EARLY ONLINE RELEASE

This is a preliminary PDF of the author-produced manuscript that has been peer-reviewed and accepted for publication. Since it is being posted so soon after acceptance, it has not yet been copyedited, formatted, or processed by AMS Publications. This preliminary version of the manuscript may be downloaded, distributed, and cited, but please be aware that there will be visual differences and possibly some content differences between this version and the final published version.

The DOI for this manuscript is doi: 10.1175/2010BAMS2868.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.
The Aerosol Modeling Testbed

A Community Tool to Objectively Evaluate Aerosol Process Modules


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

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

E-mail: Jerome.Fast@pnnl.gov
Capsule:

A new computational framework to streamline the process of testing and evaluating aerosol process modules over a range of spatial and temporal scales
Abstract:

The current paradigm of developing and testing new aerosol process modules is haphazard and slow. Aerosol modules are often tested for short simulation periods using limited data so that their overall performance over a wide range of meteorological conditions is not thoroughly evaluated. While several model inter-comparison studies quantify the differences among aerosol modules, the range of answers provides little insight on how to best improve aerosol predictions. Understanding the true impact of an aerosol process module is also complicated by the fact that other processes such as emissions, meteorology, and chemistry are often treated differently. To address this issue, we have developed an Aerosol Modeling Testbed (AMT) with the objective of providing a new approach to test and evaluate new aerosol process modules. The AMT consists of a more modular version of Weather Research and Forecasting (WRF) model and a suite of tools to evaluate the performance of aerosol process modules via comparison with a wide range of field measurements. Our approach systematically targets specific aerosol process modules, while all other the processes are treated the same. The suite of evaluation tools will streamline the process of quantifying model performance and eliminate redundant work performed among various scientists working on the same problem. Both the performance and computational expense will be quantified over time. The use of a testbed to foster collaborations among the aerosol scientific community is an important aspect of the AMT; consequently, the long-term development and use of the AMT needs to be guided by users.
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 (IPCC 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 (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), is 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 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 due entirely 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 multi-year 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 has been haphazard. Only a few model inter-comparison studies have been conducted to examine uncertainties in regional-scale aerosol predictions (e.g. McKeen et al. 2007, 2009; Stern et al. 2008). These inter-comparison 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 3-D 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 in http://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 campaigns 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. Since the AMT is a community tool, it also provides a means of enhancing collaboration and coordination among aerosol modelers.

RELATIONSHIP TO OTHER TESTBEDS. Table 1 lists two other testbed activities related to modeling aerosols: the CCPP-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 Atmospheric Radiation Measurement (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 testbed 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 WRF model to simulate atmospheric conditions using grid spacings ranging from microscale to mesoscale applications (Δx ~ 50 m to 30 km), while 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. The Weather Research and Forecasting (WRF) model (Skamarock et al. 2005) was selected for the AMT to test and evaluate new aerosol process modules because its on-line coupling of meteorology and chemistry permits the simulation of aerosol radiative forcing (Grell, et al. 2005; Fast et al. 2006). Since 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 inter-comparisons 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, while changes in the AMT version originate from users adding or modifying aerosol process modules. Version control system software, called SubVersion (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 their own “branch” when making their 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.

**Testbed Cases.** The amount and type of aerosol property measurements are limited compared 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 semi-direct 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 (~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 NASA and DOE web sites that make their field campaign archives accessible by the public. Many different types of file formats are also employed, including ASCII, binary, NetCDF, 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 testbed 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 VAMOS Ocean-Cloud-Atmosphere-Land Study (VOCALS), and the Indirect and Semi-Direct Aerosol Campaign (ISDAC). CHAPS was conducted in Oklahoma to better under-
stand how fair-weather cumulus clouds affect the properties of anthropogenic aerosols (Berg et al. 2009). While 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://acrff-campaign.arm.gov/isdac/). The extensive measurements from NASA’s ARCTAS (http://www.espo.nasa.gov/arctas/) and NOAA’s ARCPAC (http://www.esrl.noaa.gov/csd/arcpac/) field campaigns are also included in this testbed 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://acrff-campaign.arm.gov/cares/) and the Research at the Nexus of Air Quality and Climate Change (CalNex) (http://www.esrl.noaa.gov/csd/calnex/) will be added to the AMT in the near future.

