Quality Assurance Project Plan

Project 14-025
Development and Evaluation of an Interactive Sub-Grid Cloud Framework for the CAMx Photochemical Model

Chris Emery
ENVIRON International Corporation

Dr. John Nielsen-Gammon
Dr. Kenneth Bowman
Dr. Renyi Zhang
Texas A&M University

Summary of Project

QAPP Category Number: III
Type of Project: Research or Development (Modeling)

QAPP Requirements: This QAPP includes descriptions of the project and objectives; organization and responsibilities; model development approach; model evaluation procedures and quality metrics; data analysis, interpretation, and management; reporting; and references.

QAPP Requirements:
Audits of Data Quality: 10% Required
Report of QA Findings: Required in final report

May 29, 2014

DISTRIBUTION LIST

Gary McGaughey, Project Manager, Texas Air Quality Research Program

Cyril Durrenberger, Quality Assurance Project Plan Officer, Texas Air Quality Research Program

Khalid Al-Wali, Project Liaison, Texas Commission on Environmental Quality

Chris Owen, Quality Assurance Project Plan Officer, Texas Commission on Environmental Quality

Maria Stanzione, Program Manager, Texas Air Quality Research Program

1. PROJECT DESCRIPTION AND OBJECTIVES

1.1 Problem Statement

Small-scale clouds are often widespread but they are not explicitly resolved by the grid scales employed in regional meteorological and photochemical modeling applications. The physical effects from these sub-grid clouds are difficult to characterize accurately, but they can substantially influence many different atmospheric processes, including: boundary layer mixing, ventilation, and deep vertical transport of heat, moisture, and chemical tracers; radiative transfer and surface heat budgets; spatio-temporal precipitation patterns, intensity and wet scavenging rates; chemistry via photolysis and aqueous reactions; and certain environmentally-sensitive emission sectors (e.g., biogenics).

Sub-grid cloud parameterizations are employed in meteorological models to adjust grid-resolved vertical profiles of heat and moisture from the effects of moist convection. Alapaty et al. (2012) have made improvements to the Weather Research and Forecasting model (WRF; Skamarock et al., 2008) by using cloud parameters derived from the Kain-Fritsch (K-F) parameterization (Kain, 2004) to internally adjust the grid-scale cloud fields that feed into WRF's radiation treatments. They show that simulated surface temperature and precipitation fields are improved relative to the unmodified version of WRF. The group is also developing new multi-scale treatments for K-F to extend its applicability over a wider range of grid resolutions. These updates are scheduled to be available in the next public distribution of WRF in Spring 2014.

From the air quality perspective, moist convection is an important component for long-range transport of ozone, particulate matter (PM), and precursors. However, the spatial/temporal distributions of such clouds must be re-diagnosed because meteorological models do not export necessary information from their sub-grid cloud parameterizations. This leads to potentially large inconsistencies between the models. The Texas Commission on Environmental Quality (TCEQ) uses the Comprehensive Air quality Model with extensions (CAMx; ENVIRON, 2014) for research and regulatory photochemical modeling. CAMx implicitly addresses the influence of sub-grid clouds by diagnosing their presence according to resolved wind and thermodynamic fields from meteorological models, and blending their properties into the resolved cloud fields (Emery et al., 2010). The final blended cloud fields are used to adjust photolysis rates, perform aqueous chemistry, and remove pollutants via wet scavenging at grid scale – no separate sub-grid cloud processes are explicitly treated. Furthermore, CAMx does not include a cloud convective mixing treatment. Two recent studies conducted by Kemball-Cook et al. (2012, 2013) concluded that improvements to the CAMx modeling system, including a sub-grid convection treatment, are necessary to reduce the low bias in upper tropospheric nitrogen dioxide (NO_2) and peroxyacetyl nitrate (PAN) and to more accurately represent the total oxidized nitrogen (NOy) budget aloft.

1.2 Project Objectives

The project team will develop a sub-grid convective cloud mixing scheme for CAMx, incorporating chemistry, wet deposition, and vertical mixing processes into an interactive cloud modeling system. Our approach will be complementary with the WRF updates developed by Alapaty et al. (2012), whereby additional sub-grid cloud output fields will be passed to CAMx to define spatial/temporal distributions and mixing rates for the sub-grid cloud algorithm in CAMx. Through a cooperative arrangement, Dr. Kiran Alapaty's group at the Environmental Protection Agency (EPA) has agreed to interact with our team to provide interim and final versions of WRF and to make specific variables available from WRF to support sub-grid cloud modeling in CAMx.

