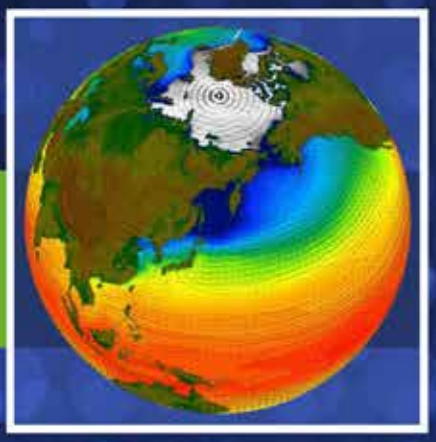
U.S. DEPARTMENT OF ENERGY | Office of Science

BIOLOGICAL AND ENVIRONMENTAL RESEARCH

Climate and Environmental Sciences Division

ATMOSPHERIC TESTBED WORKSHOP



DOE/SC-0163

BIOLOGICAL AND ENVIRONMENTAL RESEARCH

Climate and Environmental Sciences Division

ATMOSPHERIC TESTBED WORKSHOP

August 5-6, 2013

Convened by

U.S. Department of Energy
Office of Science
Office of Biological and Environmental Research

ORGANIZERS

Sally McFarlane
Atmospheric System Research

Dorothy Koch
Earth System Modeling Modeling

Renu Joseph
Regional and Global Climate Modeling

Wanda Ferrell
ARM System Research Research Facilities

Ashley Williamson
Atmospheric System Research

DISCUSSION LEADS AND WHITE PAPER CONTRIBUTORS

Jerome Fast
Pacific Northwest National Laboratory

Steve Klein
Lawrence Livermore National Laboratory

Yangang Liu
Brookhaven National Laboratory

William Gustafson
Pacific Northwest National Laboratory

Michael Jensen
Brookhaven National Laboratory

U.S. DEPARTMENT OF ENERGY | Office of Science

Contents

1.0 Introduction

A primary goal of the Climate and Environmental Sciences Division (CESD) is to improve the predictive ability of regional and global climate models (GCMs). One set of tools used for evaluating and improving model predictions and projections are testbeds that combine data derived from observations and model simulations. In these testbeds, outputs of model simulations can be compared to observations of the climate system in order to identify errors and/or biases in the model simulations and determine the specific model processes that need improvements. Testbeds can also be used to provide scientific insights into dominant processes and process interactions as well as increase our understanding of the role of various physical processes involved in a particular case study or meteorological event. CESD currently funds multiple activities, which are developing testbeds to evaluate and improve atmospheric model parameterizations across a range of models, scales, and processes. For each of these testbeds, the Atmospheric Radiation Measurement (ARM) Climate Research Facility represents a primary source of observational data for evaluating the models.

The Atmospheric System Research (ASR), Earth System Modeling (ESM), Regional and Global Climate Modeling (RGCM), and ARM program managers co-sponsored a workshop on atmospheric testbed activities across CESD. The workshop was held August 5-6, 2013, at the U. S. Department of Energy (DOE) Germantown facility. The principal objectives of the workshop were to survey the CESD testbed portfolios; identify the distinctive capabilities of each testbed as well as commonalities between the testbeds; identify areas for coordination and collaboration among testbeds; improve linkages between testbed activities and the ARM facility; identify gaps in testbed activities; and identify ways to improve coordination and communication among the testbeds and between the testbeds and the broader scientific community.

The workshop agenda is provided in Appendix A, and the list of attendees are provided in Appendix B. The workshop started with overview presentations on the first morning, then consisted of discussion sessions for the remainder of the meeting. Discussion sessions focused on overviews of the testbeds, data sets, visualization and analysis software, model frameworks, and coordination opportunities. In order to promote useful discussion at the workshop, leaders for each session were identified before the workshop and provided with several discussion questions (given in Appendix A and relevant sections below). The workshop ended with a discussion session to identify workshop outcomes and action items.

DOE will use the workshop as a first step toward making strategic improvements in testbed efficiency, e.g., where ARM and other observational data are more readily used to test and improve the representation of atmospheric processes in climate models.

CESD Mission Statement:
To advance a robust predictive understanding of Earth's climate and environmental systems and to inform the development of sustainable solutions to the Nation's energy and environmental challenges.

2.0 Testbed Descriptions

The workshop opened with each testbed project leader giving an overview presentation of their testbed. These talks were planned so that each attendee would have basic background information on each of the testbeds and could better identify coordination opportunities between the testbeds. The first discussion session focused on defining what is meant by a testbed, articulating the primary science goals of each testbed activity, and identifying distinctive and common elements of each testbed. This discussion served as the basis for identifying key areas of coordination in later discussion sessions.

The group defined a testbed as, *"a systematic, automated framework, involving a combination of model and observations, used to understand physical processes, and to evaluate and identify sources of error in a model during its development."* Although a given testbed may not be fully automated, it is expected that each testbed should be automated for at least some aspects of the model evaluation process. In addition, scientific analysis using testbeds should reveal whether a model produces physically meaningful answers; provide information to assess and quantify uncertainties; facilitate debugging; provide insights into model characteristics; and help with model calibration. Testbeds could include a tool to configure, run, and view output; include post-processing scripts; provide indications of how model changes would impact results; and evaluate whether model features resemble observations.

CESD currently sponsors four atmospheric testbeds, which each aid in developing and improving a range of atmospheric models spanning regional to global scales. Below we give a brief overview of each testbed. More details on each testbed, as provided by the testbed leads in their presentations and follow-up documentation, are given in Appendix C.

2.1 Aerosol Modeling Testbed

The Aerosol Modeling Testbed (AMT) is a computational framework for the atmospheric sciences community that streamlines the testing and evaluation of treatments of aerosol formation, transformation, and removal processes over a wide range of spatial and temporal scales (Fast et al. 2011). Current work focuses on: (1) evaluating the representation of secondary organic aerosol (SOA) in models; and (2) determining the importance of aerosol chemistry treatments in terms of aerosol number, size, and effects of hygroscopic properties on droplet and ice nucleation, and subsequently on cloud properties and precipitation.

The AMT consists of a fully coupled meteorology-chemistry-aerosol model, the Weather Research and Forecasting (WRF) model with chemistry (WRF-Chem), and a suite of tools to evaluate the performance of aerosol process modules through comparisons with a wide range of measurements collected during DOE field campaigns. 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. Meteorology, trace gas chemistry, initial and boundary conditions, and emissions are held constant while various aerosol schemes or aerosol process modules are varied. In this way, the differences in the simulations are due solely to the sensitivity of the assumptions employed by various aerosol models or modules, rather than other components of a coupled modeling system, such as meteorology, that arise during traditional model intercomparison studies. The

performance of new aerosol treatments can then be quantified and compared to existing treatments before they are incorporated into regional and global climate models.

While the focus of the AMT is on aerosols, it can also be used to test and evaluate parameterizations for meteorology (e.g., clouds) and trace gas chemical mechanisms. Since the AMT is a community tool that does not have to be used in a central computing facility, it also provides a means of enhancing collaboration and coordination among aerosol modelers. The primary users of the AMT are aerosol modelers at Pacific Northwest National Laboratory (PNNL), as part of their DOE-funded projects (Fast et al. 2014; Shrivastava et al. 2011, 2013; Qian et al. 2010; Yang et al. 2011; Ma et al. 2013). The AMT also has multiple users outside of PNNL, such as Alma Hodzic (National Center for Atmospheric Research [NCAR], Knote et al. 2013), Jennie Thomas (Laboratoire Atmosphères, Thomas et al. 2013), Joseph Ensberg (California Institute of Technology, Ensberg et al. 2013), Yang Zhang (North Carolina State University), and Zhijin Li (Jet Propulsion Laboratory).

The AMT has evolved into a form where model generated and observed data are efficiently organized for subsequent analysis. Referring to Figure 1 as an illustration, the AMT is used to test and evaluate new treatments of SOA over regional spatial scales using data collected during recent field campaigns conducted in California during May and June 2010. Treatments shown to improve model performance could then be implemented in GCMs over longer periods of time.

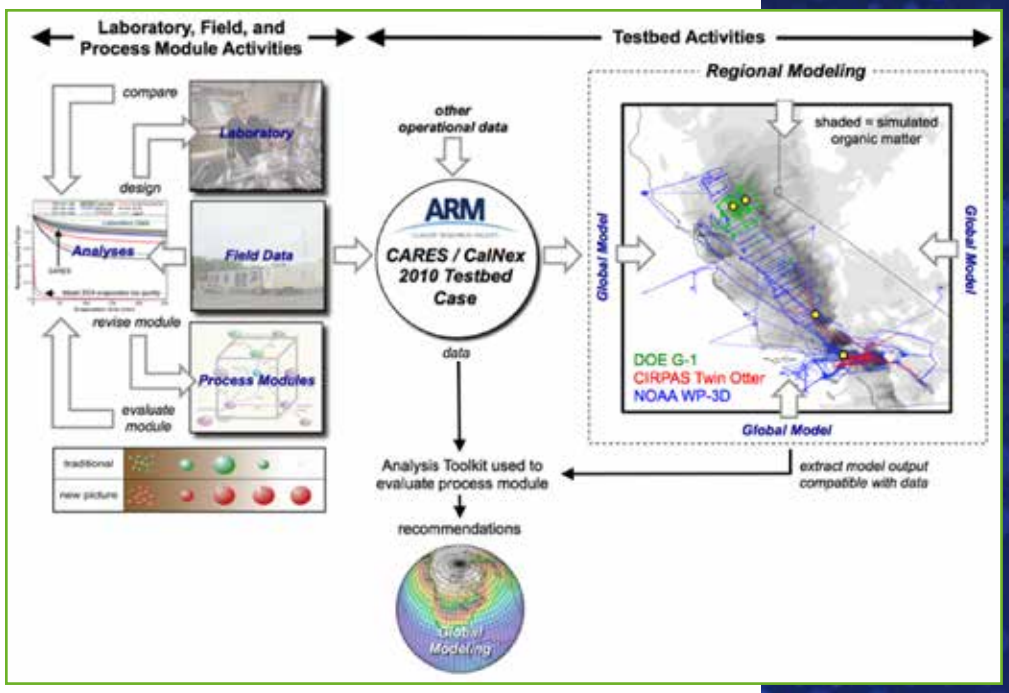


Figure 1. Schematic diagram depicting how the AMT could be used in conjunction with other CESD research activities.

A unique feature of AMT, relative to the other CESD testbeds, is its focus on aerosol microphysical and chemistry processes. AMT is also the only testbed that currently exists in a format that can be downloaded and run end-to-end on a user's own computer.

