, performing statistical analyses, and creating graphical output. The script eliminates these redundant and tedious tasks

For step 2, WRF is run with the new aerosol process modules using input for the desired case study period. Once the simulation is completed, the Analysis Toolkit is executed as part of step 3, extracting variables from model output consistent with the measurements, generating plots comparing observed and simulated quantities, and performing statistical evaluations. The user examines the results in relation to previous simulations as part of step 4 and can repeat the process if needed.

A key component of the AMT is the “dynamic archive.” Model simulations are normally not saved for a long period of time after a study has been published, despite the decreasing cost of computer storage. Users are likely to perform many simulations that debug and test new code; however, only the final simulations for a new treatment will be saved in the repository. A central repository that saves field campaign datasets as well as hundreds of terabytes of model output in an encyclopedic fashion is needed to provide a systematic means of documenting the performance of various aerosol treatments via the Analysis Toolkit. Saving local and regional aerosol simulations using the approach of the Earth System Grid (www.earthsystemgrid.org/), archiving climate simulations in a distributed fashion over several storage sites, could be adopted.

Skill scores for operational weather forecasts over the past several decades (e.g., Kalnay 2002) have been quantified by the National Centers for Environmental Prediction (NCEP). This tracking of incremental performance improvements has not yet been employed by the aerosol modeling community but is needed to produce higher public confidence in aerosol radiative forcing estimates from climate models. The aerosol modeling community could benefit greatly by adapting a similar approach, especially when defending the credibility of aerosol treatments used in global climate models.

We envision users having the ability to test and run their new aerosol modules within the AMT framework, either by using a central computing site or by downloading a more limited version of the AMT to their own computing systems. Network transfer speeds currently prohibit downloading the entire contents of the dynamic archive, which would likely grow to hundreds of terabytes in size; therefore, a limited version of the AMT will contain code, test bed cases (input files and observations), and statistical measures from previous simulations instead of the entire model output library. For example, the total size of the input files and observations from the MILAGRO test bed case is ~100 Gb. Some potential users may not be proficient

in running 3D models but will have new theoretical relationships or box-model treatments for an aerosol process that they wish to test. In this situation, it would be more efficient for AMT personnel to implement and evaluate new aerosol treatments for the investigator using SubVersion to track software changes.

EXAMPLE: SIMPLE VERSUS COMPLEX.

Objective. To demonstrate some of the AMT capabilities, we compare the performance of two aerosol models, the Modal Aerosol Dynamic Model/Secondary Organic Aerosol Module (MADE/SORGAM; Ackermann et al. 1998; Schell et al. 2001) and the Model for Simulating Aerosol Interactions and Chemistry (MOSAIC; Zaveri et al. 2008), previously implemented in WRF, for the MILAGRO test bed case. MADE/SORGAM adopts a modal approach for the aerosol size distribution and an equilibrium treatment for gas-particle partitioning, whereas MOSAIC adopts a sectional approach for the size distribution and a dynamic treatment for gas-particle partitioning. MOSAIC also includes comprehensive treatments for modeling aerosol thermodynamics for inorganic species (Zaveri et al. 2005b,a). To represent the aerosol size distribution, three modes are used by MADE/SORGAM and eight size bins are used by MOSAIC in the present application as shown in Fig. 4. In version 3.1 of WRF available to the public, the two aerosol models also employ different treatments for nucleation, coagulation, and dry deposition and are coupled to different photochemical mechanisms.

MOSAIC is theoretically more complex and ~1.8 times more computationally expensive (including input/output time) than MADE/SORGAM, even though it has 2.7 times as many prognostic species. On a bin or mode basis, MOSAIC would be faster than MADE/SORGAM. Comparing the models within the AMT will help determine whether higher complexity and number of prognostic species is worth the computational expense.

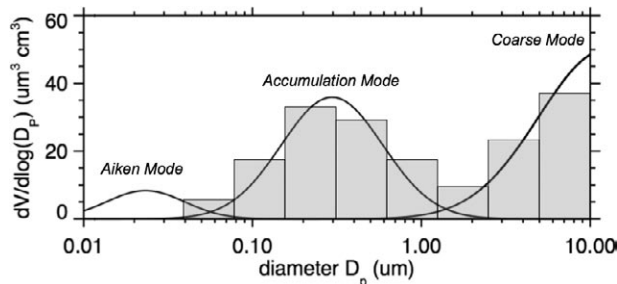


Fig. 4. Aerosol volume as a function of diameter for modal (lines) and sectional (gray boxes) size distributions.

The simulation period for the MILAGRO test bed case is from 6 to 30 March 2006, and two domains are used: an outer grid that encompasses Mexico and the surrounding ocean with a 12-km grid spacing and an inner grid that encompasses the central Mexican plateau with a 3-km grid spacing. In general, the boundary layer depth, local and regional circulations, and the transport of the Mexico City pollutant plume based on carbon monoxide concentrations were reproduced reasonably well as described in Fast et al. (2009). It is important that the simulated meteorology agree reasonably well with observations before assessing predictions of trace gases and aerosols.

Interoperability. As part of our systematic evaluation methodology, the standard version of WRF has been modified so that aerosol processes modules for MADE/SORGAM and MOSAIC are more interoperable. Here, both aerosol models employ the same emissions (Fast et al. 2009; Lei et al. 2007), meteorology, and boundary conditions from the Model for Ozone and Related Chemical Tracers (MOZART) global chemistry model (Pfister et al. 2008) as well as the same treatments for photochemistry Carbon Bond Mechanism (CBM)-Z; Zaveri and Peters (1999), aerosol optical properties (Fast et al. 2006), aerosol-radiation-cloud interactions (Chapman et al. 2009; Gustafson et al. 2007), and dry deposition [a modified version of Binkowski and Roselle (2003)]. In this study, the primary differences between the aerosol models

are for the aerosol size distribution and gas-particle partitioning. Although the treatments for nucleation differ as well, it will not likely be important in terms of total aerosol mass and direct radiative forcing. Treatments for secondary organic aerosol formation have also been turned off in both models.

Dry deposition. To illustrate the interoperability of aerosol process modules in WRF, a series of sensitivity simulations were first performed with various dry deposition treatments. Because simulated black carbon (BC) does not react with other compounds, the concentrations and distributions of BC have been examined to ensure that emissions, meteorology, and boundary conditions are handled identically for MADE/SORGAM and MOSAIC. Figure 5 shows the total mass of BC within the outer modeling domain over time. Both aerosol models produce the same mass when dry deposition is turned off, confirming that the emissions, meteorology, and boundary conditions are the same.

A new generic dry deposition module was added to WRF as part of the AMT to permit three treatments to be used for either MADE/SORGAM or MOSAIC. Dry deposition velocities v_d that vary according to aerosol diameter, surface properties, and meteorology are computed in this module and passed into the aerosol models. All code related to v_d is now in one location rather than distributed among the aerosol models, making it easier for users to refine these treatments or add new options. The differences in predicted BC mass among the three dry deposition options for MADE/SORGAM differ by as much as 12%. When both models employ the Zhang et al. (2001) parameterization, the difference in mass is ~3%. Similar differences were also produced when the other dry deposition schemes were used in both MOSAIC and MADE/SORGAM (not shown). These differences can be attributed to how the continuous modes and discrete size bins handle the aerosol size distribution (Fig. 4) as well as to coagulation and gas-particle partitioning processes that affect the growth of aerosols. The results also show that, for the same dry deposition treatment, the differences between MADE/SORGAM and MOSAIC are smaller than the differences among the various dry deposition treatments themselves.