The new CAMx treatment will be tested for a convective period that occurred during the September 2013 Houston DISCOVER-AQ field study (Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality). Testing will particularly address convective cumulus impacts to tropospheric profiles of NOx (nitric oxide plus nitrogen dioxide), ozone, and other chemical tracers by comparing to in situ concentration and tropospheric column profiles from aircraft measurements. Through a cooperative arrangement, Dr. Ken Pickering's group at the National Aeronautics and Space Administration (NASA) Langley Research Center has agreed to provide data products from the DISCOVER-AQ campaign. Convection occurred during the first week of the field campaign, particularly on September 4, 6, and 11. These data products should be available by Spring 2014. CAMx will also be run and evaluated against aircraft data from the 2008 Stratosphere-Troposphere Analyses of Regional Transport 2008 (START08) monitoring campaign, for which field data are already available at the Texas A&M University (TAMU). The new model will be provided to TCEQ to support future regulatory and research-oriented ozone and PM modeling.

2. ORGANIZATION AND RESPONSIBILITIES

2.1 Personnel and Responsibilities

This project is a collaborative effort between ENVIRON International Corporation (ENVIRON) and Texas A&M University (TAMU). Mr. Chris Emery of ENVIRON and Dr. John Nielsen-Gammon of TAMU will serve as co-Principal Investigators on this project with overall responsibility for the research and associated quality assurance. Other co-Principal Investigators include Dr. Bowman and Dr. Zhang of TAMU. The project will be overseen by AQRP Project Manager Mr. Gary McGaughey and TCEQ Project Liaison Dr. Khalid Al-Wali. The scientists working on this project and their specific responsibilities are listed in Table 1 below.

Table 1. Project participants and their affiliations and key responsibilities.

Participant (Organization)	Key Responsibilities
Mr. Christopher Emery (ENVIRON)	Co-Principal Investigator: project oversight; lead model design, development and testing; lead CAMx applications; lead reporting
Dr. Sue Kemball-Cook (ENVIRON)	Assistance with convective model design
Mr. Jeremiah Johnson (ENVIRON)	Lead code development and testing
Dr. John Nielsen-Gammon (TAMU)	Co-Principal Investigator: TAMU oversight; lead WRF meteorological modeling; reporting
Dr. Kenneth Bowman (TAMU)	Co-Principal Investigator: Model comparisons to field study data
Dr. Renyi Zhang (TAMU)	Co-Principal Investigator: CAMx application to field study periods

2.2 Schedule

The project is divided into five major tasks. The schedule for specific tasks is listed in Table 2, and includes interim progress reports and project deliverables.

Table 2. Schedule of project activities.

Task	2014									2015					
	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Task 1: Preparation & Software Design	→	→	→												
WP/QAPP	●														
Executive Summary		●													
Task 2: Convective Model System				→	→	→									
Task 3: Chemistry and Wet Deposition							→	→	→						
Task 4: Model Evaluation				→	→	→	→	→	→	→	→	→			
Task 5: Reporting										→	→	→	→	→	→
Draft Report														●	
Final Report															●
Project Presentation															●
Progress Reports		■	▲	▲	■	▲	▲	■	▲	▲	■	▲	▲	■	

Project Report Material	●
Monthly Progress Report	▲
Quarterly Progress Report	■

3. MODEL DEVELOPMENT APPROACH

3.1 Functional Requirements

3.1.1 Required Functions

The functional requirements of the CAMx sub-grid cloud model will include:

- Direct linkage to the latest version of WRF containing the K-F improvements and new output fields incorporated by EPA (Alapaty et al., 2012) as of January 1, 2015;
- Vertical convective treatment for pollutant mixing and transport within the sub-grid cloud column, and entrainment/detrainment with the ambient column;
- Aqueous chemistry and wet scavenging of pollutants within the sub-grid cloud column.
- Integration of cloud and ambient column pollutant mass at each hour for writing to CAMx time-averaged output files.