2.2 Cloud-Associated Parameterizations Testbed

The Cloud-Associated Parameterizations Testbed (CAPT) (**http://www-pcmdi.llnl.gov/projects/capt/index.php**) is a framework utilized at both Lawrence Livermore National Laboratory (LLNL) and NCAR that aims to: (1) identify the source of climate model problems primarily in the simulation of clouds, precipitation, radiation, and aerosols; and to (2) test new parameterizations for these processes through comparison to ARM and

other observations (Phillips et al. 2004). The defining characteristic of CAPT is the use of a hindcast technique in which climate model simulations are initialized with Numerical Weather Prediction (NWP) analyses and where the climate model is subsequently used to perform short-duration simulations (usually less than 6 days). This is similar to the approach used by weather-forecasting centers to evaluate NWP models. The hindcast technique is useful for climate model diagnosis because it: (1) facilitates evaluation with ARM site data, which is localized in space and time; (2) averages hindcast errors that in many circumstances are the same errors that the model makes in climate simulations; and (3) encourages a process-oriented evaluation of model simulations by examining variability at the hours-to-days time scales, which are characteristic of cloud processes that are not the monthly mean fields typically examined by climate modelers.

The primary users of the CAPT framework are developers of atmospheric model parameterizations, particularly developers of the Community Atmospheric Model (CAM). Over the years, the project has tested many parameterizations (primarily of cloud microphysics and convection) for inclusion in GCMs (e.g., Williamson et al. 2005; Hannay et al. 2009; Boyle and Klein 2010; Xie et al. 2012). While previous work included other GCMs (e.g., the Geophysical Fluid Dynamics Laboratory [GFDL] Atmospheric Model [AM2]), the CAPT project currently focuses solely on evaluation of parameterizations, mainly developed by external scientists, for the DOE/NCAR CAM. The project has compared simulations to observations from numerous ARM field campaigns as well as global satellite and analysis data. Often, individual studies are focused on specific cloud regimes, and the project has examined simulation quality in environments that range from tropical deep convection to Arctic mixed-phase clouds. The regimes of interest are chosen to coincide with an analyst's areas of expertise to maximize the scientific insight gained and to provide useful guidance for the parameterization developer. Currently, the CAPT project is developing the capability to diagnose the contribution of cloud and related processes to climate model biases that develop when an atmospheric model is coupled with an interactive land, sea ice, or ocean model. In addition, an automatic metrics and diagnostics package is under development, facilitating the comparison of model output with ARM, global satellite and reanalysis data, and the assessment of new physical parameterizations.

Unique features of CAPT, relative to the other CESD testbeds, are the inclusion of dynamical global climate feedbacks, the tracking and evaluation of error growth with time, and the ability to perform global model evaluation.

2.3 Climate Science for a Sustainable Energy Future

The Climate Science for a Sustainable Energy Future (CSSEF) project began in April 2011 and is a multi-institution effort involving eight DOE national laboratories (ANL, BNL, LANL, LBNL, LLNL, PNNL, ORNL, and SNL[1]) and NCAR. The overarching goal of CSSEF is to develop and test the "next-plus-one" generation of the Community Earth System Model, including developing variable resolution climate-system components, building an automated framework for building-testing-analyzing the climate system, and developing new methods of quantifying climate-system uncertainties. One component of the

(1) Acronyms are defined in Appendix E.

project, the CSSEF Atmospheric Testbed, aims to: (1) improve prediction of precipitation and the hydrologic cycle at regional scales; (2) utilize uncertainty quantification (UQ) techniques (e.g., perturbed-parameter studies) for sensitivity analysis and rapid model calibration with observations; (3) test new atmosphere parameterizations from precipitation and related processes in very high resolution models; and (4) develop an automated end-to-end testbed workflow to permit rapid testing and calibration of new model configurations.

The CSSEF testbed involves a unique modeling framework, namely, the regionally refined CAM with high resolution (typically 1/8 degree latitude) in a small-portion of a global model, which is usually at low-resolution (typically 1 degree latitude). So far, the CSSEF testbed's high-resolution portion is centered on the ARM Southern Great Plains (SGP) site. When the coarse grid region is nudged to analysis data (CAPT-like), predictions on the high-resolution grid may be compared in a deterministic way to ARM site and regional observations. The initial testbed simulations focused on spring and summer 2011 to compare with multiple observations from the ARM Midlatitude Continental Convective Clouds Experiment (MC3E) field campaign. Some data set development has been undertaken, most notably including the acquisition and quality control of the National Weather Service Next Generation Radars (NEXRAD) hourly precipitation for the eastern half of the United States at 0.1 degree spatial resolution for 2009-2011.

A significant component of the CSSEF testbed is the extensive exploration of the sensitivity of precipitation processes to parametric uncertainty. Large ensembles of climate simulations with CAM Version 5 at low-resolution (i.e., not the regionally refined model described previously) have been performed that perturb the uncertain parameters in the parameterizations of aerosol, clouds, and precipitation physics. Project participants have attempted to determine the parametric sensitivity of many characteristics of precipitation, including the mean, diurnal cycle, intensity distribution, and low-frequency phenomena, such as the Madden-Julian Oscillation (MJO). A challenging goal of this work is to move beyond sensitivity analysis to model calibration. Because precipitation simulations exhibit significant structural as well as parametric uncertainties, it is not obvious that parameter calibration, through a UQ technique, is useful in all circumstances. Some "idealized model" tests of parameter calibration techniques are being performed to explore aspects of model calibration. In these tests, model output is used in place of observations to calibrate the model, which removes the impact of structural or observational uncertainty and isolates the uncertainty in the calibration techniques themselves.

Relative to the other testbeds, CSSEF is unique in both its focus on a regionally refined global model as well as its use of perturbed-parameter studies for sensitivity analysis and model calibration.

2.4 Fast-Physics System Testbed and Research

As part of the Fast-Physics System Testbed and Research (FASTER) project (**http://www.bnl.gov/faster**), the FASTER Fast-Physics Testbed (FASTER testbed) aims to provide a web-based integrated platform to enhance and facilitate utilization of the detailed, long-term, high-resolution ARM measurements to evaluate and test parameterizations of cloud-related sub-grid processes (i.e., fast physics) in climate models. The FASTER project seeks to identify

and understand the couplings among processes and compensating errors due to inconsistent parameterizations. It also aims to identify the processes responsible for model biases. Particular parameterizations of interest include cloud microphysics, entrainment, turbulence and microphysics interactions, and radiation.

The current FASTER testbed consists of two major components that complement computationally expensive global GCM evaluations and capitalize on the long-term and high-resolution measurements at the ARM sites: a single column model (SCM) testbed and a NWP model testbed. Currently, the SCM-testbed features three versions of the Community Atmosphere Model (CAM3, CAM4, and CAM5), two versions of the GFDL GCM (AM2 and AM3), and the GISS ModelE. Users can compare different models, parameterizations and/or measurements against each other. The unique use of multiple models and multiple versions facilitates diagnosising parameterization deficiencies and tracking model development. The NWP-testbed is built on, and expands upon, the European Cloudnet concept (Illingworth et al. 2007) by taking advantage of routine NWP forecasts. The SCM-NWP integration allows researchers to comprehensively assess the performance of existing or newly developed fast-physics parameterizations, making use of not only rich ARM measurements but also a vast pool of NWP results. The integration of the NWP-testbed also offers a unique opportunity for evaluating NWP models in collaboration with major NWP centers.

Through the testbed, researchers can interactively and effectively evaluate and test their parameterizations against ARM observational data and other complementary measurements (e.g., Blossey et al. 2013; Li et al. 2013; Liu et al. 2013; Lu et al. 2013; Song et al. 2013). To aid in the evaluation and development process, a multi-regime case library is being developed to integrate and categorize available forcing and evaluation data. The regime-based approach serves to enhance the robustness of the evaluation, and isolate the challenges faced by cloud parameterizations.

In addition to the SCM and NWP models, the WRF model has been reconfigured to run as a cloud-resolving model (CRM) or large eddy simulation (LES) model, driven by the same forcing as the SCMs. This version of WRF, known as WRF-FASTER, is in the process of being integrated into the FASTER testbed. The three components (SCM, NWP, and CRM/LES) can each be used as stand-alone testbeds to address issues facing the corresponding communities. More importantly, the integration of WRF-FASTER can facilitate diagnosis of the physical sources of error in fast-physics parameterizations, provide high-resolution information unavailable with current measurements, and inform development of high-resolution climate models. The FASTER project has also developed a multi-scale data assimilation system that can use WRF or WRF-Chem as base models. This multi-scale data assimilation system generates hydrometeor and multi-scale forcing. Therefore, it improves model inputs. This system might also be used to produce a high-resolution reanalysis data set or “4D data cube” for parameterization development.

Unique features of FASTER, relative to the other testbeds, are the inclusion of LES scale models, the integrative multi-model evaluation framework, the online tool for interactive simulations with SCMs, and the multi-scale data assimilation component.

3.0 Testbed Data Sets

As discussed in the preceding section, the CESD-sponsored testbeds use a variety of models (Table 1). Given the variety of models, many workflows and data sets are employed, depending on the particular application. Meteorological states, used to initialize the models, can come from model analyses and reanalysis products. Some applications, such as testing of global model parameterizations at higher resolution, also use simulations from CAM5 to drive limited area model initial and boundary conditions. Additional data sets are often needed beyond the three-dimensional (3D) atmospheric state such as trace gas and particulate emissions, sea surface temperatures, and sea ice. The large-scale forcing data sets generated by ARM for its sites are also important for running SCMs and CRMs. More details of the testbed workflows, including model initialization procedures, are provided in Appendix C.

Table 1. Models currently used by the DOE testbeds.

Testbed	Global	Limited Area (specified boundaries)	Large Eddy (periodic boundaries)	Single Column
AMT		WRF		
CAPT	CAM5			SCAM
CSSEF	CAM5			SCAM
FASTER	output from NWP	WRF	WRF	GISS, GFDL, SCAM, WRF

NOTE: Acronyms are defined in Appendix E.

The testbed projects have assembled and collected a variety of data sets from both DOE observations and other sources (e.g., satellite, operational meteorological and air quality monitoring, and model analyses) for the purpose of model testbed data development activities and related research. Data sets are used both as initialization/boundary conditions for the models and to evaluate the model simulations. In the workshop, participants presented the key data sets developed for each testbed (Table 2), identified high priority future data sets, and discussed ways to improve the coordination of development and distribution of data products across the DOE community.

Table 2. Data sets used by each testbed.