Aerosol composition. An example of the simulated sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), and organic matter (OM) at the supersite in Mexico City is shown in Fig. 6 along with measurements obtained from an aerosol mass spectrometer (AMS; Aiken

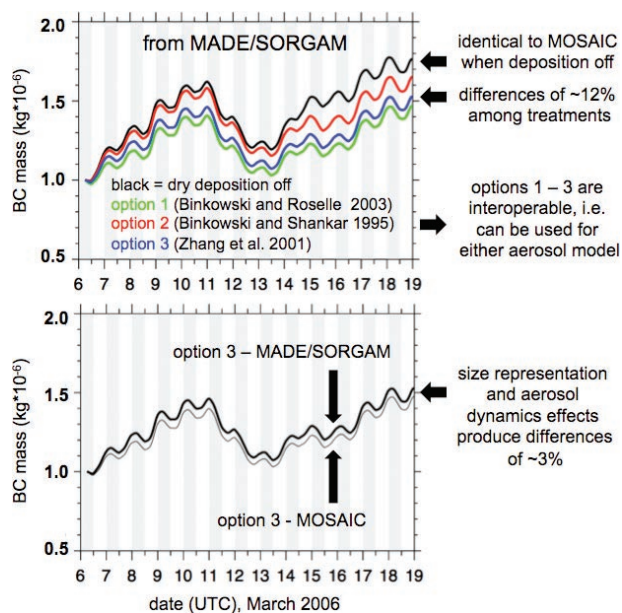


FIG. 5. Sensitivity of total BC mass within the outer modeling domain to three treatments for dry deposition.

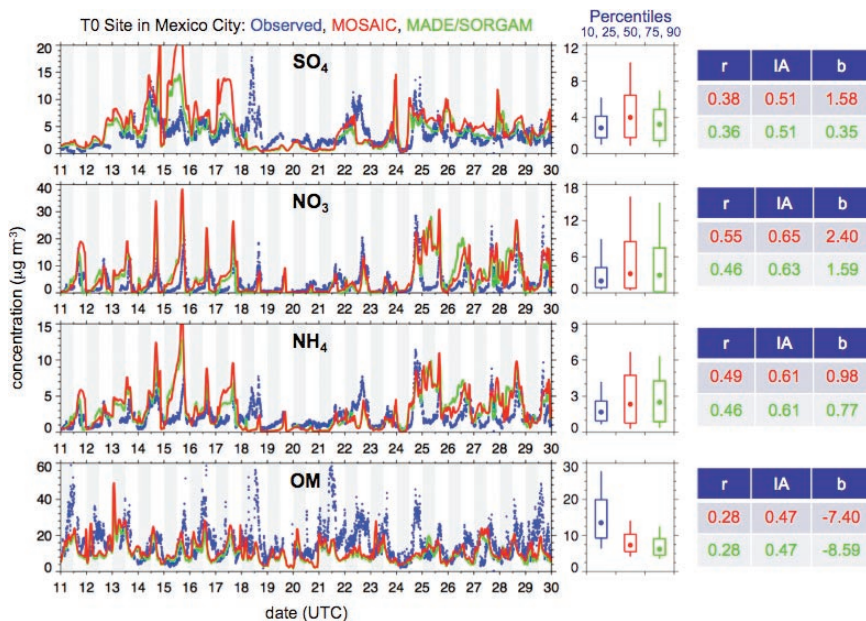


FIG. 6. Comparison of observed and predicted aerosol composition at the supersite in Mexico City. Observed composition obtained from an aerosol mass spectrometer. Statistics presented in terms of percentiles, r , IA, and b in $\mu\text{g m}^{-3}$.

et al. 2009). Rather than show the AMT's quick-look plots here, Figs. 6–9 were made by porting and combining the observations, simulated quantities, and statistics into another graphics package. The time series indicate that both simulations captured much of the diurnal and multiday variation in these aerosol components. Statistics generated by the Analysis Toolkit, including percentiles, correlation coefficient r , index of agreement (IA), and bias b , indicate that the performance of both aerosol models was similar in Mexico City. MOSAIC represents the temporal variability somewhat better than MADE/SORGAM as indicated by the correlation coefficients, whereas MADE/SORGAM has a bias that is somewhat lower for SO_4 , NO_3 , and NH_4 . Simulated OM from both aerosol models are nearly the same and lower than observed because SOAs are neglected; however, the overall magnitude of simulated primary organic aerosols (POAs) agree with derived POAs from the AMS

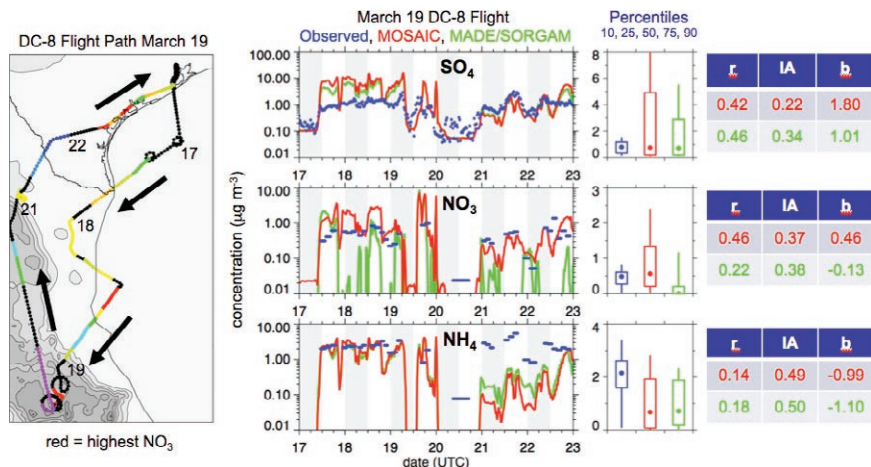


FIG. 7. Comparison of observed and predicted aerosol composition along the DC-8 flight on 19 Mar. Observed composition obtained from the University of New Hampshire's mist chamber and filter and ion-chromatography measurements.

measurements as described by Fast et al. (2009).

The Analysis Toolkit's "aircraft simulator" extracts model output corresponding to flight paths by interpolating in space and time. Statistics on aerosol composition between 0.7 and 2.3 km above the ground and within 50 km of Mexico City for 14 G-1 aircraft flights (Kleinman et al. 2008) were somewhat poorer than at the surface in the city (not shown). Aerosol composition from the C-130 (DeCarlo et al. 2008) and DC-8 aircraft (Singh et al. 2009) was examined to determine whether the differences between the two models were the same farther downwind

of Mexico City, and an example from the DC-8 aircraft using the University of New Hampshire's mist chamber (Scheuer et al. 2003) and filter and ion-chromatography measurements (Dibb et al. 2003) on 19 March is shown in Fig. 7.

SO_4 and NH_4 from both models were very similar along the entire flight. Simulated and observed SO_4 concentrations were similar to observed along the western flight leg, but along the eastern flight leg the

simulated values were too high. Most of the SO_4 in this region was produced from SO_2 emitted from the Popocatepetl volcano, suggesting that the estimated emissions (Grutter et al. 2008) were too high or that simulated vertical mixing in the midtroposphere was incorrect. In contrast, simulated NO_3 from the two models were very different. NO_3 from MADE/SORGAM was nearly zero along much of the DC-8 flight path. Although the variation in NO_3 from MOSAIC was much closer to observed, the bias was too high along the eastern flight leg. Other DC-8 flight legs (not shown) also indicated that the variability of NO_3 from MOSAIC was better than MADE/SORGAM over the Gulf of Mexico.