3.1.2 Functionality, Interfacing, Performance, and Constraints

The project team will discuss with Dr. Alapaty's group at EPA to define the specific WRF K-F cumulus variables that will be output to support the CAMx cloud model framework. From this we will further refine details of the methodology to incorporate a sub-grid cloud model in CAMx. The WRF-to-CAMx interface (WRFCAMx) will be modified to pass the additional WRF sub-cloud data to CAMx via a new cloud/rain input file format. New user-defined variables will be needed in the CAMx namelist control file to select and/or configure the sub-grid cloud model. No modifications to other input or output files are anticipated. For this initial version, the new sub-cloud model will not interact with Plume-in-Grid or any probing tools, although code development will need to consider linking to these components in the future.

The new sub-grid module must be coded to ensure numerical stability and to balance accuracy and efficiency so that impacts to overall model speed performance is minimized as much as feasible. Ideally, the new treatment should be implemented to minimize changes to the CAMx program flow and subroutine calling infrastructure.

ENVIRON will employ the latest model FORTRAN codes available for CAMx (v6.10) and its related pre- and post-processors, and implement modifications in such a manner as to minimize impacts to other components of the modeling system. To remain consistent with CAMx code, all modified and new modules will be written in the FORTRAN90 (F90) standard, with extensions compatible with today's most widely used FORTRAN compilers, specifically Portland Group and Intel F90. The code will adhere to the CAMx coding/format style, including the use of appropriate in-code documentation (comment statements), loop indentation, and memory management techniques. All variables must be type-declared consistent with the FORTRAN "implicit none" statement for all new routines.

3.1.3 Hardware and Operating System Constraints

CAMx and associated pre- and post-processors are expected to be run on workstations and cluster environments running common distributions of the Linux operating system; they will not be expected to run on the Windows operating system. Model code will include parallelization using Message Passing Interface (MPI) and Open Multi-Processing (OMP) protocols, and thus must be compatible with common Linux MPI library installations and FORTRAN compilers supporting OMP in-code directives.

3.2 System Design

3.2.1 System Overview

The program to be updated is the CAMx photochemical grid model. We will develop a comprehensive interactive sub-grid cloud framework in CAMx that addresses shallow mixing, deep convective transport, gas and aqueous chemistry, and wet scavenging. All processes will be driven by specific data obtained from output fields generated by the WRF K-F scheme.

The K-F cumulus parameterization scheme is fundamentally a mass flux scheme (Kain 2004). Changes to grid-scale temperature and moisture are calculated from the parameterized properties of entraining/detraining plumes that constitute convective updrafts and convective downdrafts, and from ambient mass fluxes (i.e., compensating subsidence outside the cloud) necessary to maintain mass conservation. This places the K-F scheme within a subset of cumulus parameterization schemes for which constituent transport is already implicit. Outside the convective volume, the ambient mass fluxes will be output from WRF as an effective compensating vertical velocity that would be added to grid-resolved vertical wind for the core CAMx vertical advection scheme. Within the convective volume, updraft and downdraft plumes produce non-local (multi-layer) transport, so WRF will be modified to calculate, store, and output the fractional sources of air from each model level that constitute the resulting mixture of post-convection air at each model level. A developer-specified parameter will control the fractional mixing within the updraft plume. Decisions will need to be made on the relative mixing within the updraft and downdraft plumes, and on interactions between smaller-scale, short-lived convection and larger-scale, steady-state convection.

The CAMx sub-grid cloud model framework will operate separately from the normal grid processes in a manner similar to the Plume-in-Grid (PiG) model (Emery et al., 2013a; ENVIRON, 2014). This “cloud-in-grid” (CiG) approach will define at each hour the physical attributes of a multi-layer cloud “reactor” (much like a PiG “puff”) according to the hourly cloud data provided by WRF. Each CiG reactor configuration would be unique to each grid column (or entirely absent from it) and characterize a steady-state sub-grid cloud environment between each meteorological update time. Fractions of pollutant vertical mass profiles from each host grid column will be allocated to each CiG reactor layer, which would then operate on that mass to include vertical transport, entrainment/ detrainment with the ambient grid column, chemistry, and wet removal.

3.2.2 Component Description

The CAMx photochemical model is fully described in the User’s Guide (ENVIRON, 2014).

The WRFCAMx meteorological interface program will accommodate new output fields from the latest WRF version from EPA. The CAMx cloud/rain file format will be updated to optionally include the new K-F variable fields according to whether all needed WRF K-F variables are available and whether the user wishes to invoke the sub-grid cloud model in CAMx. Basic process testing and debugging will be performed by running short (~1 day) WRF test cases through the updated WRFCAMx code.