Testbed	Data Sets for Initial and Boundary Conditions	Evaluation Data Sets	Existing Campaigns	Future Campaigns
AMT	Global analyses and reanalyses for meteorology, e.g., GFS, MERRA, ERA-40, CAM5; global chemical transport models for trace gases and aerosols, e.g., MOZART, CAM5	Aerosol microphysical, optical, and cloud nucleating properties, meteorological, cloud, and trace gas measurements from in situ and remote sensing instruments deployed on the surface and aircraft; satellite platforms	MILAGRO CARES BNL Aerosol IOP CHAPS ISDAC VOCALS	TCAP GOAmazon
CAPT	ECMWF operational analysis; ECMWF ERA-Interim, MERRA	ARMBE; variational analysis forcing and evaluation data; satellite cloud, precipitation, and radiative fluxes; two-dimensional (2D) gridded data set of surface fluxes, radiation, and soil properties	1997 Summer IOP MC3E 2003 Aerosol IOP M-PACE ISDAC TWP-ICE YOTC	AMIE MAGIC GOAmazon
CSSEF	Same as, or similar to, those used in CAPT	Precipitation metrics from ARM and NEXRAD; ARMBE with uncertainties; variational analysis forcing and evaluation data; moisture flux convergence and humidity profiles; stability indices; low-level jet	MC3E	AMIE
FASTER	ARM soundings; ARM variational analysis; FASTER data assimilation products; NWP analyses and reanalyses	High-resolution ARMBE; convective-stratiform partitioning, cloud fraction; hydrometeor classification; vertical velocity; convective indices; cloud condensation nuclei and aerosol measurements	2000 Cloud IOP 2003 Aerosol IOP RACORO MC3E AMF Point Reyes AMF China	MC3E MAGIC GOAmazon

NOTE: Acronyms are defined in Appendix E.

AMT data sets focus on the characterization of aerosol microphysics, optical, and cloud nucleating properties at the surface and aloft, using both surface- and aircraft-based observations collected by advanced instrument suites deployed as part of ARM/ASR field campaigns. In addition to these aerosol properties, the AMT data sets also include a range of meteorological, cloud, and trace gas measurements. These measurements are supplemented with available operational measurements from surface networks, NWP products, and satellites.

The CAPT team works closely with the ARM Infrastructure Team at LLNL, which develops data sets for the ASR modeling community, including the ARM Best Estimate Data Product (ARMBE) and the variational analysis model forcing and evaluation data sets. In addition to these data sets, CAPT archives and processes several analysis/reanalysis products specifically for initializing and evaluating CAM5. Satellite-derived data sets of precipitation, clouds, and radiative fluxes are also reformatted and archived for model evaluation. The CAPT team will work with the LLNL ARM Infrastructure Team to develop a 2D gridded data set over the SGP network that will include surface fluxes, surface radiation, and soil properties for land-atmosphere coupling studies.

The CSSEF data set development has focused on hydrological metrics, including hourly rain accumulation and the spatial patterns to the diurnal cycle statistics (phase, amplitude), intensity distribution (e.g., 95th percentile) of precipitation from NEXRAD and ARM precipitation radars. They have also produced a version of the ARMBE product that includes uncertainty estimates for surface meteorology data. In addition to these derived data products, the CSSEF project has archived various products for model evaluation purposes, including reanalysis output for moisture flux convergence and humidity profiles, radiosonde data products, diabatic heating, and stability indices from ARM radiosondes and the Modern-Era Retrospective Analysis for Research and Applications (MERRA) reanalysis. CSSEF is currently analyzing wind profiles from radiosondes and Doppler lidar observations to characterize the low-level jet, which may become a data product. Future plans involve the development of similar data products for the ARM MJO Investigation Experiment (AMIE) field campaign.

The FASTER project investigators have developed a suite of cloud, precipitation, aerosol, and atmospheric state data products focused on the ARM sites. The long-term data sets include a cloud condensation nuclei power law fit spectral analysis (at SGP); convective indices; a version of the ARMBE product that includes higher time-resolution averages and higher order moments (skewness, kurtosis); convective-stratiform rainfall partitioning; and estimates of cloud fraction derived from a variety of instruments. Additional data sets have been constructed for specific IOPs, including hydrometer classification and vertical velocity (MC3E); an extensive cloud, aerosol, and atmospheric state product interpolated to the GISS GCM model grid for SGP (2003 Aerosol IOP, RACORO) and AMF1 (Point Reyes; China); a set of comprehensive radar and surface-based observations for the RACORO field campaign; and an assembly of Environmental Protection Agency observed chemical speciation data over the SGP region for data assimilation studies. New data products under development include an estimate of the vertical profile of entrainment in convective clouds and simultaneous retrieval of albedo, cloud optical depth, and cloud fraction.

Participants identified additional data needs, including both data sets that could be assembled from current observations and new measurement needs. In particular, a number of potential new data sets made possible by new ARM radar capabilities were proposed, including:

- overlap statistics of cloud/rain
- cloud-scale vertical velocity
- depth of the rain layer
- height and depth of clouds (including frequency of occurrence)
- cloud condensate phase

- total raining volume
- frequency of occurrence of rain
- covariance of rain and cloud
- identification of melting layer height.

In addition, the collection and analysis of the chemical composition of rainwater was suggested as an important measurement for determining the type of aerosols being scavenged.

Given the effort needed to develop new data sets, there is a need to improve coordinating development and distribution of data products across the DOE community, including awareness among the testbed projects of the data development projects underway and planned, as well as communication of CESD testbed activities to the broader community. One avenue for improved communication among the testbed projects is through the ARM translator team. Each of the testbed projects includes a member of the ARM translator team who plays a significant role in the data development activities. CESD testbed data development activities could be discussed during ARM translator teleconferences on a regular basis in order to coordinate activities and leverage expertise. Currently, some of the testbed projects have included their data in the ARM Data Archive (AMT, CAPT), while other testbeds maintain their own local archive (CSSEF, FASTER). In order to improve coordination with the ARM Data Archive, participants suggested that mature testbed data products be shared with the ARM Data Archive as principal investigator (PI) data products (**http://www.arm.gov/data/pi**). PI data products are data sets developed by PIs within the ARM community that are distributed through the ARM Data Archive. If they are seen as broadly useful to the ARM community, these PI products may be elevated for value-added product activities by the ARM Infrastructure Team.

The release of CESD testbed products, as ARM PI data products, will increase awareness of testbed data development activities within the larger DOE and science communities.

4.0 Software/Visualization

Workshop participants described existing testbed software, discussed future plans for analysis and visualization tools, and identified capabilities that could be shared with other testbeds. Software tools associated with testbeds include: (1) programs and scripts to download and interpolate large-scale data to provide initialization and boundary conditions for model simulations; (2) scripts to run the models in certain configurations, such as hindcasts; (3) programs to process simulated output and observational data sets into comparable formats and variables; (4) tools to perform UQ analysis; and (5) software to visualize the model/observational comparisons. Testbeds varied widely in their range of automation, their visualization capabilities, and the types of software used.

4.1 Aerosol Modeling Testbed

The AMT provides Perl scripts that automatically extract model output compatible with observations and produce "quick-look" plots that compare observed and simulated meteorological, chemical, and aerosol quantities. The graphics are based on Gnuplot, a

publically available software program easily installed on most Linux-based systems. Currently, the AMT can generate plots for data types that consist of time series, such as those for surface and aircraft measurements. The model and observations are compared using plots of time series, scatter plots, and box-and-whisker plots for percentiles. There is currently no quick-look plotting capability for radar, profiles (e.g., radar wind profiler, radiosonde, lidar), or satellite measurements, but it would be useful to add this capability. No attempt was made to create more sophisticated graphics, since users have a wide range of graphics software preferences for publication-quality plotting. In addition, the text files of compatible observed and simulated quantities can easily be read by a wide range of existing graphics software for analyses and presentations that address the needs of the user. While the script runs programs in serial, some parallelism can be introduced by running various functions of the "analysis toolkit" at the same time on different processors.

Ideas for future improvements to the AMT software tools include:

- completing the suite of "quick-look" plots to include those compatible with profile and satellite measurements. The AMT also needs radar simulators for visualization, particularly ones consistent with the new ARM radars.
- changing the format of the observational testbed cases from ASCII to netCDF and tailoring the format for the storage access used by new supercomputers
- extending the AMT interface from WRF-Chem to other models, including CAM5. An interface with CAM5 would permit direct comparisons of global model predictions with field campaign data. Some considerations are needed to process the testbed case data to be more compatible (temporal and spatial averaging) with coarse GCM output. In addition, the variability of high-resolution measurements (e.g., aircraft) could be used to test sub-grid scale parameterizations being developed for global models.
- developing a formal methodology of storing model simulations and testbed analyses for documenting improvements associated with parameterization development over time. Such an archive would also enable other data-mining activities of the simulation results from other investigators, similar in concept to those associated with the Coupled Model Intercomparison Project (CMIP).

4.2 Cloud-Associated Parameterizations Testbed

CAPT currently uses Ultrascale Visualization Climate Data Analsyis Tools (UV-CDAT); a Python based visualization tool; Python scripts; and Fortran programs for downloading and interpolating data to provide initial and boundary conditions, performing hindcast simulations, and processing simulation outputs. The NCAR component of CAPT is currently making the CAPT code available on the CAM trunk, making it easily accessible to all researchers. CAPT currently has no automated quick-look or visualization tools. Scientists use their own software packages to graph the output metrics and diagnostics. Future plans include development of an automatic diagnostics package that will compare a large ensemble of hindcast simulations to ARM and global satellite and reanalysis data. This will permit rapid diagnostic assessment of the effect of new atmospheric parameterizations on simulation quality. The package will involve the calculation of both process-oriented diagnostics that help one to understand the sources of model errors, as well as metrics to provide a quantitative score used to judge relative model performance. A website will also be developed to display CAPT results.

4.3 Climate Science for a Sustainable Energy Future

A unique aspect of the CSSEF testbed is the implementation of UQ techniques, which involve distinct software for model simulation and analysis. An example simulation tool is LLNL's "UQ pipeline," which facilitates the generation and management of ensemble simulations. This includes the generation of perturbed-parameter values with characteristic sampling distributions (e.g., Latin-hypercube), as well as the scripting environment for performing ensemble simulations and archiving model output. Many UQ analysis tools are also used, such as tools to generate surrogate models fitted to simulation results, facilitating both sensitivity analysis and model calibration. Future plans for software development within CSSEF include development of an automatic diagnostics package that compares a large ensemble of hindcast simulations with the regionally refined CAM to ARM and regional satellite and reanalysis data. This diagnostic effort has been focused on precipitation and related processes from the spring 2011 ARM SGP MC3E field campaign. The automatic diagnostics package will permit rapid assessment of the effect of new atmospheric parameterizations on simulation quality.