Aerosol size distribution. Predictions of mass and number size distributions are very important in terms of computing aerosol radiative forcing. For example, the impact of aerosols on scattering is highest for particle diameters around $0.4 \mu\text{m}$ at a wavelength of 550 nm (Seinfeld and Pandis 1998). Aerosol number concentrations are also closely related to cloud droplet number concentrations, influencing cloud albedo (Twomey 1974), lifetime (Albrecht 1989), extent of clouds (Ramanathan et al. 2001), and precipitation (Rosenfeld 2000). An example of the observed versus simulated number distribution in terms of percentiles among all the G-1 flights between 7 and 27 March is shown in Fig. 8. Because MOSAIC employs eight discrete size bins, the observations and MADE/SORGAM predictions have been lumped together in the same size bins so that they all can be compared to one another as shown in Fig. 8a. The Differential Mobility Analyzer (DMA) and Passive Cavity Aerosol Spectrometer Probe (PCASP) instrumentation (Kleinman et al. 2009) collected data over smaller particle diameter ranges, and MADE/SORGAM has

a continuous size distributions based on its Aiken, accumulation, and coarse modes, as shown in Fig. 8b. The number distribution from both MOSAIC and MADE/SORGAM were qualitatively similar to the measurements, although the models differ somewhat from one another. The range of values between the 25th and 75th percentiles from MADE/SORGAM is less than observed for particle diameters less than $0.625 \mu\text{m}$, whereas the range from MOSAIC is more consistent with the observations. Both MOSAIC and MADE/SORGAM underestimate aerosol number for bin 3 ($0.156225\text{--}0.3125 \mu\text{m}$), likely because of neglecting SOAs in both models. Aerosol number from both models is too high for bin 5 ($0.625\text{--}1.25 \mu\text{m}$), but for bin 6 ($1.25\text{--}2.5 \mu\text{m}$) MOSAIC is too high and MADE/SORGAM is too low, suggesting that the size distribution of coarse dust emissions is not treated adequately.

Aerosol optical depth and extinction. Satellite AOD measurements are frequently used by global climate models to evaluate simulated spatial variations of aerosols in the atmosphere (e.g., Kinne et al. 2006); therefore, this capability is included in the AMT. A “satellite simulator” either derives simulated AOD at the same grid spacing and overpass times as the *Terra* ($\sim 1700 \text{ UTC}$) and *Aqua* ($\sim 1900 \text{ UTC}$) satellites or averages satellite measurements to the model’s grid cells. Figure 9 shows a composite of the level-2 Moderate Resolution Imaging Spectroradiometer (MODIS) product (Levy et al. 2007) and the simulated AOD distributions at 550 nm from the 3-km domain for the *Terra* overpasses between 8 and 29 March. White regions denote areas with either $\text{AOD} < 0.1$ or no valid satellite measurements during the entire period.

Both models produce large AOD in the vicinity of large cities and gradients in AOD over the

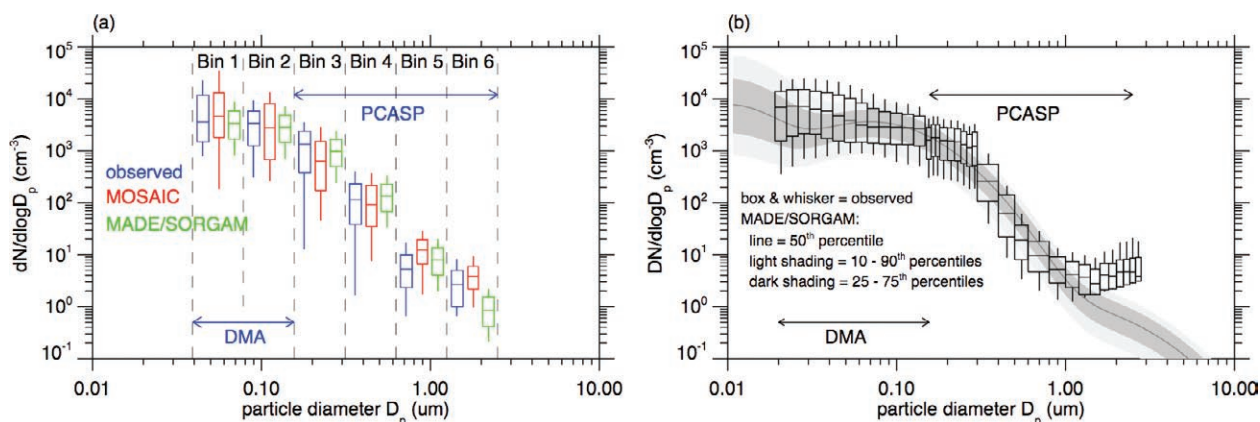


FIG. 8. Comparison of observed and predicted aerosol number among the 13 G-1 flights between 7 and 27 Mar expressed in terms of percentiles.

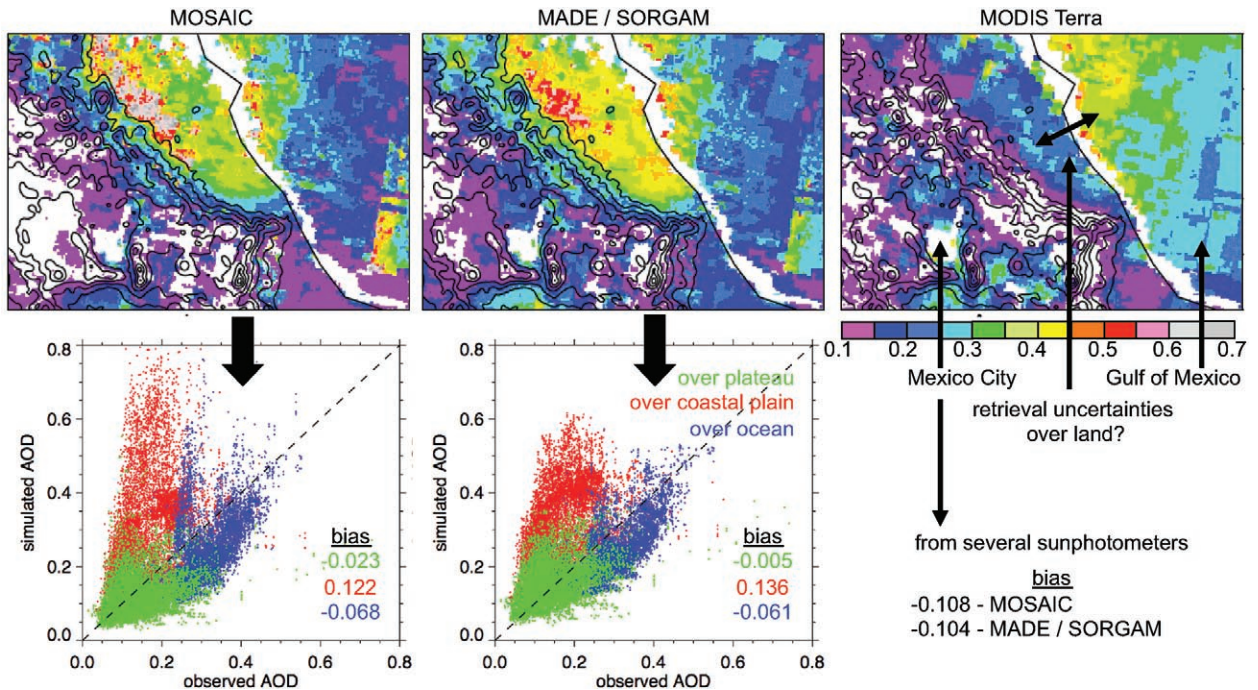


FIG. 9. Comparison of observed average AOD between 6 and 29 Mar and corresponding simulated values from the 3-km domain over central Mexico. Scatterplots show differences between observed and simulated AOD by region.