The new CiG model framework will include necessary input/output (I/O) infrastructure and the convective mixing/transport and entrainment/ detrainment components. The framework will call existing gas and aqueous chemistry for each layer within the CiG column, and the existing wet scavenging treatment spanning the entire sub-grid cloud column. Special consideration will be given to photolysis rate adjustments and column wet scavenging coefficients within the CiG reactors. With the amount and composition of air participating in updrafts and downdrafts known, such processes can be simulated as long as the trajectory of the air in time and thermodynamic space is known or parameterized.

A new routine will be developed to blend cloud-column and ambient column concentrations during time integration according to the fractional convective cloud cover in each grid column. Resulting column-averaged concentrations will contribute to the time-averaged concentration array for output to either 2-D or 3-D concentration files. No changes in output file formats are anticipated.

3.2.3 Rationale for Selected Hardware/Software Tools

The software and hardware selected for this project are consistent with the current CAMx and related utility programming code, as well as the compilers and platforms used to develop, build and run these systems, respectively. This will ensure modeling system compatibility with TCEQ's current computer system.

3.3 Implementation

3.3.1 Software System Development

Code development will be directed by a single ENVIRON staff member, with assistance from one or two other staff members as needed to develop and test specific process modules. The lead developer will oversee construction of all facets of new code, ensure seamless integration among new subroutines and within the CAMx program flow, and lead all testing and quality assurance steps. The lead developer will maintain close communication with ENVIRON's co-Principal Investigator to report progress, technical issues, and possible solutions.

3.3.2 Verification and Validation

Basic process testing and debugging will be performed by first running the CiG framework inside a standalone test-bed driver program, in which simple user-controlled cloud environments are provided to the CiG to ensure proper operation and data passing between the cloud column and ambient column. Functionality, interfacing, performance and design constraints for the new module will be evaluated. Good FORTRAN coding practices and FORTRAN compile-time checks will help to confirm that the PiG subroutines are coded properly.

Upon successful testing, the CiG framework will be implemented as a called sub-model into the CAMx model, and tied into the model's MPI and OMP parallelization. CAMx will be tested by running short (~1 day) test cases using input data streams from the updated WRFCAMx system. This testing will focus on identifying implementation bugs and performance issues. Potential alternatives will be considered and tested to improve speed, including modifications to the parallelization approach.

Final CAMx system evaluation will be conducted by applying the updated model for two existing modeling datasets covering the September 2013 DISCOVER-AQ and May-June 2008 START08 field campaigns. This testing and associated datasets are described in Section 4.

3.3.3 Release and Delivery Management

The testing described in Section 3.3.2 will encompass "beta" testing of the new CAMx model and associated utilities. Once the system is confirmed to be working correctly and results from the test bed cases are fully evaluated, ENVIRON will transfer the revised program codes to TAMU for installation on their computer system. TAMU will then commence the final model evaluation against data from the DISCOVER-AQ and START08 field study campaigns. At the end of the project, the final CAMx model and all associated support utilities (e.g., WRFCAMx, etc.) will be delivered to the AQRP (see Section 6).

3.3.4 Version Control and Documentation

All model development and modifications to be undertaken in this project will be coordinated with ENVIRON's CAMx model development team so that future public model releases can include this work. All model modifications will be considered to be open source.

ENVIRON will apply its standardized in-house protocols for CAMx version control, documentation, and archival. All developmental versions (i.e., "alpha" and "beta") are managed separately. Once a version is determined to be ready for public distribution, additional code checks are applied to ensure that standard FORTRAN techniques are used throughout all model routines, the code can be built successfully with several common compilers, and that all functions are properly tied into the OMP and MPI parallelization schemes. All routines and programs are annotated with version and release date. Then the model is subjected to a series of standard tests using an in-house database. The User's Guide is updated to include descriptions of any new or modified processes, and any changes or new requirements for I/O.

The CAMx source code, User's Guide, and test dataset are then made available at www.camx.com. Public CAMx releases are provided with a detailed set of "Release Notes" in flat text format, which accompanies the source code, and describes the specific new features, updates to existing routines, changes in I/O, and detailed descriptions of all bug fixes since the last release.