In addition, there are plans to automate the CSSEF testbed workflow (from code definition, through simulations, through diagnostic assessment, and eventually calibration) in collaboration with computational scientists. This will likely involve some of the following items: workflow capture, model provenance, data publication, and embedding UQ techniques within the workflow.

4.4 Fast-Physics System Testbed and Research

FASTER has the most advanced visualization capabilities of the testbeds discussed here. The FASTER testbed uses server-client style communication for model configuration and simulations. The client (user) end uses Javascript language and Ajax technology to gather user inputs and communicate with the server, while the server end uses Perl script to process user inputs, control SCM simulations and post-processing, including generation of quick-look plots. Perl-based thread-parallel control in a shared memory computing platform is used to handle multitasking, such as ensemble simulations or concurrent reinitialized short-term simulations over an extended period of time. A separate, but similar, set of scripts is also available to automate various SCM simulations without the use of a web interface.

In addition to the instant quick-look plot generation using NCAR command language (NCL) scripts, a Java-based 3D data analysis and visualization toolkit has been developed. This will fulfill the need for easy-to-use, platform-independent, web-deployed, dynamic interactivity, with large volumes of 3D observational and model data, at a range of scales. It is complemented by collocated 2D and in situ data. The toolkit is designed to allow users to load 3D and four-dimensional (4D) gridded netCDF files with a few clicks. Data is rendered in synchronized panels so that the user may rotate, zoom in and out, and slice 3D volumes in real time. Plots and statistics are dynamically generated, and available for export as the user interacts with the data. An alpha version of the application, along with a short tutorial and sample data, has been released.

A multi-dimensional data analysis and visualization system has also been developed, in collaboration with computer scientists at BNL and Stony Brook University. The system

incorporates innovative methods of visualizing multi-variable data, including parallel coordinate and dynamic scatter plots, and can effectively grasp the complicated associations among environmental variables, control processes, physical properties, and metrics of interest. This concept of multi-dimensional analysis and the supporting visualizations can enhance the exploratory research of parameterization problems that involve known and unknown dependencies of parameterized physics on variables at the resolved-scale, as well as the interdependencies among the parameterized ones.

Figure 2. A screenshot of the FASTER 3D visualization and analysis toolkit.

The future visualization plans will focus on refining the two advanced visualization systems described previously, integrating them, and implementing them in the online testbed. Specifically, the FASTER project plans to refine the 3D Java-based toolkit to allow for additional file formats, render isosurfaces, and improve plotting features; implement functions to overlay collocated aircraft trajectories; integrate multi-dimensional data visualization through parallel coordinates and dynamic scatter plots into the 3D toolkit; and deploy the integrated visualization system in the online testbed.

4.5 Sharing Tools

Due to the differing software, data formats, and model interfaces used by each of the testbeds, direct sharing of most software and visualization tools is not possible without additional software development efforts. However, participants identified several tools and methods that could potentially be shared among testbeds. This includes:

- sharing scripts for developing initial and boundary conditions from large-scale reanalysis data sets or forcing data sets for case studies already developed
- sharing scripts and/or documentation for running CAM in particular configurations, such as hindcast mode
- sharing existing and future CAM diagnostic package software between all testbeds using CAM
- sharing UQ techniques and tools with other testbeds
- applying website ideas and tools from other testbeds to display CAPT results
- processing model and observational data into comparable variables using methodologies and best practices.

In addition, some of the tools identified as common needs among testbeds (e.g., radar simulators, methodology for archiving testbed results) could be developed to be applicable to all testbeds.

5.0 Invited Presentations: What Do Modelers Want in Testbeds?

In order to get some perspective from the members of the scientific community not directly involved in the CESD testbed activities, two modeling scientists (a cloud-process modeler and a global-climate modeler) were invited to attend the workshop and provide input on what climate modelers would like to see in model testbeds.

Professor Vince Larson (University of Wisconsin, Milwaukee) discussed testbed needs from a process-modeling perspective. He noted that a key issue in model development is not determining that there are errors in the model, but instead determining the source of these errors. In order to do this, testbeds need to have not just metrics, but diagnostics that go beyond a single variable. For example, knowing that cloud fraction is incorrect in a model is not enough to diagnose why it is incorrect. Additional variables, such as the terms of the moisture budget, are needed to understand why the model is under- or over-producing clouds. Two major problems with using observations to evaluate models are that: (1) the most relevant variables needed to model processes are often not measured; and (2) for those variables that are measured, the sampling density is generally inadequate. Professor Larson also emphasized that observations of the inputs to the parameterization (not just the outputs) are needed. This will allow modelers to determine whether the error is in the formulation of the parameterization or in the forcing of the parameterization. Model developers would really like a 4D data cube," which contains all relevant variables at all locations and times.

Professor Larson proposed using a high-resolution LES model as an intermediary between the observations and SCM simulations to produce this 4D data cube. The LES could be run in continuous mode over an ARM site using the same forcing data set as the SCM, so they could be directly compared. The LES could then be compared to observations. If there were large discrepancies, the forcings and input data sets could be adjusted. When the LES and observations are in good agreement, the LES output could then be used as metrics and diagnostics for the SCM parameterizations. Going a step further, the LES could be combined with the ARM observations through data assimilation techniques to produce a high-resolution reanalysis product that could be used for parameterization evaluation and development. A reanalysis product is essentially a best fit of the numerical model to the available data, taking into account errors in both the model and the data. Additional recommendations were that more efforts should be put into data quality, uncertainty, documentation, and visualization to make data sets more user-friendly for model developers.

Dr. Phil Rasch (PNNL) discussed testbed needs from a global modeling perspective. He emphasized that the key purpose of testbeds is to aid in the development stage of modeling research. If a model is already fully developed, then nothing needs to be tested. Therefore, testbeds must be used to determine if a model is producing predictable, physically meaningful answers; facilitate debugging efforts; and provide insight into characteristics of the model. Testbeds must also be treated as a tool to help identify the best values for adjustable parameters within the model. Dr. Rasch described the workflow of the typical GCM development process and gave examples of how observations are currently used within that process. Additional needs of current testbeds were presented, including more metrics

and diagnostics associated with extremes, capability to evaluate coupled model simulations, and development of a 'scorecard' for each model or version, allowing objective evaluation of the model's fidelity and improvement. It was also noted that archiving testbed output was needed so model versions could be compared to each other more easily. Sensitivity to perturbations of input data or parameter values could be better understood.

Dr. Rasch noted that the role of each of the existing CESD testbeds, their overall goals, and the processes they are evaluating should be clearly identified. He also noted that the end users of the testbeds should be clarified. Are testbeds intended to be used by model developers or by the broader scientific community to evaluate the model output? Are the testbeds tools to be provided to the community for them to run themselves? Or should the testbeds be run only by the testbed teams and their results provided as a service to the community? If the testbeds are intended to be a tool for others, then effort needs to be put in to make them more user-friendly. This would include documentation, easy porting to new platforms, and tools to assess that the software is working correctly. The testbeds should also make more use of metadata and provenance tools.

6.0 Needs and Gaps in Current Testbeds

The current testbed capabilities successfully meet many needs of model developers by enabling quick comparison of model output with observations, access to visualization tools, and automated workflows. However, workshop participants identified several areas as deserving of further attention so that the testbeds might better serve the model development community.

6.1 Methodologies and Best Practices for Using Testbeds

A particularly difficult problem in model development is separating model output errors into deficiencies caused by input errors versus those from particular model components or processes. For example, when clouds in a CRM differ from the observed clouds, is this difference because the forcing data set did not adequately capture the large-scale meteorological state? Because the surface fluxes were incorrect? Because the clouds were not represented properly in the radiation component, leading to incorrect cloud radiative cooling? Because the autoconversion formulation within the microphysics component was not accurate enough? Or did the problem arise from another part of the model? Given sufficient observations, many of these problems can be constrained. However, work is needed to identify what specific observations are required, and possible to acquire, to constrain each process of interest. Further improvement and application of uncertainty characterization techniques would also assist in narrowing down sensitivities within the model. A particular need identified in this area is the ability to identify compensating errors within the model to help confirm that the model produces the correct answers for the correct reasons, or if deficiencies in one part of the model mask deficiencies in a different part.

A methodological issue raised during this portion of the workshop is how to better use the range of testbeds in the model development process. The general idea is to start with an SCM to test a model component in a very constrained environment. Once the desired behavior is

achieved in the SCM, the component can be tested in limited area models where some of the degrees of freedom in the meteorological state are relaxed. The large-scale is still constrained, but mesoscale and smaller feedbacks are allowed to occur between the tested component and the other components. Next, the new component is implemented in an atmospheric global model where all spatial scales can influence, and be influenced, by the component behavior. Finally, tests are done in a coupled atmosphere-ocean climate model where the final boundary constraints are removed, allowing for full climate prediction. While this is straightforward in principle, it is much more complicated in practice. Best practices should be developed to make this process efficient, and the value added at each step needs to be more clearly elucidated.

Overcoming many model deficiencies requires additional data to both understand physical processes as well as constrain model behavior. While this workshop was not tasked with identifying specific data products, several broad categories were emphasized that are of particular need. These include concurrent in situ observations of cloud and aerosol properties to better understand cloud-aerosol interactions, and concurrent observations of cloud and precipitation properties to improve upon the current limitations in cloud data (detailed cloud properties are often available when the cloud is not precipitating, but more limited cloud data is available when precipitation is present). These two categories are of particular importance due to the synergy for studying clouds and cloud-aerosol interactions with the suite of DOE testbeds. Need was also expressed for additional soil measurements to supplement the existing ARM products—this data is needed to address the land-atmosphere interaction research highlighted below.

For aerosols, a zero-dimensional box model testbed would be useful to evaluate aerosol thermodynamics and microphysics treatments in relation to laboratory data. Much work is being done by ASR in this area, but is not coordinated as a testbed activity.

6.2 Instrument Simulators

To enable better comparison between observations and models, the testbeds also require simulator codes that allow the models to simulate the observed quantities. An existing example of this type of simulator is the Cloud Feedback Model Intercomparison Project Observation Simulator Package (COSP) software. Additional software development is needed to provide simulators for ARM lidar instruments and the new scanning cloud and precipitation radars. Satellite simulators are also needed (e.g., to simulate aerosol information from satellite lidar measurements) to aid in the development of these simulators, the testbeds need to identify specific use scenarios and required variables for the particular problems being solved. Of particular importance is differentiating the type of simulators needed for model development versus what is needed by radar scientists.