Gulf of Mexico, similar to MODIS measurements. Scatterplots for these regions indicate the bias is lowest over the central plateau, whereas simulated AOD is somewhat too low over the ocean. In contrast with MODIS, MADE/SORGAM and MOSAIC produce AOD of 0.3–1.0 over the coastal plain, whereas MODIS values were usually less than 0.3. MODIS is known to have higher uncertainties associated with the higher albedo over land (Levy et al. 2005). In the model, the high humidity over the coastal plain permits aerosols to uptake water in both MADE/SORGAM and MOSAIC, greatly increasing the amount of scattering; therefore, the strong gradient in AOD across the coast from MODIS does not seem realistic. Although the peak AOD from MOSAIC is higher than MADE/SORGAM, the mean AOD over the coastal plain from MADE/SORGAM is somewhat higher than MOSAIC. In addition, both aerosol models neglect SOA in the present simulations and therefore a large fraction of total aerosol mass. Consequently, the bias in AOD over land that is close to zero is likely too high. When simulated AOD was compared with AOD derived from various sun photometer measurements in the vicinity of Mexico City, both aerosol models had a bias of close to zero when observed AOD was <0.3, but simulated AOD from MOSAIC and MADE/SORGAM was much lower

than observed for larger observed AOD, with overall biases of -0.108 and -0.104 , respectively.

Fortunately, profiles of aerosol backscatter, extinction, depolarization, and other aerosol properties were obtained from the High Resolution Spectral Lidar (HRSL) deployed on NASA's B-200 aircraft (Rogers et al. 2009) provided information on vertically and horizontally varying aerosol properties. On 12 March, the B-200 flew over the Gulf of Mexico, the coastal plain, and the central plateau as shown in Fig. 10. In contrast with MODIS data, the HRSL obtained relatively high aerosol optical thickness (AOT) along the coast and over the coastal plain. The spatial variation in simulated AOT from both aerosol models was similar to the HRSL data, although MADE/SORGAM was usually too high over much of the flight path.

Vertical profiles of backscatter from the HRSL show that the AOT in the region is made up of several aerosol layers. These layers are likely the result of varying boundary layer depth associated with topography (e.g., De Wekker et al. 2004; Langford et al. 2010) as well as biomass burning plumes that originate from different elevations and have different injection heights. The simulated layering is qualitatively similar to HRSL observations, but the model's vertical grid spacing cannot resolve the thin aerosol layers in the midtroposphere. Both models also fre-

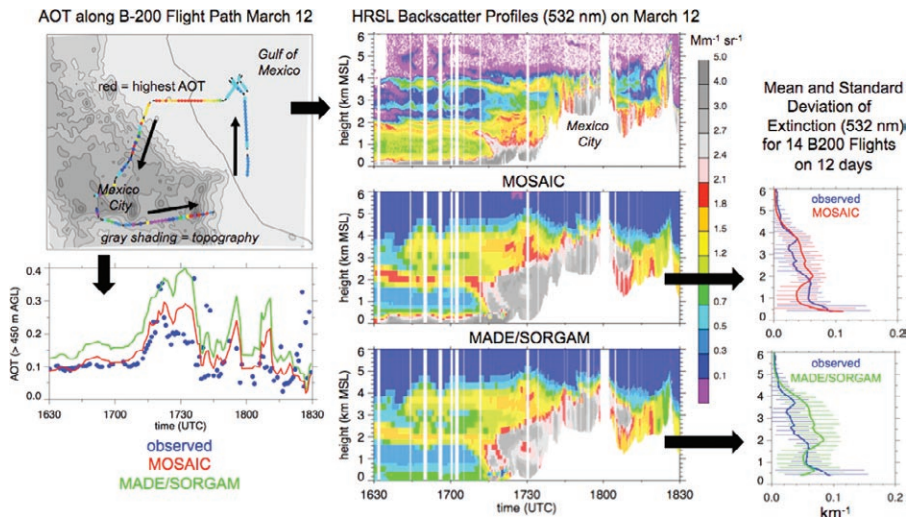


FIG. 10. Comparison of observed and simulated backscatter profiles and AOT between the surface and ~9 km MSL for the B-200 flight on 12 Mar. (right) Profiles of the mean and standard deviation of extinction among all the B-200 flights over central Mexico and the Gulf of Mexico.

quently overestimated backscatter in the convective boundary layer over land on this day.

The observed mean profiles of extinction and the standard deviation from all the B-200 flights also shown in Fig. 10 indicate the highest extinction was observed near the ground and gradually decreased with height to small values at ~5 km above mean sea level (MSL). MOSAIC produced extinction profiles adjacent to the ocean surface that were very similar to those from HRSL, suggesting that aerosol water uptake on sea salt was better represented by MOSAIC. Both models usually overestimated extinction between 2- and 4-km MSL, although MOSAIC was somewhat closer to HRSL than MADE/SORGAM. Errors in simulated relative humidity likely contribute to some of the differences between observed and simulated backscatter and extinction profiles. The bias, correlation coefficient, and index of agreement for relative humidity were -6%, 0.73, and 0.83, respectively; therefore, the model was drier than observed on average during the period. The slight underestimation of relative humidity suggests that either both models overestimated the uptake of aerosol water or that predicted total dry mass was too high; however, aerosol water and dust (a large component of dry mass) cannot be evaluated directly using the available measurements.

Differences in how well the aerosol models simulate vertical variations of mass, composition, and size distribution and consequently extinction and single-scattering albedo could affect local heating in the atmosphere (e.g., Johnson et al. 2008; Gadhavi and

Jayaraman 2006). This effect could be especially important depending on whether the aerosol layer is above or below clouds (e.g., Chand et al. 2009; Stone et al. 2008).

Impact on net shortwave radiation. To assess whether the differences in the gas-particle partitioning and size distribution between MADE/SORGAM and MOSAIC lead to significantly different estimates of direct radiative forcing, the average change in net shortwave radiation resulting from aerosols at

1500 UTC (0900 LT) from 8 and 29 March is shown in Fig. 11. Net shortwave radiation under clear skies without aerosols is 550–600 W m⁻² at this time. An offline code (not part of the AMT) that employs the same aerosol optical properties and radiation schemes as in WRF determined the impact of specific aerosol compositions on net shortwave radiation. In addition to differences in aerosol burden between the aerosol models, the modal and sectional size distributions affect single-scattering albedo (not shown) and consequently net shortwave radiation. Dust is a large fraction of the PM_{2.5} burden, especially over northwestern Mexico, but it is omitted here to demonstrate the impact of aerosols resulting from other sources.

As seen in the first two panels, both models produced similar spatial distributions of changes in net shortwave radiation due to aerosols, with the largest reductions occurring along the plateau edge 150–200 km northeast of Mexico City. On average, the reduction in net shortwave radiation from MOSAIC was as much as 27 W m⁻² greater than from MADE/SORGAM in that region, but the differences between the two models decreased to ~3 W m⁻² over the northern Gulf of Mexico. Note that clouds frequently formed along the eastern plateau edge; therefore, cloud albedo often has a larger impact on local radiation than aerosols. Synoptic-scale winds usually transported aerosol plumes northeastward along the coast and over the Gulf of Mexico, although there were a few periods in weak synoptic forcing and strong cold surges (Fast et al. 2007) in which aerosols were transported south of Mexico City.