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

Analysis Toolkit. While software exists to evaluate meteorological predictions from WRF using operational monitoring data, there are no community tools that handle non-standardized 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 direct comparison of field campaign observed and simulated quantities, 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 in 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 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 in time to the aircraft data points, although other protocols could easily be employed to evaluate the model results. While 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 grid cell 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 consensus on protocols for statistical evaluations using field campaign data.

Graphical tools, based on Gnuplot freeware for platform compatibility, provide users “quicklook” plots of observed and simulated quantities, since 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 [http://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 Center 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, testbed cases (input files and observations), 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 testbed case is ~100 Gb. Some potential users may not be proficient in running 3-D 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 VS COMPLEX**

*Objective.* To demonstrate some of the AMT capabilities, we compare the performance of two aerosols models, MADE/SORGAM (Ackermann et al. 1998; Schell et al. 2001) and MOSAIC (Zaveri et al. 2008) previously implemented in WRF, for the MILAGRO testbed case.
MADE/SORGAM adopts a modal approach for the aerosol size distribution and an equilibrium treatment for gas-particle partitioning, while 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. 2005a,b). 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.

The simulation period for the MILAGRO testbed case is from March 6 to 30, 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, as well as 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 MOZART global chemistry model (Pfister et al. 2008) as well as the same treatments for photochemistry (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. Since 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 \nu_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 \sim 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 (Aiken 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 multi-day 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, while 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 SOA is ne-
neglected; however, the overall magnitude of simulated primary organic aerosols (POA) agree with derived POA from the AMS 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 further 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 March 19 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 mid-troposphere 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. While 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 scatter-
ing is highest for particle diameters around 0.4 \( \mu m \) 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 March 7 and 27 is shown in Fig. 8. Since 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 25 and 75\(^{th}\) percentiles from MADE/SORGAM is less than observed for particle diameters less than 0.625 \( \mu m \), while the range from MOSAIC is more consistent with the observations. Both MOSAIC and MADE/SORGAM underestimate aerosol number for bin 3 (0.15625 - 0.3125 \( \mu m \)), likely due to neglecting SOA in both models. Aerosol number from both models is too high for bin 5 (0.625 – 1.25 \( \mu m \)) but for bin 6 (1.25 – 2.5 \( \mu m \)) 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” de-
rives either simulated AOD at the same grid spacing and overpass times as the Terra (~17 UTC) and Aqua (~19 UTC) satellites or averages satellite measurements to the model’s grid cells. Figure 9 shows a composite of the Level-2 MODIS product (Levy et al. 2007) and the simulated AOD distributions at 550 nm from the 3-km domain for the Terra overpasses between March 8 and 29. White regions denote areas with either 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 Gulf of Mexico, similar to MODIS measurements. Scatter plots for these regions indicate the bias is lowest over the central plateau, while 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, where 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 an overall bias 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 March 12 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 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 HSRL observations, but the model’s vertical grid spacing cannot resolve the thin aerosol layers in the mid troposphere. Both models also frequently 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 - 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 was -6%, 0.73, and 0.83; 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 15 UTC (09 LT) from March 8 and 29 is shown in Fig. 11. Net shortwave radiation under clear skies without aerosols is 550 to 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 PM2.5 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 since 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. While 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 impact 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 18 UTC between March 8 and 29 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 since 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. March 18 – 20) 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. March 15). 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. March 19).