6.3 Data Formats and Unstructured Grids

All the testbeds use different formats for their evaluation measurement data. It is possible that using common data formats may reduce duplication of effort and facilitate sharing of data sets, although this would require reformatting existing data sets and modifying the testbed ingest routines to use them. It is not clear how much duplication of effort in data set

creation and processing is being conducted at present, but since there are only a few instances of common cases, it does not seem likely there has been much duplication of effort so far. It was also noted that while FASTER and AMT employ quick-look plots, it would be useful for CAPT to include this capability for users.

An important issue that will affect the operation of testbeds in the future is the use of unstructured grids by climate models. Changing the grid structure will likely affect software employed to facilitate model-data comparisons. There also needs to be standardization of tools for easier visualization of simulations that employ unstructured grids. This issue affects the larger modeling community, and not just the testbeds. It was suggested that a representative from the appropriate DOE national laboratories be contacted to discuss existing activities and software for sampling model output on unstructured grids.

6.4 Data Assimilation

Data assimilation is a tool that has had some use in the FASTER testbed, but could also be used to great advantage in future testbeds. Data assimilation has been attempted by NCAR scientists using the CAPT testbed, but more knowledge is needed to understand how the errors corrected by the assimilation process change the model's internal error characteristics. A potential advantage of the Data Assimilation Research Testbed (DART; Anderson et al. 2009) as applied in the Community Atmosphere Model includes the ability to provide an ensemble of initial conditions more consistent with the model uncertainties. This advantage over CAPT's standard procedure of providing a single initial condition from the analysis of a foreign model must be weighed against downsides, including: (1) DART is costly and complex to use; (2) DART typically uses a less complete set of observations than those available to NWP centers, thus producing lower quality analyses; and (3) some model errors may be hidden because model errors are retained in the data assimilation process. Nonetheless, contacts between the DART and CAPT teams do exist and it is likely that in the future there will be a comparison of the strengths and weaknesses of the different model initialization procedures taken by DART and CAPT.

Currently, the AMT testbed does not use data assimilation. The AMT "spins up" its aerosol fields based on a combination of a 3D spatial field acquired from a global chemical transport model and basic initial assumptions that evolve over the first hours or days of the simulation. Assimilation of aerosol observations could improve initial conditions for the model simulations. However, existing aerosol assimilation methods focus on diagnostic variables, such as aerosol optical depth, and require many assumptions to link the optical depth to the underlying aerosol characteristics, thus offering limited benefit over the existing initialization method. Development of new methods to assimilate 3D aerosol observations in a manner consistent with the representation of particular species and size distributions in the model would be useful to the AMT.

Related to data assimilation, an interest was expressed to be able to use assimilation techniques with LES. This would be needed to enable a "4D data cube" approach to filling the data void between ARM observations by using a very-high-resolution model in a way that is fully consistent with the available observations. Part of this process would involve the need to identify the optimized siting of instruments within the overall ARM site. Optimal

siting would involve identifying scientific questions that require overlapping measurements versus a distributed configuration that would provide more spatial detail. It could also involve the use of adjoint techniques to show the value added by particular instrument locations versus alternate locations.

6.5 Land-Atmosphere Interactions

All of the CESD testbed projects discussed in this report were originally developed to examine atmosphere-only processes. However, a greater understanding of the importance of land-atmosphere interactions on climate has driven an increased focus on related processes by CESD funded projects, including the testbeds. The mesoscale NWP community has known for decades that properly initializing soil moisture and thermal properties has an impact on the accuracy of weather forecasts. As climate models begin encroaching on the spatial resolution used for weather prediction, the same issues affect the simulated climate. Biases in the soil moisture and temperature impact fluxes between the land and atmosphere, potentially reducing accuracy of the simulated climate. Vegetation and the characterization of soil features also impact emissions of trace gases and particulates. The treatment of land-atmosphere interactions is seen as a current gap in the CESD testbed portfolio. However, several of the testbed projects have plans to address this gap in the near future.

Because climate biases in near surface air temperature over land can be the result of land-atmospheric interactions, in addition to separate errors, in either the atmosphere of land models (Klein et al. 2006), CAPT is currently working on the central United States summertime warm bias through a new joint ASR/GASS project: Clouds Above the United States and Errors at the Surface (CAUSES). (GASS is the Global Atmospheric Systems Studies project focused on improving the representation of cloud, convection, and precipitation processes in atmospheric models.) The major focus is to evaluate the role of cloud, radiation, and precipitation processes in contributing to the surface temperature biases using the short-term hindcast approach. The CAPT team will lead the effort, analyzing simulated precipitation and surface energy budget. A team from United Kingdom's Met Office will focus on the errors in clouds and radiation. The science questions that will be addressed by the CAPT team in this study are: (1) what is the relative contribution of precipitation errors to the temperature errors? (2) Does the atmosphere provide the correct amount of precipitation for the soil? Which type of precipitating convection systems dominate the errors in the surface precipitation? Does the surface energy balance reveal signs that evaporation is underestimated due to the lack of soil moisture? Through the close collaboration with other domestic and international research groups, it is expected to obtain a better insight into this long-lasting problem seen in many weather and climate models.

Given the aerosol focus of AMT, it is particularly interested in improving the surface trace gas and particulate emissions. A current area of research is on SOAs. A large proportion of the precursor trace gases for these organic particles are emitted from vegetation. A full representation of this process could be as involved as using dynamic vegetation combined with a detailed soil model, or as simple as an empirical relationship between available soil moisture, sunlight, and 2-m temperature. The level of detail needed is an open area

of research that involves both understanding the plant ecology and the necessary level of trace gas speciation required to adequately represent the physiochemical processes of SOA formation in the atmosphere.

Land-atmosphere interactions are also important to science goals of the CSSEF testbed. Using the regionally refined grid methodology available with the CAM Spectral Element dynamical core, CSSEF will investigate the impact of regionally refining the land model within the refined atmospheric region versus using a uniform, coarse land model. This will help elucidate the impact of small-scale eddies generated by heterogeneity in surface fluxes. Potential feedbacks exist between the land and precipitation characteristics, such as precipitation extremes that might be driven by a small-scale convergence due to surface heterogeneity. Interest was also expressed by CSSEF to investigate small-scale influences over ice covered surfaces. These studies will ultimately help answer the question of whether biases in the fluxes are due to problems in the atmosphere or the land component of the model, which will help guide further model development priorities.

The FASTER project also has begun exploratory work on the evaluation of surface-flux parameterizations by performing offline comparisons of commonly used parameterization schemes against long-term surface flux measurements over the ARM SGP site. Several deficiencies, especially in parameterized sensible and latent heat fluxes, have been identified (Liu et al. 2013). Although current SCM and CRM/LES simulations are constrained by measured surface fluxes to focus on atmospheric processes and minimize potential problems associated with land-surface representations, it is straightforward, in principle, to run these models with coupled surface modules and evaluate them against surface-flux measurements (or other related measurements). Such an extended framework provides a unique testbed for investigating the land-surface representation and land-atmosphere coupling, which are essential to getting atmospheric processes right. In the future, FASTER will place increased focus on surface fluxes and land-surface-atmosphere interactions by running coupled models and comparing simulated surface fluxes against measurements.

Given the broad interest in understanding land-atmosphere interactions, ARM has started to develop necessary land data sets to support such studies. The ARMBELAND data set at its SGP site has been released by the ARM Data Archive. A more comprehensive 2D gridded surface data set over the ARM SGP surface network is being developed.

6.6 Documentation

The need for better documentation from the testbeds was expressed by workshop participants. Each testbed has a range of workflows with varying levels of documentation. The AMT has a user's guide and overview documentation for published field campaign data sets. CAPT recently recorded their overall workflow to assist training new staff on the CAPT team. However, a new user often needs to interact with the testbed developers to fully take advantage of the tools. If DOE desires to make the testbeds more widely used within the greater scientific community, development of adequate documentation must be made a priority.

7.0 Collaboration/Coordination Opportunities

Participants discussed potential common science interests among the testbed projects that they could work on collaboratively. Three potential science topics were discussed: cloud-aerosol interactions, hydrological cycle, and land-atmosphere interactions.

Because each testbed is designed to address science issues relating to either clouds or aerosols, cloud-aerosol interaction processes was a likely candidate for possible collaboration. Each testbed could test a candidate parameterization for a climate model using their particular methodology. Parameterization performance could then be compared and contrasted among the testbeds, providing more robust conclusions regarding the advantages and disadvantages of the parameterization.

Since DOE climate programs are moving to emphasize the hydrologic cycle as a "grand challenge" for climate models, it would be useful to study the partitioning of stratiform and convective precipitation in models and how it affects the hydrologic cycle. The role of and mechanism by which testbeds reduce uncertainties in the hydrological cycle has not been articulated, however, and there are many processes that affect the hydrologic cycle that could be studied. There may also be field campaign cases in which all the testbed tools and utilities could be applied and demonstrated. The GOAmazon experiment to be conducted in 2014 is a possible candidate since the planned measurements will address a wide range of issues associated with clouds, aerosols, cloud-aerosol interactions, and land-atmosphere interactions.

Workshop participants noted that before coordinating testbeds around a single case and/or parameterization, the value of such an activity will need to be articulated in detail. It was also apparent that additional discussion is needed regarding coordination of science issues among the testbeds, so it was suggested that there be breakout meetings at upcoming science meetings such as the ASR or Modeling PI meetings.

7.1 Service to Broader Community

Participants discussed whether the testbeds should be viewed as a tool or a service to the community. They came to a consensus that the current testbeds are research tools and should be viewed as tools, rather than a service. There are cases in which the individual testbeds may provide services to individual scientists, but that depends on the available funding to provide such a service. Each testbed group currently interacts with scientists and responds to university calls for proposals on an ad hoc basis. While the AMT currently works with individual scientists to tailor the testbed tools for their specific application, the AMT would like to form larger, more coordinated efforts around a few important science questions. Such an effort is hampered by coordination among various projects with different timelines. FASTER has a vision of using the SCM version of CAM for educational purposes. They would also like to make the SCM easier to use for the community, and some efforts are underway. FASTER would like to use a SCM that employs the same dynamic core that will ultimately be used by a DOE climate model. CAPT has a long history of working with key developers of CAM to test and assess candidate parameterizations during the model

development process. They would like to continue this effort to help future community or DOE climate model development.

To better communicate the objectives of the testbeds to the larger scientific community, it was suggested that a central web page from the DOE BER site would be useful. This page would provide potential users high-level information on the testbeds and point them to individual testbed web pages and contact information. Other avenues of communicating to the larger scientific community could be town halls at the American Geophysical Union (AGU) or American Meteorological Society annual meetings and smaller testbed sessions at the ASR and Modeling PI meetings. Creating Digital Object Identifier (DOI) numbers for the cases used by the testbeds can be a means of better documenting how observational data is being used and by whom.