The next three panels show the impact of carbonaceous (from OM and BC), inorganic (from SO_4 , NO_3 , NH_4 , and sea salt), and aerosol water on the net shortwave radiation. Not surprisingly, the effect of carbonaceous aerosols was very similar between the models because the emissions, boundary conditions, and dry deposition treatment were identical. Variations in the modal and sectional aerosol size distribution through the dry deposition treatment produced small differences in total OM and BC mass (Fig. 5) and subsequently differences in net shortwave radiation of 1 W m^{-2} or less. Larger differences in net shortwave radiation between the two models can be attributed to secondary inorganic aerosols. Although size distribution is still a factor, differences in net shortwave radiation of as much as 6 W m^{-2} can be attributed mostly to the gas-particle partitioning treatments in MADE/SORGAM and MOSAIC. The biggest impact on net shortwave radiation in MOSAIC was from aerosol water. Total aerosol water mass in MOSAIC over the outer modeling domain was often twice as much as MADE/SORGAM, leading to as much as 10 W m^{-2} more radiation scattered back to space on average. This effect is most pronounced along the eastern edge of the plateau.

The impacts of aerosols on the local net shortwave radiation in the vicinity of Mexico City are better resolved by the 3-km domain. The reduction in average net shortwave radiation at 1800 UTC between 8 and 29 March was -24 and -27 W m^{-2} for MADE/SORGAM and MOSAIC, respectively, when dust is

neglected. In contrast with the regional-scale results, OM and BC are largely responsible for this reduction, because relative humidity over the plateau is usually very low so that simulated aerosols do not deliquesce. Although emission rates are the same every weekday (slightly lower on weekends), the impact of aerosols on local radiation varies significantly from day to day, depending on the meteorology. On days with strong ambient winds (e.g., 18–20 March) that quickly transported aerosols and their precursors out of the basin, the reduction in net shortwave radiation was between -5 and -10 W m^{-2} . The effect of aerosols on radiation was as large as -60 W m^{-2} on days with weak ambient winds (e.g., 15 March). Large fires also led to high AOD reductions in net shortwave radiation that were often as large or larger than those from anthropogenic aerosols (e.g., 19 March).

It is interesting to note where the largest differences occur when comparing MADE/SORGAM and MOSAIC (i.e., uncertainties) in relation to the field campaign measurements. Most of the surface sampling sites were located in the vicinity of Mexico City, where the predictions between the two models were very similar. The largest differences in net shortwave radiation between the two models occurred farther downwind, as shown in Fig. 12. Although the largest differences on average occurred along the eastern plateau edge, the location of the peak difference between the two models varied from day to day, depending on the synoptic conditions. For example, on 19 March the largest difference occurred close to the border

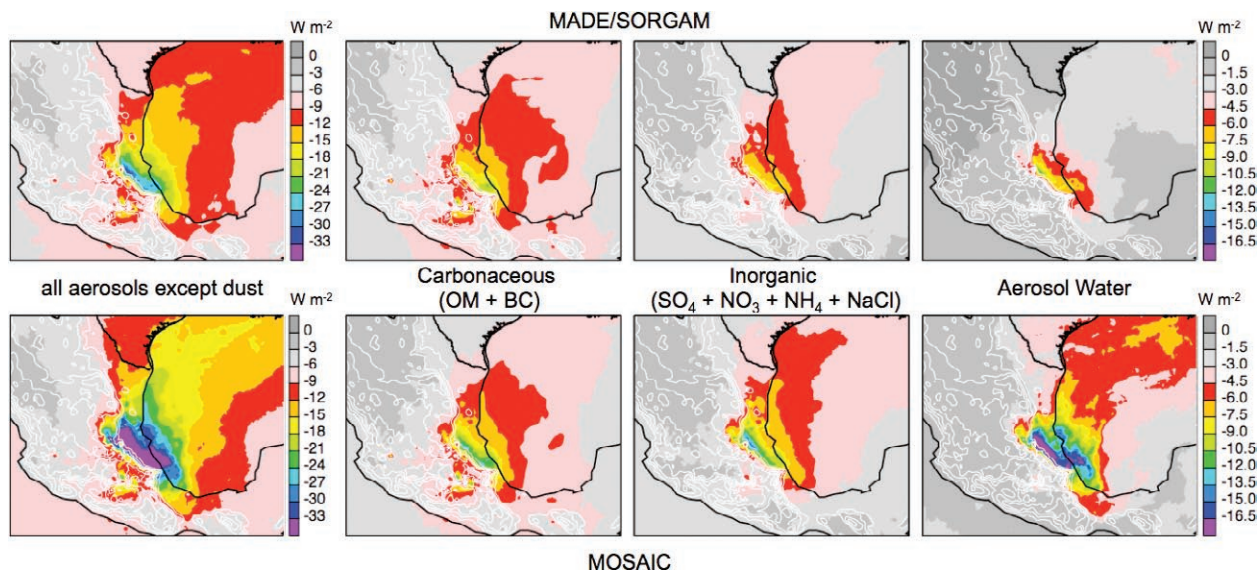


FIG. 11. Effect of aerosols on clear-sky downward shortwave radiation at 1500 UTC averaged between 8 and 29 Mar produced by (top) MADE/SORGAM and (bottom) MOSAIC attributable to various aerosol composition. Negative values indicate cooling resulting from aerosols.

Difference in Downward Shortwave Radiation
(MOSAIC – MADE/SORGAM) $W m^{-2}$

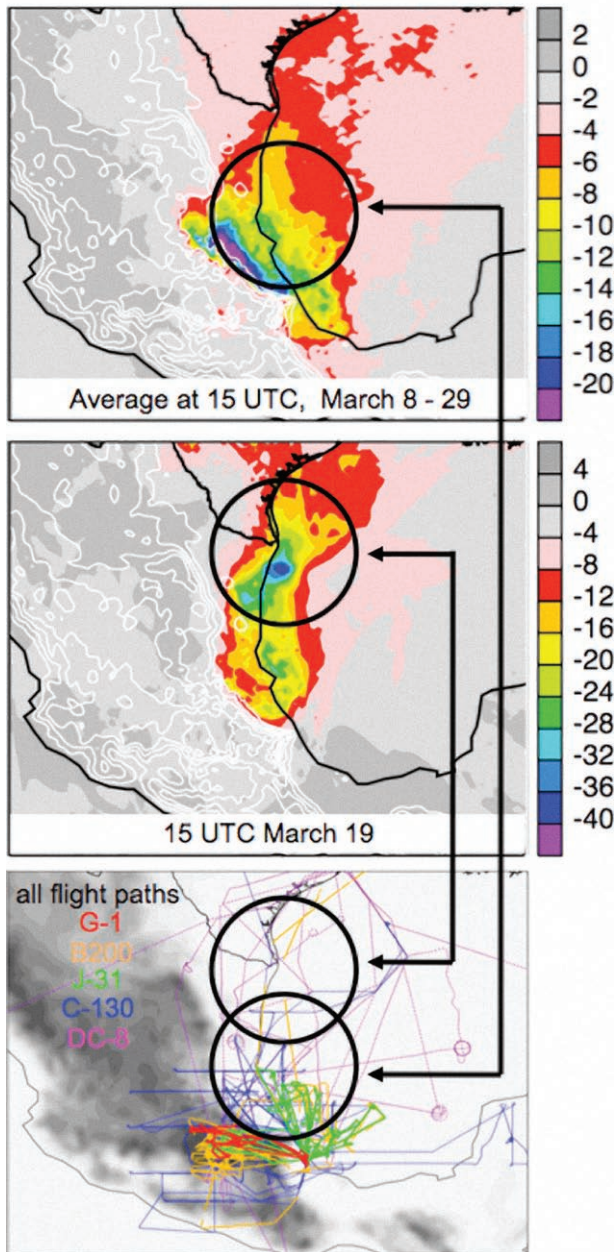


FIG. 12. (top) Average difference in shortwave radiation (MOSAIC – MADE/SORGAM) at 1800 UTC between 8 and 29 Mar and (middle) instantaneous difference at 1800 UTC 19 Mar in relation to (bottom) aircraft flight paths during the MILAGRO field campaign.