It is interesting to note where the largest differences occur when comparing MADE/SORGAM and MOSIAC (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 further downwind as shown in Fig. 12. While 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 March 19 the largest difference occurred close to the border with Texas when the ambient south-
erly 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. While some statistics for MOSAIC were somewhat better than MADE/SORGAM and visa versa, neither model out-performed the other over the central Mexican plateau. Larger differences in aerosol mass (mostly nitrate and aerosol water) between the two models were produced further 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 expense 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 at http://www.pnl.gov/atmospheric/research/aci/amt/index.stm. While 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 testbed 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 tested as listed in Table 2 so that aerosol process modules can be evaluated over a wider range of atmospheric conditions and aerosol sources. The website also includes documentation for users to create other testbed 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 testbed 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:

*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 SOA). The intent of the AMT is to incorporate those findings into 3-D 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, and most physicochemical processes, is 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. Since 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, e.g. the Program for Climate Model Diagnosis and Intercomparison funded by DOE’s Climate Change Research Division [http://www-pcmdi.llnl.gov](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, i.e. 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 of developing parameterization for use in those applications and in the long-term used directly in global climate models when sufficient computing power becomes available.

**ACKNOWLEDGEMENTS.** 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. The National Center for Atmospheric Research is operated by the University Corporation of Atmospheric Research under the sponsorship of the National Science Foundation.
REFERENCES


The time evolution of aerosol composition over the Mexico City plateau. *Atmos. Chem. Phys.*, 8, 1559-1575.


Volkamer, R., Jimenez, J.L., San Martini, F., Dzepina, K., Zhang, Q., Salcedo, D., Molina, L.T.,
Worsnop, D., and Molina, M.J.: Secondary organic aerosol formation from anthropogenic air

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.

Pacific regional experiment (REx): Scientific program overview. Available at

Xie, S., J. Boyle, S.A. Klein, X. Liu, and S. Ghan, 2008: Simulations of Arctic mixed-phase clouds

Yu, S., R. Mathur, K. Schere, D. Kang, J. Pleim, J. Young, D. Tong, G. Pouliot, S.A. McKeen, and
S.T. Rao, 2008: Evaluation of real-time PM2.5 forecasts and process analysis for PM2.5 forma-
tion over the eastern United States using the Eta-CMAQ forecast model during the 2004


coefficients of electrolytes in aqueous atmospheric aerosols, *J. Geophys. Res.*, 110,


Figure Captions:

Fig. 1. Schematic diagram depicting how the SVN software is used to manage changes in the WRF code distributed by NCAR and NOAA and code employed by the AMT.

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

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

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

Fig. 5. Sensitivity of total black carbon 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, correlation coefficient (r), index of agreement (IA), and bias (b) in \( \mu g \ m^{-3} \).

Fig. 7. Comparison of observed and predicted aerosol composition along the DC-8 flight on March 19. Observed composition obtained from University of New Hampshire’s Mist Chamber and filter and ion-chromatography measurements.

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

Fig. 9. Comparison of observed average AOD between March 6 and 29 and corresponding simulated values from the 3-km domain over central Mexico. Scatter plots show differences between observed and simulated AOD by region.
Fig. 10. Comparison of observed and simulated backscatter profiles and aerosol optical thickness (AOT) between the surface and ~9 km MSL for the B-200 flight on March 12. Right panels are profiles of the mean and standard deviation of extinction among all the B-200 flights over central Mexico and the Gulf of Mexico.

Fig. 11. Effect of aerosols on clear-sky downward shortwave radiation at 15 UTC averaged between March 8 and 29 produced by MADE/SORGAM (top) and MOSAIC (bottom) attributable to various aerosol composition. Negative values indicate cooling resulting from aerosols.

Fig. 12. Average difference in shortwave radiation (MOSAIC - MADE/SORGAM) at 18 UTC between March 8 and 29 (top panel) and instantaneous difference at 18 UTC March 19 (middle panel) in relation to aircraft flight paths during the MILAGRO field campaign (bottom panel).
Table 1. Comparison of the AMT capabilities in relation to the CAPT and AeroCom test-bed activities.