When using the testbeds, users need to have a basic knowledge of modeling, processing data, and post-processing. For CAPT, a document is needed that describes the steps needed to run a climate model in a NWP mode. It would be useful to share the procedures for other climate models, such as the GFDL model, as was done in the past. Currently, the CAPT approach has been used in one major international modeling project – the Transpose Atmospheric Modeling Intercomparison Project II (Transpose-AMIP II). Five other climate modeling groups have employed a similar strategy to CAPT for running their climate models in forecast mode. While CAPT simulations produce large amounts of output (multiple terabytes), only the files necessary to create the simulations (which are much smaller) need to be provided by CAPT. The ability to use CAPT for parameterization development will depend on the case study selected, and this issue is common to all of the testbeds.

8.0 Workshop Outcomes and Action Items

In the final session of the workshop, the participants discussed potential action items resulting from the workshop. A range of actions, including both short-term and long-term activities, were identified that could improve coordination and efficiency among the testbed projects and better engage the broader research community. Key recommendations and proposed activities to address these recommendations are listed. Testbed leads also identified the highest priority science questions that their testbeds can address over the next several years.

8.1 Communication

Workshop participants recommended increased communication among the testbed projects, as well as between the testbed projects and the ARM Infrastructure Team, noting that it would reduce potential redundancy in data set development efforts and provide scientific guidance to ARM for prioritizing data set and measurement activities. Specific activities that could be undertaken to increase communication among the testbeds and between the testbeds and ARM include:

- holding breakout sessions on testbed activities at ASR and Modeling PI meetings to discuss collaborative activities among testbeds and engage the broader DOE research community

- using ARM translator teams as the primary means of communication between testbed groups and ARM Infrastructure; discuss testbed activities routinely on ARM translator calls
- identifying high priority data sets and variables for each testbed; provide this information to ARM to assist in their prioritization of data set development and measurement activities.

Several of the testbed groups suggested that the development of new instrument simulators would be useful for enabling better comparisons between observations and models. In particular, groups were interested in simulators for the new ARM scanning and precipitation radars and in simulators for aerosol measurements. The participants noted that the ARM observational community has the key expertise needed for developing simulators and recommended that the testbed groups work with ARM to develop instrument simulators suitable for the model testbeds. Specific activities that could be conducted to address this recommendation include:

- providing specifications for simulator needs from each testbed to ARM/ASR program managers
- holding instrument simulator breakout sessions at ASR meetings to increase communication and collaboration between testbed groups and the ARM/ASR instrument community.

8.2 Sharing of Data Sets/Tools

Workshop participants recommended that testbed groups consider sharing data sets, scripts, and/or visualization tools among testbed groups. While testbeds have generally not examined the same test cases or field campaigns to date, sharing data sets (including model initialization data sets) among the testbed groups could reduce the amount of effort for a given testbed to expand their set of cases. Sharing such data sets through the ARM or Earth System Grid Federation (ESGF) archives, rather than on individual project websites or through email, would increase awareness of testbed data development activities within the larger DOE and science communities. Creating DOI numbers for the testbed data sets would also provide useful documentation on how the data sets are being used by the broader community.

Specific actions to meet this recommendation include:

- identifying mature testbed data sets (including initialization and/or evaluation data sets); submit to ARM PI archive or ESGF
- developing DOI and/or metadata tags to identify testbed data sets in order to track use by broader community
- developing a path forward to climate and forecast-compliant data formats for testbed data sets, so testbeds can more easily use data sets created by other groups
- sharing methodologies for processing model and observational data into comparable variables
- identifying visualization and web tools from existing testbeds that could be applied to other testbeds.

8.3 Engagement with Broader Community

Currently, each testbed group interacts with external research scientists on an ad hoc basis. Participants noted that there would be benefits in more coordinated activities with external scientists, but that providing general support for the broader community to use the testbeds would require additional resources. Participants identified several activities that could be undertaken with existing resources:

- holding AGU Town Halls to describe testbed capabilities to the broader scientific community
- developing a single CESD testbed website that provides detailed information on testbed capabilities, goals, and plans
- surveying ASR and DOE modeling communities on testbed needs and ways the communities would like to use/interact with the testbeds.

With additional resources, potential activities that were identified to engage the broader community include:

- improving documentation of testbeds so that they are more easily used by external users
- adding provenance and metadata to testbed workflows
- improving visualization tools and automation of testbed scripts
- developing coordinated testbed cases with the user community
- extending testbeds to other models and/or computer architectures
- developing a formal methodology and archive for storing testbed simulations and analyses to allow data-mining by the community.

8.4 Coordinate Research Activities Among Testbeds

Workshop participants agreed that coordinating activities among the testbeds around a particular science question, case study, or parameterization could be useful as each testbed evaluates model performance in different, but complementary ways. However, given the potential effort required in developing a group activity (due to the different formats, models, and workflows of the testbeds), such an activity would have to be well defined. Three scientific areas that were of mutual interest to two or more testbeds were identified as areas of potential coordination: aerosol indirect effects, GOAmazon case studies, and land-atmosphere interactions. Activities to further develop coordinated activities among the testbeds include:

- coordinating data set development (including initial/boundary condition files), test cases, and testbed activities for GOAmazon
- defining specific science questions on aerosol indirect effects of land-atmosphere interactions that would benefit from a coordinated testbed approach.

8.5 Continuous High-Resolution LES/CRM Simulations at ARM Sites

One challenge in developing parameterizations for large-scale models is that parameterizations have to account for sub-grid scale processes. Therefore, they require information on relationships between all relevant variables at high temporal and spatial

resolution. Essentially large-scale model developers would like a 4D data cube of all variables of interest. However, no observation system is able to provide all needed variables simultaneously at the needed resolution. High-resolution models (such as LES or CRM) can provide a consistent set of variables at a given spatial and temporal resolution, but their ability to accurately simulate reality may vary with the specific model used, the meteorological regime, or the quality of the initial and boundary conditions. Workshop participants proposed exploring the idea of running continuous high-resolution simulations at ARM sites to provide this 4D data cube. Such simulations could include continuous comparison to observations to identify regimes for which model/observations are in good agreement or could use data assimilation to directly integrate the simulations with ARM observations. It might be necessary to intercompare multiple LES or CRM models to determine which model is most suitable for the meteorological regimes observed at the ARM sites. In addition, participants suggested holding a workshop to explore specifications for continuous high-resolution (LES or CRM) simulations at ARM sites.

8.6 High Priority Science Questions

Finally, testbed leads identified the highest priority science questions that their testbeds might address in the upcoming years.

- How well does new theory on new particle formation, particularly related to organics, represent observed size and number distributions? (AMT)
- How well do current treatments of wet and dry removal perform, and how do those uncertainties affect the lifetime of aerosols in the atmosphere? (AMT)
- What processes control the formation of brown carbon, and what is the best way to represent brown carbon and its effect on optical and cloud nucleating properties in models? (AMT)
- What factors cause CAM5 to simulate a warm bias in the summer surface climate over the Central United States? (CAPT)
- What factors cause the Community Earth System Model 1 to simulate a double Intertropical Convergence Zone in the Atlantic and eastern Pacific Oceans? (CAPT)
- Where and when can model calibration of precipitation to observations be usefully performed through a perturbed-parameter technique? (CSSEF)
- How do different physics options perform for precipitation-related processes in simulations of the regionally refined CAM model over the Central United States and central Indian Ocean? (CSSEF)
- How well are boundary layer processes, convection, and sub-grid variability/structure represented in SCM and/or CRM/LES models, and to what extent can CRM/LES be used to inform/improve the related parameterizations in climate models? (FASTER)
- How well are aerosol-cloud interactions represented in SCM and/or CRM/LES models and what are the influences of considering cloud systems and their evolution on evaluation of aerosol indirect effects? (FASTER)
- How well are the process couplings represented in existing parameterizations, and what are the associated compensating errors? (FASTER)

Appendix A
Agenda and Discussion Questions

CESD Testbed Workshop Agenda

August 5-6, 2013

Location: DOE (Room A-410), Germantown, Maryland

Goal: Increase Communication and Efficiency Across the CESD Testbed Portfolio

Monday, August 5

8:00	Arrive at DOE for Badging
8:30-8:45	Introduction and Charge for Workshop
8:45-10:15	Overview of CESD Testbeds (15-minute talks + 5-minute questions; order TBD) AMT (Jerome Fast), CAPT (Steve Klein), CSSEF (Steve Klein), FASTER (Yangang Liu)
10:15-10:45	Coffee Break
10:45-11:45	What do we want/need from testbeds? (20-minute talks + 10 minute discussion) Vince Larson – What do process modelers need in a testbed? Phil Rasch – What do global modelers need in a testbed?
11:45-12:30	Discussion and Afternoon Charge
12:30-2:00	Lunch
2:00-3:15	Topical Discussion 1 – Testbed Overview (Dorothy Koch)
3:15-3:30	Coffee Break
3:30-4:45	Topical Discussion 2 (Breakouts) 2a – Data Sets (Mike Jensen) 2b – Visualization and Analysis Software (Ashley Williamson/Renu Joseph)
4:45-5:30	Wrap-up of Day 1

Tuesday, August 6

8:30-10:00	Topical Discussion 3 – Model Frameworks (William Gustafson)
10:00-10:30	Coffee Break
10:30-12:00	Topical Discussion 4 – Collaboration/Coordination Opportunities (Jerome Fast/Steve Klein/Yangang Liu)
12:00-1:30	Lunch
1:30-3:00	Outcomes and Action Items Charge for Testbed Overview Talks: • What is the goal of your testbed? • What processes are being studied? • What model(s) are you targeting, and what model framework is being used? • What data sets and time periods are being used and planned? • What visualization and analysis tools are being used?

- What outreach or collaborations do you have and do you plan?
- What are high priority future directions?

Topical Discussion 1: Testbed Overview

- What is a testbed?
- What are science goals of the testbeds?
- What are the distinctions of the testbeds?
- What are the similarities of the testbeds?
- What are current science questions that testbeds do not, but could, address?

Topical Discussion 2a: Data sets (ARM and other atmospheric, including land-atmosphere and forcing)

- What are existing data sets, how are these formatted?
- What are planned data sets?
- What data sets are needed?
- How might data set distribution be better coordinated across the DOE community?
- How might we improve coordination of ARM data set development; coordination with ARM Data Archive?

Topical Discussion 2b: Visualization and Analysis and Methods

- What are the existing analysis/visualization capabilities?
- What are plans for further development of these?
- Might the current tools/technologies be shared among groups?
- What analysis is done, UQ, bias calculations, tuning, etc.