3.3.5 Archival

CAMx source codes for specific versions are compressed into a single Linux "tar" archive file. All CAMx system tar files are stored on a specific hard disk drive on ENVIRON's data-protected firewalled computer network. This hard disk drive is regularly backed up for off-site storage.

3.4 Validation, Verification, and Testing

3.4.1 Testing Strategy

The code testing strategy is presented in Section 3.3.2. The model application and performance testing against data from two field campaigns are presented in Section 4.

3.4.2 Checking Correctness of Outputs

Per requirements for Category III projects, we will conduct audits of data quality at a level of at least 10% of the data generated by the updated software. Temporary diagnostic output code will be written in the test-bed system to allow for a visual inspection of chemical evolution and all affected variables going into the modified algorithms and solvers. Additional details on the approach to checking correctness of outputs are described in Section 3.3.2.

The new CAMx model will be run and thoroughly evaluated for two field study campaigns (Section 4). Data generated by CAMx will be compared to the output from the original version to ensure that design changes result in expected outcomes. CAMx sensitivity results will be evaluated visually using PAVE and other graphical systems to identify and report the impact of the program changes. Concentrations will be graphically and statistically compared to available measurement data to gauge impacts to model performance.

3.4.3 Determining Conformance to Requirements

The ENVIRON co-Principal Investigator will work closely with the chief code developer to review all new code developments, and with the TAMU co-Principal Investigators to review all testing configurations, applications, and results from the CAMx test applications. Results of all tests and QA/QC procedures will be documented in the Final Report that comprises the project deliverables.

3.5 Documentation, Maintenance, and User Support

3.5.1 Project Documentation Requirements

The project documentation requirements are listed in Section 2, Table 2. A more complete description of all documentation is given in Section 6. ENVIRON will document all program modifications and procedures so that an experienced model user can understand and follow the methodology.

3.5.2 Maintenance and User Support

Code maintenance will be conducted following ENVIRON's standard in-house protocols. These are detailed in Section 3.3.4. ENVIRON provides user support through various channels, including an e-mail user group (camxusers@environ.org) and direct e-mail access (ask-camx@environ.org). The "camxusers" group is meant for broadcasting useful information about CAMx and its pre- and post-processors to all users that subscribe to this group, or to ask about available datasets or issues that other users may have come across. The latter is used to report specific problems or bugs found while using CAMx.

The CAMx web site (www.camx.com) provides the model code, test case I/O, access to the most current CAMx User's Guide, contact information (including references to the above e-mail groups), and access to a series of important CAMx papers and reports that have been published over the past decade.

TCEQ and AQRP staff may contact ENVIRON's CAMx support staff directly for user support. Contact information is listed below:

- www.camx.com
- Ask-camx@environ.org
- Gary Wilson: gwilson@environcorp.com, 415/899-0719
- Chris Emery: cemery@environcorp.com, 415/899-0704

3.5.3 Methods and Maintenance Facilities

The methods and facilities used to maintain, store, secure, and document code versions and related items are described in Sections 3.3.4, 3.3.5, and 3.5.2.

4. MODEL EVALUTION APPROACH AND QUALITY METRICS

The project team will conduct a detailed evaluation of the CAMx results against ambient measurements. TAMU will develop graphical and statistical products with which to facilitate the evaluation, specifically including surface concentration measurements from the TCEQ monitoring system, and in situ concentration measurements and remotely-sensed tropospheric column profiles from the START08 and DISCOVER-AQ databases. The team will collaborate on evaluating these products and identifying any clear problems in the modeling results. Solutions will be developed to improve certain details within the CAMx CiG framework or to fix bugs or performance issues not caught during the code development and testing tasks. Additional model runs will be performed and evaluated as necessary. Particular focus will be placed on how well CAMx captures or evolves features in the profiles of NOx, ozone and other key tracers during monitored convective periods, and how the WRF/CAMx modeling system performs overall in characterizing convection and its impacts to spatial and temporal pollutant distributions.

Results will be compared to CAMx runs using the original (unmodified) program to quantify (through statistical and graphical means) changes in the resulting output fields and model speed. Job scripts and standard output and error files will be reviewed to ensure that the model tests were constructed and applied properly, and that the model ran correctly. Statistical assessments for ozone, precursor and product species will be calculated and compared to the results using the original model inputs.