<table>
<thead>
<tr>
<th></th>
<th>AMT</th>
<th>CAPT</th>
<th>AeroCom</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model</strong></td>
<td>WRF</td>
<td>CAM</td>
<td>Multiple</td>
</tr>
<tr>
<td><strong>Domain Extent</strong></td>
<td>LES to mesoscale</td>
<td>Global and SCM mode</td>
<td>Global</td>
</tr>
<tr>
<td><strong>Primary Processes Addressed</strong></td>
<td>aerosol properties, cloud-aerosol interactions, aerosol radiative forcing</td>
<td>cloud properties</td>
<td>aerosol properties, direct radiative forcing</td>
</tr>
<tr>
<td><strong>Primary Data Used for Evaluation</strong></td>
<td>field campaign, operational</td>
<td>operational, field campaign, satellite</td>
<td>satellite, AERONET</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Field campaigns currently being developed as testbed cases for the AMT. The MILAGRO field campaign is completed, while the other campaigns are at various stages of development.

<table>
<thead>
<tr>
<th>Field Campaign</th>
<th>Time Period</th>
<th>Location</th>
<th>Processes Relevant to the AMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MILAGRO</td>
<td>March 2006</td>
<td>central Mexico and western Gulf of Mexico</td>
<td>aerosol aging, organic aerosols, optical properties, oxidant chemistry, megacity emissions, biomass burning</td>
</tr>
<tr>
<td>CHAPS</td>
<td>June 2007</td>
<td>vicinity of Oklahoma City, Oklahoma</td>
<td>effect of shallow cumulus on aerosol properties, urban emissions</td>
</tr>
<tr>
<td>VOCALS</td>
<td>mid October – mid November 2008</td>
<td>southeastern Pacific Ocean near Chile</td>
<td>cloud-aerosol interactions, anthropogenic point source and biogenic emissions</td>
</tr>
<tr>
<td>ISDAC / ARCTAS / ARCPAC</td>
<td>March - April 2008</td>
<td>North Slope of Alaska</td>
<td>effect of aerosols on mixed-phase stratus clouds, long-range transport of aerosols</td>
</tr>
<tr>
<td>CARES / CalNex</td>
<td>May - June 2010</td>
<td>vicinity of Sacramento, California for CARES; primarily Los Angeles area and San Joaquin Valley for CalNex</td>
<td>organic aerosols, aerosol mixing state, optical properties, urban and biogenic emissions</td>
</tr>
</tbody>
</table>
Fig. 1. Schematic diagram depicting how the SVN software is used to manage changes in the WRF code distributed by NCAR and NOAA and code employed by the AMT.
Fig. 2. Types of instrument simulators and data categories included in the Analysis Toolkit.
Fig. 3. Conceptual diagram depicting how the AMT would operate for the user community.
Fig. 4. Aerosol volume as a function of diameter for modal (lines) and sectional (gray boxes) size distributions.
Fig. 5. Sensitivity of total black carbon 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, correlation coefficient (r), index of agreement (IA), and bias (b) in µg m$^{-3}$.
Fig. 7. Comparison of observed and predicted aerosol composition along the DC-8 flight on March 19. Observed composition obtained from University of New Hampshire’s Mist Chamber and filter and ion-chromatography measurements.
Fig. 8. Comparison of observed and predicted aerosol number among the 13 G-1 flights between March 7 and 27 expressed in terms of percentiles.
Fig. 9. Comparison of observed average AOD between March 6 and 29 and corresponding simulated values from the 3-km domain over central Mexico. Scatter plots show differences between observed and simulated AOD by region.
Fig. 10. Comparison of observed and simulated backscatter profiles and aerosol optical thickness (AOT) between the surface and ~9 km MSL for the B-200 flight on March 12. Right panels are profiles of the mean and standard deviation of extinction among all the B-200 flights over central Mexico and the Gulf of Mexico.
Fig. 11. Effect of aerosols on clear-sky downward shortwave radiation at 15 UTC averaged between March 8 and 29 produced by MADE/SORGAM (top) and MOSAIC (bottom) attributable to various aerosol composition. Negative values indicate cooling resulting from aerosols.
Fig. 12. Average difference in shortwave radiation (MOSAIC - MADE/SORGAM) at 18 UTC between March 8 and 29 (top panel) and instantaneous difference at 18 UTC March 19 (middle panel) in relation to aircraft flight paths during the MILAGRO field campaign (bottom panel).