Topical Discussion 3: Model Frameworks

- What are current models using DOE testbeds?
- How are models forced, initialized, configured?
- What are needs/gaps in current testbeds that models require?
- How might modeling efforts be shared across groups?
- What are plans for including land-atmosphere interactions in your testbed?

Topical Discussion 4: Collaboration/Coordination Opportunities

- What are the most pressing science issues?
- What gaps exist in current testbeds?
- What are best opportunities to exploit synergies in data, software, and models?
- What is vision for service to the broader community?

Appendix B
CESD Testbed Workshop Attendees

Attendee	Institution	Testbed
Satoshi Endo	BNL	FASTER
Jerome Fast (AMT lead)	PNNL	AMT
William Gustafson	PNNL	AMT
Mike Jensen	BNL	FASTER
Steve Klein (CAPT, CSSEF lead)	LLNL	CSSEF and CAPT
Vince Larson	University of Wisconsin, Milwaukee	
Wuyin Lin	BNL	FASTER and CSSEF
Yangang Liu (FASTER lead)	BNL	FASTER
Hsi-Yen Ma	LLNL	CAPT
Jim Mather	PNNL	ARM
Brian Medeiros	NCAR	CAPT
Phil Rasch	PNNL	
Laura Riihimaki	PNNL	CSSEF
Chitra Sivaraman	PNNL	ARM
Jimmy Voyles	PNNL	ARM
Shaocheng Xie	LLNL	CAPT
Yunyan Zhang	LLNL	CAPT

Consultants	Institution
Vince Larson	University of Wisconsin, Milwaukee
Phil Rasch	PNNL

Appendix C
Additional Testbed Details and Workflow

AMT

Since the host model of the AMT is WRF-Chem, the user needs to be familiar with running WRF and WRF-Chem. The code is publically available from NCAR at **http://www.wrf-model.org**. The user can then get the AMT software and testbed field campaign data cases from the ARM Data Archive at **http://www.arm.gov/data/eval/59**, or directly from the AMT principal investigator. User's guides are provided with the AMT software. Additional guidance is also available from the AMT PI.

The following are the steps for users to run a model experiment utilizing the AMT that assesses the performance of a new or modified parameterization:

(1) Decide which testbed case is most relevant to the particular science question that will be addressed by the new or modified parameterization to be tested.

(2) Set up and run a WRF-Chem "control" simulation that contains an existing parameterization, which the new or modified parameterization will be compared to. Setting up the simulation will involve configuring the 3D computational domain, generating the initial and boundary conditions for meteorology, trace gas chemistry, and aerosols, and generating input emission files. If the user desires, the AMT PI can provide these files to the user for preconfigured cases. Depending on the testbed case, the initial and boundary conditions are provided by a variety of models including global reanalyses, global chemical transport model simulations, or CAM5 simulations. The emission files will also depend on the testbed case and are obtained from the most appropriate source for that case. The "control" simulation can also be provided to the user.

(3) Implement the new or modified parameterization in WRF-Chem, and perform an additional "sensitivity" simulation using it.

(4) Modify the script before executing to indicate the location (directory path) where the AMT software and WRF-Chem simulations are located. The AMT includes a script that automatically runs various extraction and statistics programs using the AMT "analysis toolkit" software. The script is run first to extract simulated meteorological, chemical, and aerosol quantities in a manner compatible with the available measurements. The user has the option of extracting model variables of one type for surface, aircraft, profile, and satellite instrument platforms, or all of the possible combinations at one time that are run sequentially by the script. Once the model variables have been extracted, the script is modified and run again to activate the statistics programs that produce quick-look plots and tables of statistics that compare observed and simulated quantities.

(5) Inspect the results to determine the differences in the performance from the "control" and "sensitivity" simulations. The user will likely perform additional analyses in addition to those generated by the AMT. This is also an appropriate point to compare the computational time required by the simulations to assess the computational cost of the new or modified parameterization.

(6) Decide if the parameterization needs further development. If so, steps (3)–(5) can be repeated.

(7) Test the new or modified parameterization with an additional testbed case, if desired, to obtain performance for a different set of conditions. The user may wish to repeat the "control" simulation with another existing parameterization for additional comparison purposes.

(8) Document findings from the new or modified parameterization in terms of a journal publication.

(9) Make the new or modified parameterization available to the AMT user community, or to the larger WRF modeling community, if it is clearly superior to an existing scheme.

If the user is not proficient in WRF-Chem, the AMT principal investigator can test and evaluate the new or modified parameterization for the user. However, that will depend if sufficient resources are available to the AMT principal investigator to conduct the study.

Additional information on the AMT can be found at **http://www.pnl.gov/atmospheric/research/aci/amt**. The website contains PDF documents that can be downloaded, describing how the testbed case data are generated and how to run the analysis toolkit software. The AMT was originally developed under PNNL's Laboratory Directed Research and Development program through the Aerosol Climate Initiative between FY2007 and FY2009. Since FY2010, research associated with the AMT has been supported by the ASR and ARM programs.

CAPT

Workflow

(1) Download and process operational analysis/reanalysis data needed for generating initial and boundary conditions.

(2) Perform nudging simulations with atmospheric state variables from operational analysis/reanalysis to generate atmosphere and land initial conditions.

(3) Perform hindcast experiments.

(4) Process and analyze results from the hindcast experiments.

Input Data Sets: State variables (U, V, T, Q, and PS, optionally), and SST and sea ice are from operational analysis/reanalysis (ERA-Interim, YOTC analysis, or MERRA) for atmospheric initial and boundary conditions, respectively.

Initialization Methodology. A 6-hourly nudging simulation with atmospheric state variables from operational analysis/reanalysis is first performed for the spin-up of certain atmospheric and land variables, such as clouds and aerosols for the atmospheric model, and soil temperature and moisture for the land model. These variables, along with the state variables from operational analysis/reanalysis, are later used as the initial conditions for the atmospheric and land models in the hindcast experiments. The spin-up run usually starts 3-6 months earlier than the period of interest in which the hindcast experiments are performed.

Hindcasts: A series of short-term hindcasts (e.g., 6-day long) are carried out with the initial and boundary conditions from the above methodology.

Evaluation/Visualization Performed: The results from hindcast experiments are processed and evaluated with certain metrics and diagnostics for the model performance.

Types of Scripts Used: For processing initial and boundary conditions and evaluating the hindcasts' output, sets of scripts and programs based on UV-CDAT/Python, Fortran,

Earth System Modeling Framework, and Unix shell are utilized. The programs for plotting/visualization are from UV-CDAT/Python, NCL, the Integrated Design Lab, and the Grid Analysis and Display System.

CSSEF

At present, the CSSEF testbed workflow for a single-integration of the regionally refined model is very similar to that of the CAPT testbed, with the only minor difference being the use of nudging instead of direct initialization. Thus, an extensive discussion of CSSEF workflow will not be given here. It is worth noting that the project aspires to greatly improve the efficiency of the workflow by automation, provenance, documentation, etc. However, it remains to be seen whether this will be accomplished in the context of the new DOE climate modeling project.

A unique aspect of the CSSEF tested is the use of large ensembles of perturbed parameter integrations. This modifies the workflow in that one must modify the simulation environment by having numerous model instances, generating the perturbed parameter values for each model instance, and managing the model simulation jobs and output.

FASTER

The general evaluation workflow consists of the following steps:

(1) Choose the ARM observational site (e.g., SGP) and period (e.g., RACORO) to be evaluated

(2) Identify (or download if necessary) the observational data used for initial and boundary conditions and evaluation purpose, including the set of large-scale forcing necessary for driving the SCM and CRM/LES

(3) Perform simulations (no need for this step for the direct evaluation of NWP results)

(4) Process and generate quick-look plots

(5) Perform detailed analysis to identify/quantify model-observation differences and relate the differences to physical sources underlying specific parameterizations, including idealized sensitivity investigation, as needed

(6) Fix/update the corresponding parameterizations or develop new ones.

Note that as a side effort, in collaboration of computer experts, we have briefly explored a more formal workflow framework based on individual physical modules, which allows a more sophisticated management of workflow and running models with several computers in different locations (Wu et al. 2012). The work could be continued and expanded if there is such a programmatic need/desire.

The online FASTER testbed adopts server-client style programming, with the server end heavy on using Perl script to control SCM simulations, while the client end is heavy on using Javascript language and Ajax technology to gather user inputs and communicate with the server. Perl-based thread-parallel control in a memory computing platform is used to handle multitasking, such as ensemble simulations or concurrent reinitialized short-term simulations over an extended period of time.

Quick-look plots are generated on the server end using the NCL. The Javascript-enabled client-end browser can interactively display the plots by category. D3js, a Javascript library for manipulating documents based on data, is being adopted to further enhance the user experience on interactive visualization of the results, by generating plots on the client end using processed data fed by the server.

Appendix D
References

Anderson, J, T Hoar, K Raeder, H Liu, N Collins, R Torn, and A Avellano. 2009. "The data assimilation research testbed: A community facility." *Bulletin of the American Meteorological Society* 90: 1283-1296.

Blossey, PN, CS Bretherton, M Zhang, A Cheng, S Endo, T Heus, Y Liu, AP Lock, SR de Roode, and K-M Xu. 2013. "Marine low cloud sensitivity to an idealized climate change: The CGILS LES intercomparison." *Journal of Advanced Modeling Earth Systems* 5(2): 1-25, doi: 10.1002/jame.20025.

Boyle, J, and SA Klein. 2010. "Impact of horizontal resolution on climate model forecasts of tropical precipitation and diabatic heating for the TWP-ICE period." *Journal of Geophysical Research* 115: D23113, doi:10.1029/2010JD014262.

Ensberg, JJ, JS Craven, AR Metcalf, JD Allan, WM Angevine, R Bahreini, J Brioude, C Cai, H Coe, JA de Gouw, RA Ellis, JH Flynn, CL Haman, PL Hayes, JL Jimenez, BL Lefer, AM Middlebrook, JG Murphy, JA Neuman, JB Nowak, JM Roberts, J Stutz, JW Taylor, PR Veres, JM Walker, and JH Seinfeld. 2013. "Inorganic and black carbon aerosols in the Los Angeles Basin during CalNex." *Journal of Geophysical Research* 118: 1777-1803, doi: 10.1029/2012JD018136.

Fast, JD, WI Gustafson Jr., EG Chapman, RC Easter, J Rishel, RA Zaveri, G Grell, and M Barth. 2011. "The Aerosol Modeling Testbed: A community tool to objectively evaluate aerosol process modules." *Bulletin of the American Meteorological Society* 92: 343-360.