with Texas when the ambient southerly winds were strong. However, relatively fewer aircraft missions were conducted this far downwind of Mexico City.

The model uncertainties associated with specific treatments and illustrated by the AMT could be used in the future to guide aircraft operations. In this way, sufficient information can be collected where

the largest uncertainties among models occur. This information is needed to determine why a specific aerosol treatment performs better than others and to further improve the treatment.

Discussion. By employing the same host model and treating as meteorology, emissions, photochemistry, and many aerosol processes identically permitted a more objective comparison of the two aerosol models than the traditional modeling paradigm. Although some statistics for MOSAIC were somewhat better than MADE/SORGAM and vice versa, neither model outperformed the other over the central Mexican plateau. Larger differences in aerosol mass (mostly nitrate and aerosol water) between the two models were produced farther downwind along the coast and over the Gulf of Mexico. These differences also affected the regional distribution of surface shortwave radiation. Given that MOSAIC is ~1.8 times more computationally expensive than MADE/SORGAM, the benefits of MOSAIC are more apparent at the regional scale. A version of MOSAIC that employs a modal representation of the aerosol size distribution is still needed to more fairly compare the gas-particle partitioning treatments.

The results shown in this study are just a small fraction of the results available from the Analysis Toolkit. Thousands of quick-look plots are generated automatically that compare predicted and observed time series of meteorological, trace gas, and aerosol properties for surface sites and along aircraft flight tracks, including standard statistics.

A comprehensive approach that evaluates all model components is needed to assess the true performance of specific aerosol process modules. Evaluations that focus on a subset of field measurements neglect the potential interactions of meteorology, trace gas chemistry, aerosol chemistry, radiation, and cloud-aerosol interactions and the impact of a new aerosol process module could have on other quantities. Improvements in one simulated quantity could be accompanied by a decrease in performance in other simulated quantities.

SUMMARY AND FUTURE DIRECTIONS.

Further information on the AMT and its current status can be found on the Pacific Northwest National Laboratory Web site (www.pnl.gov/atmospheric/research/aci/amt/index.stm). Although many of the components of the AMT have been completed, several key features needed for the user community remain to be developed. Most of the Analysis Toolkit software has been completed, except for the radar

simulation shown in Fig. 2. Standard statistical measures (bias, index of agreement, root-mean-square error, correlation coefficient, percentiles) and simple quick-look graphics are produced by the Analysis Toolkit, but additional statistics will be developed based on user needs. The MILAGRO test bed case is complete, including porting all the datasets into a common format and creating configuration files for WRF simulations. As stated previously, we are currently porting four other field campaigns into the test bed as listed in Table 2, so that aerosol process modules can be evaluated over a wider range of atmospheric conditions and aerosol sources. The Web site also includes documentation for users to create other test bed cases for their own research purposes. The ability to manage a large number of the model simulations and the accompanying statistics as part of the dynamic archive shown in Fig. 3 has yet to be completed. The AMT is expected to evolve based on the needs of the user community, new types of test bed cases, and additional capabilities developed for the Analysis Toolkit.

By improving aerosol process modules, the AMT will also benefit the air quality modeling community. Parameterizations for meteorological processes, such as those for boundary layer mixing and cloud microphysics, can also be tested. The primary long-term goal of the AMT, however, is to facilitate the development of improved aerosol process modules for climate applications by the following subsections:

Translating findings from laboratory and field studies. Analysis of data collected from laboratory and field studies can lead to a range of new aerosol process treatments that range from simple (e.g., revised rate constants or empirical parameterizations) to the complex (e.g., representation of SOAs). The intent of the AMT is to incorporate those findings into 3D models to quantitatively assess their impact on predicted particulate properties, as well as their effect on trace gases and meteorological parameters.

Evaluating aerosol treatments employed by global climate models. Global climate models already contain simple aerosol treatments that are now being replaced with new, more complex treatments. It is attractive to evaluate new treatments in a local-to-regional-scale framework compatible with field measurements and quantify their performance prior to use in GCM predictions.

Evaluating aerosol treatments for regional climate models. Current regional-scale climate models use boundary

conditions (obtained from global climate models) to control climate forcing. Direct and indirect forcing, as well as most physicochemical processes, are currently neglected in regional-scale climate models for computational reasons. As computer power continues to increase, regional climate models will need to incorporate treatments of aerosol direct and indirect forcing similar to global climate models. Because the AMT is designed for regional scales, it will provide guidance as to the performance of aerosol process modules suitable for regional climate models.

Data mining. The collection of all the field campaign measurements and primary simulations into a single archive is itself a valuable activity. Non-AMT users can access these data, likely enabling other types of analyses that cannot be anticipated a priori. This type of activity has already been employed by the global climate modeling community and proven useful: for example, the Program for Climate Model Diagnosis and Intercomparison funded by the DOE Climate Change Research Division (<http://www-pcmdi.llnl.gov>).

The systematic evaluation methodology employed by the AMT will measure improvements both in accuracy and in computational requirements. The advantage and disadvantages of various aerosol process modules can be fairly assessed so that modelers can judge which treatments should be incorporated into global climate models. Some aerosol process modules will likely be developed for research only; that is, computational requirements will be such that they are not feasible for direct use in long-term climate simulations. However, in the near term such modules would serve as the basis for developing parameterization for use in those applications, and in the long term they would be used directly in global climate models when sufficient computing power becomes available.

ACKNOWLEDGMENTS. Development of the Aerosol Modeling Testbed was supported by Pacific Northwest National Laboratory's (PNNL) Laboratory Directed Research and Development program through the Aerosol Climate Initiative. The authors thank Stuart McKeen (NOAA) for providing a dry deposition treatment and the numerous MILAGRO investigators for providing data, emissions, and boundary conditions used in this study, including Jose Jimenez and Allison Aiken (University of Colorado), Larry Kleinman (Brookhaven National Laboratory), and Richard Ferrare (NASA). PNNL is operated for the U.S. Department of Energy by Battelle Memorial Institute.