Fast, JD et al. 2014. "Sensitivity of simulated regional aerosol distributions over California to emissions and long-range transport during the CalNex and CARES campaigns." *In preparation*.

Gustafson, WI Jr., Y Qian, and JD Fast. 2011. "Downscaling aerosols and the impact of neglected sub-grid processes on aerosol radiative forcing for a representative GCM grid spacing." *Journal of Geophysical Research* 116: D13303, doi: 10.1029/2010JD015480.

Hannay, C, DL Williamson, JJ Hack, JT Kiehl, SA Klein, CS Bretherton, and M Kohler. 2009. "Evaluation of forecasted southeast Pacific stratocumulus in the NCAR, GFDL, and ECMWF models." *Journal of Climate* 22: 2871–2889.

Illingworth, AJ, RJ Hogan, EJ O'Connor, D Bouniol, ME Brooks, J Delanoe, DP Donovan, JD Eastment, N Gaussiat, JWF Goddard, M Haeffelin, H Klein Baltink, OA Krasnov, J Pelon, J-M Piriou, A Protat, HWJ Russchenberg, A Seifert, AM Tompkins, G-J van Zadelhoff, F Vinit, U Willen, DR Wilson, and CL Wrench. 2007. "CLOUDNET - Continuous Evaluation of Cloud Profiles in Seven Operational Models Using Ground-Based Observations." *Bulletin of the American Meteorological Society* 88(6), doi: 10.1175/BAMS-88-6-883.

Klein, SA, X Jiang, J Boyle, S Malyshev, and S Xie. 2006. "Diagnosis of the summertime warm and dry bias over the U. S. Southern Great Plains in the GFDL climate model using a weather forecasting approach." *Geophysical Research Letters* 33: L18805, doi: 10.1029/2006GL027567.

Knote, C, A Hodzic, JL Jimenez, R Volkamer, S Baidar, J Brioude, JD Fast, JB Gilman, A Goldstein, J de Gouw, P Hayes, BT Jobson, WB Knighton, WC Kuster, H Oetjen, J Orlando, C Song, H Stark, PS Stevens, R Thalman, G Tyndall, C Warneke, R Washenfelder, E Waxman, and Q Zhang. 2013. "Simulation of semi-explicit mechanisms of SOA formation from glyoxal in a 3-D model." *Atmospheric Chemistry and Physics Discussions* 13: 26699-26759.

Li, Z, Z Zang, Q Li, Y Chao, D Chen, Z Ye, Y Liu, and K-N Liou. 2013. "Three-dimensional variational data assimilation system for multiple aerosol species with WRF/Chem and an application to PM2.5 prediction." *Atmospheric Chemistry and Physics* 13: 4265-4278, doi:10.5194/acp-13-4265-2013.

Liu, G, Y Liu, and S Endo. 2013. "Evaluation of surface flux parameterizations with long-term ARM Observations." *Monthly Weather Reviews* 141: 773-797.

Lu, C, Y Liu, S Niu, S Krueger, and T Wagner. 2013. "Exploring parameterization for turbulent entrainment-mixing processes in clouds." *Journal Geophysical Research* 118: 185-194, doi:10.1029/2012JD018464.

Ma, PL, PJ Rasch, JD Fast, RC Easter, WI Gustafson Jr., X Liu, SJ Ghan, and B Singh. 2013. "Assessing the CAM5 Physics Suite in the WRF-Chem Model: Implementation, Evaluation, and Resolution Sensitivity." *Geoscientific Model Development Discussions* 6: 6167-6218.

Phillips, TJ, GL Potter, DL Williamson, RT Cederwall, JS Boyle, M Fiorino, JJ Hnilo, JG Olson, S Xie, and JJ Yio. 2004. "Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction." *Bulletin of the American Meteorological Society* 85: 1903-1915.

Qian, Y, WI Gustafson Jr., and JD Fast. 2010. "An investigation of the sub-grid variability of trace gases and aerosols for global climate models." *Atmospheric Chemistry and Physics* 10: 6917-6946.

Shrivastava, M, JD Fast, RC Easter, WI Gustafson Jr., RA Zaveri, A Hodzic, and J Jimenez. 2011. "Modeling organic aerosols in a megacity: Comparison of simple and complex representations of the volatility basis set approach." *Atmospheric Chemistry and Physics* 11: 6639-6662.

Shrivastava, M, A Zelenyuk, D Imre, R Easter, J Beranek, RA Zaveri, and JD Fast. 2013. "Implications of low volatility SOA and gas-phase fragmentation reactions on SOA loadings and their spatial and temporal evolution in the atmosphere." *Journal of Geophysical Research* 118(8): 3323-3342, doi:10.1002/jgrd.50160.

Song, H, W Lin, Y Lin, A Wolf, R Neggers, L Donner, A Del Genio, and Y Liu. 2013. "Evaluation of precipitation simulated by seven SCMs against the ARM observation at the SGP site." *Journal of Climate* 26: 5467-5489.

Thomas, JL, J-C Raut, KS Law, L Marelle, G Ancellet, F Ravetta, JD Fast, G Pfister, LK Emmons, GS Diskin, A Weinheimer, A Roiger, and H Schlager. 2013. "Pollution transport from North America to Greenland during summer 2008." *Atmospheric Chemistry and Physics* 13: 3825-3848.

Williamson, DL, J Boyle, R Dederwall, M Fiorino, J Hnilo, J Olson, T Phillips, G Potter, and SC Xie. 2005. "Moisture and temperature balances at the Atmospheric Radiation Measurement Southern Great Plains Site in forecasts with the Community Atmosphere Model (CAM2)." *Journal of Geophysical Research* 110: D15S16, doi: 10.1029/2004JD005109.

Wu, Q, M Zhu, Y Gu, P Brown, X Lu, W Lin, and Y Liu. 2012. "A distributed workflow management system with case study of real-life scientific applications on grids." *Journal of Grid Computing* 10(3): 367-393, doi: 10.1007/s10723-012-9222-7.

Xie, S, H-Y Ma, JS Boyle, SA Klein, and Y Zhang. 2012. "On the correspondence between short- and long- timescale systematic errors in CAM4/CAM5 for the Years Of Tropical Convection." *Journal of Climate* 25:7937–7955, doi: 10.1175/JCLI-D-12-00134.1.

Yang, Q, WI Gustafson Jr., JD Fast, H Wang, RC Easter, H Morrison, Y-N Lee, EG Chapman, SN Spak, and MA Mena-Carrasco. 2011. "Assessing regional scale predictions of aerosols, marine stratocumulus, and their interactions during VOCALS-REx using WRF-Chem." *Atmospheric Chemistry and Physics* 11:11951-11975.

Appendix E
Acronyms

AAF	ARM Aerial Facility
AGU	American Geophysical Union
AMIE	ARM MJO Investigation Experiment
AMF	ARM Mobile Facility
AMT	Aerosol Modeling Testbed
ANL	Argonne National Laboratory
ARM	Atmospheric Radiation Measurement
ARMBE	ARM Best Estimate
ASR	Atmospheric System Research
BNL	Brookhaven National Laboratory
CAM	Community Atmospheric Model
CAPT	Cloud-Associated Parameterizations Testbed
CARES	Carbonaceous Aerosol and Radiative Effects Study
CAUSES	Clouds Above the United States and Errors at the Surface
CESD	Climate and Environmental Sciences Division
CHAPS	Cumulus Humilis Aerosol Processing Study
CLOWD	Clouds with Low Optical Water Depths
CMIP	Coupled Model Intercomparison Project
COSP	Cloud Feedback Intercomparison Project
CRM	Cloud Resolving Model
CSSEF	Climate Science for a Sustainable Energy Future
DART	Data Assimilation Research Testbed
DOE	Department of Energy
DOI	Digital Object Identifier
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA	ECMWF Reanalysis
ESGF	Earth System Grid Federation
ESM	Earth System Modeling
FASTER	Fast-Physics System Testbed and Research
GCM	global climate models
GFDL	Geophysical Fluid Dynamics Laboratory
GFS	Global Forecast System
GISS	Goddard Institute for Space Studies
GOAmazon	Green Ocean Amazon
IOP	Intensive Operational Period

ISDAC	Indirect and Semi-Direct Aerosol Campaign
LANL	Los Alamos National Laboratory
LBNL	Lawrence Berkeley National Laboratory
LES	large eddy simulation
LLNL	Lawrence Livermore National Laboratory
MAGIC	Marine ARM Global Energy and Water Cycle Experiment Cloud Systems Study Pacific Cross-section Intercomparison
MC3E	Midlatitude Continental Convective Clouds Experiment
MERRA	Modern-Era Retrospective Analysis for Research and Applications
MILAGRO	Megacity Initiative: Local and Global Research Observations
MJO	Madden-Julian Oscillation
M-PACE	Mixed-Phase Arctic Cloud Experiment
MOZART	Model for Ozone and Related Chemical Tracers
NCAR	National Center for Atmospheric Research
NEXRAD	Next Generation Radar
NCL	NCAR command language
NWP	Numerical Weather Prediction
ORNL	Oak Ridge National Laboratory
PI	principal investigator
PNNL	Pacific Northwest National Laboratory
RACORO	Routine AAF Clouds with CLOWD Optical Radiative Observations
RGCM	Regional and Global Climate Modeling
SCM	single column model
SGP	Southern Great Plains
SOA	secondary organic aerosol
SNL	Sandia National Laboratory
SST	sea surface temperature
TCAP	Two-Column Aerosol Project
TWP-ICE	Tropical Western Pacific International Cloud Experiment
UQ	uncertainty quantification
UV-CDAT	Ultrascale Visualization Climate Data Analysis Tools
VOCALS	Variability of the American Monsoon System Ocean-Cloud-Atmos-Land Study
WRF	Weather Research and Forecasting
YOTC	Year of Tropical Convection

For More Information

Climate and Environmental Sciences Division • http://science.energy.gov/ber/research/cesd/
Gary Geernaert, gerald.geernaert@science.doe.gov

ARM Climate Research Facility • http://www.arm.gov
Wanda Ferrell, wanda.ferrell@science.doe.gov
Rick Petty, rick.petty@science.doe.gov

Atmospheric System Research • http://asr.science.energy.gov/
Ashley Williamson, ashley.williamson@science.doe.gov
Sally McFarlane, sally.mcfarlane@science.doe.gov

Earth System Modeling • http://climatemodeling.science.energy.gov/
Dorothy Koch, dorothy.koch@science.doe.gov

Regional and Global Climate Modeling • http://climatemodeling.science.energy.gov/
Renu Joseph, renu.joseph@science.doe.gov