REFERENCES

- Ackermann, I. J., H. Hass, M. Memmesheimer, A. Ebel, F. S. Binkowski, and U. Shankar, 1998: Modal aerosol dynamics model for Europe: Development and first applications. *Atmos. Environ.*, **32**, 2981–2999.
- Aiken, C., and Coauthors, 2009: Mexico City aerosol analysis during MILAGRO using high resolution aerosol mass spectrometry at the urban supersite (T0)—Part I: Fine particle composition and organic source apportionment. *Atmos. Chem. Phys.*, **6**, 6633–6653.
- Albrecht, B. A., 1989: Aerosols, cloud microphysics, and fractional cloudiness. *Science*, **245**, 1227–1230.
- Barrie, L., and Coauthors, 2001: A comparison of large-scale atmospheric sulfate models (COSAM): Overview and highlights. *Tellus*, **53B**, 615–645.
- Berg, L. K., and Coauthors, 2009: Overview of the Cumulus Humilis Aerosol Processing Study. *Bull. Amer. Meteor. Soc.*, **90**, 1653–1667.
- Binkowski, F. S., and U. Shankar, 1995: The Regional Particulate Matter Model. 1. Model description and preliminary results. *J. Geophys. Res.*, **100**, 26 191–26 209.
- , and S. J. Roselle, 2003: Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component. 1. Model description. *J. Geophys. Res.*, **108**, 4183, doi:10.1029/2001JD001409.
- Chand, D., R. Wood, T. L. Anderson, S. K. Satheesh, and R. J. Charlson, 2009: Satellite-derived direct radiative effect of aerosols dependent on cloud cover. *Nat. Geosci.*, **2**, 181–184.
- Chapman, E. G., W. I. Gustafson Jr., R. C. Easter, J. C. Barnard, S. J. Ghan, M. S. Pekour, and J. D. Fast, 2009: Coupling aerosols-cloud-radiative processes in the WRF-Chem model: Investigating the radiative impact of large point sources. *Atmos. Chem. Phys.*, **9**, 945–964.
- DeCarlo, P. F., and Coauthors, 2008: Fast airborne aerosol size and chemistry measurements above Mexico City and central Mexico during the MILAGRO campaign. *Atmos. Chem. Phys.*, **8**, 4027–4048.
- Dentener, F., and Coauthors, 2006: Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom. *Atmos. Chem. Phys.*, **6**, 4321–4344.
- De Wekker, S. F. J., D. G. Steyn, and S. Nyeki, 2004: A comparison of aerosol-layer and convective boundary-layer structure over a mountain range during STAAARTE '97. *Bound.-Layer Meteor.*, **113**, 249–271.
- Dibb, J. E., R. W. Talbot, E. M. Scheuer, G. Seid, M. A. Avery, and H. B. Singh, 2003: Aerosol chemical composition in Asian continental outflow during the TRACE-P campaign: Comparison with PEM-West B. *J. Geophys. Res.*, **108**, 8815, doi:10.1029/2002JD003111.
- Fan, J., M. Ovtchinnikov, J. M. Comstock, S. A. McFarlane, and A. Khain, 2009: Ice formation in Arctic mixed-phase clouds: Insights from a 3-D cloud-resolving model with size-resolved aerosol and cloud microphysics. *J. Geophys. Res.*, **114**, D04205, doi:10.1029/2008JD010782.
- Fast, J. D., W. I. Gustafson Jr., R. C. Easter, R. A. Zaveri, J. C. Barnard, E. G. Chapman, G. A. Grell, and S. E. Peckham, 2006: Evolution of ozone, particulates, and aerosol direct forcing in an urban area using a new fully-coupled meteorology, chemistry, and aerosol model. *J. Geophys. Res.*, **111**, D21305, doi:10.1029/2005JD006721.
- , and Coauthors, 2007: A meteorological overview of the MILAGRO field campaigns. *Atmos. Chem. Phys.*, **7**, 2233–2257.
- , and Coauthors, 2009: Evaluating simulated primary anthropogenic and biomass burning organic aerosols during MILAGRO: Implications for assessing treatments of secondary organic aerosols. *Atmos. Chem. Phys.*, **9**, 6196–6215.
- Gadhavi, H., and A. Jayaraman, 2006: Airborne lidar study of the vertical distribution of aerosols over Hyderabad, an urban site in central India, and its implications for radiative forcing calculations. *Ann. Geophys.*, **24**, 2461–2470.
- Ghan, S. J., and S. E. Schwartz, 2007: Aerosol properties and processes: A path from field and laboratory measurements to global climate models. *Bull. Amer. Meteor. Soc.*, **88**, 1059–1083.
- Grell, G. A., S. E. Peckham, R. Schmitz, and S. A. McKeen, G. Frost, W. C. Skamarock, and B. Eder, 2005: Fully coupled “online” chemistry within the WRF model. *Atmos. Environ.*, **39**, 6957–6976.
- Grutter, M., R. Basaldud, C. Rivera, R. Harig, W. Junkerman, E. Caetano, and H. Delgado-Granados, 2008: SO₂ emissions from Popocatepetl volcano: Emission rates and plume imaging using optical remote sensing techniques. *Atmos. Chem. Phys.*, **8**, 6655–6663.
- Gustafson, W. I., Jr., E. G. Chapman, S. J. Ghan, R. E. Easter, and J. D. Fast, 2007: Impact on modeled cloud characteristics due to simplified treatment of uniform cloud condensation nuclei during NEAQS 2004. *Geophys. Res. Lett.*, **34**, L19809, doi:10.1029/2007GL030021.
- Haywood, J. M., V. Ramaswamy, and L. J. Donner, 1997: A limited-area model case study of the effects of sub-grid scale variations in relative humidity and cloud upon the direct radiative forcing of sulfate aerosol. *Geophys. Res. Lett.*, **24**, 143–146.

- Johnston, B. T., B. Heese, S. A. McFarlane, P. Chazette, A. Jones, and N. Bellouin, 2008: Vertical distribution and radiative effects of mineral dust and biomass burning over West Africa during DABEX. *J. Geophys. Res.*, **113**, D00C12, doi:10.1029/2008JD009848.
- Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 364 pp.
- Kinne, S., and Coauthors, 2006: An AeroCom initial assessment—Optical properties in aerosol component modules of global models. *Atmos. Chem. Phys.*, **6**, 1815–1834.
- Kleinman, L. I., and Coauthors, 2008: The time evolution of aerosol composition over the Mexico City plateau. *Atmos. Chem. Phys.*, **8**, 1559–1575.
- , and Coauthors, 2009: The time evolution of aerosol size distribution over the Mexico City plateau. *Atmos. Chem. Phys.*, **9**, 4261–4278.
- Langford, A. O., C. J. Senff, R. J. Alvarez II, R. M. Banta, and R. M. Hardesty, 2010: Long-range transport of ozone from the Los Angeles Basin: A case study. *Geophys. Res. Lett.*, **37**, L06807, doi:10.1029/2010GL042507.
- Lei, W., B. de Foy, M. Zavala, R. Volkamer, and L. T. Molina, 2007: Characterizing ozone production in the Mexico City metropolitan area: A case study using a chemical transport model. *Atmos. Chem. Phys.*, **7**, 1347–1366.
- Levy, R. C., L. A. Remer, J. V. Martins, Y. J. Kaufman, A. Plana-Flattori, J. Redemann, and B. Wenny, 2005: Evaluation of the MODIS aerosol retrievals over ocean and land during CLAMS. *J. Atmos. Sci.*, **62**, 974–992.
- , —, S. Mattoo, E. F. Vermote, and Y. J. Kaufman, 2007: Second-generation operational algorithm: Retrieval of aerosol optical properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance. *J. Geophys. Res.*, **112**, D13211, doi:10.1029/2006JD007811.
- Lohmann, U., and J. Feichter, 2005: Global aerosol indirect effects. A review. *Atmos. Chem. Phys.*, **5**, 715–737.
- Mathur, R., S. Yu, D. Kang, and K. L. Schere, 2008: Assessment of the wintertime performance of developmental particulate matter forecasts with the Eta-Community Multiscale Air Quality modeling system. *J. Geophys. Res.*, **113**, D02303, doi:10.1029/2007JD008580.
- McKeen, S., and Coauthors, 2007: Evaluation of several PM_{2.5} forecast models using data collected during the ICARTT/NEAQS 2004 field study. *J. Geophys. Res.*, **112**, D10S20, doi:10.1029/2006JD007608.
- , and Coauthors, 2009: An evaluation of real-time air quality forecasts and their urban emissions over eastern Texas during the summer of 2006 Second Texas Air Quality Study field study. *J. Geophys. Res.*, **114**, D00F11, doi:10.1029/2008JD011697.
- Molina, L. T., S. Madronich, J. S. Gaffney, and H. Singh, 2008: Overview of MILAGRO/INTEX-B campaign. *IGACTivities Newsletter*, No. 38, IGAC Seattle Core Project Office, Seattle, Washington, 2–15.
- Penner, J. E., and Coauthors, 1994: Quantifying and minimizing uncertainty of climate forcing by anthropogenic aerosols. *Bull. Amer. Meteor. Soc.*, **75**, 375–400.
- , and Coauthors, 2006: Model intercomparison of indirect aerosol effects. *Atmos. Chem. Phys.*, **6**, 3391–3405.
- Petch, J. C., 2001: Using a cloud-resolving model to study the effects of subgrid-scale variations in relative humidity on direct sulphate aerosol forcing. *Quart. J. Roy. Meteor. Soc.*, **127**, 2385–2394.
- Pfister, G. G., and Coauthors, 2008: Contribution of isoprene to chemical budgets: A model tracer study with the NCAR CTM MOZART-4. *J. Geophys. Res.*, **113**, D05308, doi:10.1029/2007JD008948.
- Phillips, T. J., and Coauthors, 2004: Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction. *Bull. Amer. Meteor. Soc.*, **85**, 1903–1915.
- Quaas, J., and Coauthors, 2009: Aerosol indirect effects—General circulation model intercomparison and evaluation with satellite data. *Atmos. Chem. Phys.*, **9**, 8697–8717.
- Ramanathan, V., P. J. Crutzen, J. T. Kiehl, and D. Rosenfeld, 2001: Aerosols, climate, and the hydrological cycle. *Science*, **294**, 2119–2124.
- Rogers, R. R., and Coauthors, 2009: NASA LaRC airborne high spectral resolution lidar aerosol measurements during MILAGRO: Observations and validation. *Atmos. Chem. Phys.*, **9**, 4811–4826.
- Rosenfeld, D., 2000: Suppression of rain and snow by urban and industrial air pollution. *Science*, **287**, 1793–1796.
- Roy, B., R. Mathur, A. B. Gilliland, and S. C. Howard, 2007: A comparison of CMAQ-based aerosol properties with IMPROVE, MODIS, and AERONET data. *J. Geophys. Res.*, **112**, D14301, doi:10.1029/2006JD008085.
- Schell, B., I. J. Ackermann, H. Hass, F. S. Binkowski, and A. Ebel, 2001: Modeling the formation of secondary organic aerosol within a comprehensive air quality modeling system. *J. Geophys. Res.*, **106**, 28 275–28 293.
- Scheuer, E., R. W. Talbot, J. E. Dibb, G. K. Seid, L. DeBell, and B. Lefer, 2003: Seasonal aerosol distributions of fine aerosol sulfate in the North American Arctic

- basin during TOPSE. *J. Geophys. Res.*, **108**, 8370, doi:10.1029/2001JD001364.
- Seinfeld, J. H., and S. N. Pandis, 1998: *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons, 1326 pp.
- Shultz, M., and Coauthors, 2006: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations. *Atmos. Chem. Phys.*, **6**, 5225–5246.
- Singh, H. B., W. H. Brune, J. H. Crawford, F. Flocke, and D. J. Jacob, 2009: 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.*, **9**, 2301–2318.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the advanced research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 8 pp.
- Solomon, S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen, Eds., 2007: *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 996 pp.
- Stern, R., and Coauthors, 2008: A model inter-comparison study focusing on episode with elevated PM₁₀ concentrations. *Atmos. Environ.*, **42**, 4567–4588.
- Stone, R. S., G. P. Anderson, E. P. Shettle, E. Andrews, K. Loukachine, E. G. Dutton, C. Schaaf, and M. O. Roman III, 2008: Radiative impact of boreal smoke in the Arctic: Observed and modeled. *J. Geophys. Res.*, **113**, D14S16, doi:10.1029/2007JD009657.
- Textor, C., and Coauthors, 2006: Analysis and quantification of the diversities of aerosol life cycles within AeroCom. *Atmos. Chem. Phys.*, **6**, 1777–1813.
- Twomey, S., 1974: Aerosols, clouds and radiation. *Atmos. Environ.*, **25A**, 2435–2442.
- Volkamer, R., and Coauthors, 2006: Secondary organic aerosol formation from anthropogenic air pollution: Rapid and higher than expected. *Geophys. Res. Lett.*, **33**, L17811, doi:10.1029/2006GL026899.
- Williamson, D. L., and J. G. Olson, 2007: Comparison of forecast errors in CAM2 and CAM3 at the ARM Southern Great Plains site. *J. Climate*, **20**, 4572–4585.
- Wood, R., C. Bretherton, B. Huebert, C. R. Mechoso, and R. Weller, 2006: VOCALS-Southeast Pacific Regional Experiment (REx): Scientific program overview. WCRP/CLIVAR/VAMOS/GEWEX Rep., 40 pp. [Available online at www.atmos.washington.edu/~robwood/VOCALS/VOCALS_SPO_Complete.pdf].
- Xie, S., J. Boyle, S. A. Klein, X. Liu, and S. Ghan, 2008: Simulations of Arctic mixed-phase clouds in forecast with CAM3 and AM2 for M-PACE. *J. Geophys. Res.*, **113**, D04211, doi:10.1029/2007JD009225.
- Yu, S., and Coauthors, 2008: Evaluation of real-time PM_{2.5} forecasts and process analysis for PM_{2.5} formation over the eastern United States using the Eta-CMAQ forecast model during the 2004 ICARTT study. *J. Geophys. Res.*, **113**, D06204, doi:10.1029/2007JD009226.
- Zaveri, R. A., and L. K. Peters, 1999: A new lumped structure photochemical mechanism for large-scale applications. *J. Geophys. Res.*, **104**, 30 387–30 415.
- , R. C. Easter, and L. K. Peters, 2005a: A computationally efficient Multicomponent Equilibrium Solver for Aerosols (MESA). *J. Geophys. Res.*, **110**, D24203, doi:10.1029/2004JD005618.
- , —, and A. S. Wexler, 2005b: A new method for multicomponent activity coefficients of electrolytes in aqueous atmospheric aerosols. *J. Geophys. Res.*, **110**, D02201, doi:10.1029/2004JD004681.
- , —, J. D. Fast, and L. K. Peters, 2008: Model for Simulating Aerosol Interactions and Chemistry (MOSAIC). *J. Geophys. Res.*, **113**, D13204, doi:10.1029/2007JD008782.
- Zhang, L. M., S. L. Gong, J. Padro, and L. Barrie, 2001: A size-segregated particle dry deposition scheme for an atmospheric aerosol model. *Atmos. Environ.*, **35**, 549–560.