Data to be accessed and/or developed for this study are of three types:

(1) CAMx forecast modeling system data for 2013 (Johnson et al., 2013);
(2) CAMx modeling system data for 2008 (Emery et al., 2013; McDonald-Buller et al., 2013);
(3) Measurement data from the DISCOVER-AQ (Pickering, 2013) and START08 field campaigns (Pan et al., 2010).

ENVIRON fully evaluated the 2008 CAMx modeling database in a past AQRP project. The modeling system tended to over-predict ozone in the south-central US (including Texas) throughout the year, with the largest over predictions occurring during the warm season (June-August). Emery et al. (2013) attributed the ozone over predictions to chemistry issues (as opposed to transport issues) and the tendency to predict very high ozone over the Gulf, which was then brought inland.

ENVIRON developed the 2013 forecasting database for TCEQ and evaluated forecast performance against measurement data throughout Texas and the Gulf Coast. The model typically over-predicted ozone throughout Texas, and incremental changes to the forecasting configuration were able to ameliorate some of the performance issues. Meteorological modeling will be re-run by TAMU for the September 2013 period using the updated version of WRF, but emission estimates from the forecasting system will continue to be used in these tests.

TAMU will run the modified WRF with various K-F configurations for three test cases. The first subtask will be case development and meteorological validation. The cases have been selected for their variety of convective modes and the availability of measurements in the free troposphere. Then WRF output from each case will be passed through the updated WRFCAMx interface program, and CAMx will be run using pre-existing emissions and ancillary input datasets.

The first case occurred in 2013, during early DISCOVER-AQ operations between September 4 and 11, and particularly on September 6, when substantial shallow and deep convection moved inland from the Gulf of Mexico and across the Houston area (Pickering, 2013). The P3 aircraft sampled vertical columns up to 12,000-15,500 feet at several locations in the Houston-Galveston area before and after the scattered deep convection moved through the area, while the B200 aircraft flew at high altitude with a downward-looking hyperspectral sensor. This case will allow us to examine CAMx performance in the context of the complex emissions profile of the Houston-Galveston airshed. TAMU will commence measurement data acquisition by coordinating with NASA researchers to obtain necessary data from DISCOVER-AQ. Field study data are expected to be available for this project sometime during the first half of 2014 (K. Pickering, pers. comm.).

We are anticipating that validation and quality assurance procedures have been independently conducted on DISCOVER-AQ measurement datasets by the respective reporting institutions to flag missing or suspect data due to a variety of causes. Only final quality-assured un-flagged data will be employed in the comparisons to modeling data. Further information on the DISCOVER-AQ data management plan is provided by NASA (2011).

The next two cases occurred during the Stratosphere-Troposphere Analyses of Regional Transport 2008 (START08). START08 was designed to measure the effects of a wide variety of transport and mixing processes in the upper troposphere and lower stratosphere, including deep convection (Pan et al., 2010). The key platform for START08 was an NCAR/NSF Gulfstream-V aircraft equipped with sensors for measuring O_3, CO, CO_2, NO, NOy, CH_4, N_2O, and 55 other trace species. The first case, on May 6, 2008, featured a squall line over North Texas and has already been simulated with WRF to study the transport effects of explicitly-resolved convection (Bowman 1993, Bowman and Carrie 1992). The second case occurred on June 16, 2008 and was an instance of a multi-day episode of unorganized, diurnally-forced deep convection over the south-central United States. High-altitude and vertical column measurements were made within a triangle bounded by Colorado, the northern Gulf of Mexico, and Missouri.

Following the completion of the START08 field program in 2008, the project data sets were quality controlled by the instrument scientists, the project principal investigators, and the data support staff of the Earth Observation Laboratory (EOL) of the National Center for Atmospheric Research (NCAR) (Pan et al., 2010). Data are archived in the NCAR EOL online data repository (http://data.eol.ucar.edu/master_list/?project=START08). Only final quality-assured un-flagged data will be employed in the comparisons to model simulations.

5. DATA ANALYSIS, INTERPRETATION, AND MANAGEMENT

The spatial and temporal performance of the CAMx photochemical model in simulating measured ozone and precursors during the 2013 DISCOVER-AQ and 2008 START08 episodes will be evaluated using both graphical and statistical methods that describe the range of predictions among the various model combinations. Graphical methods will include spatial maps, profiles, and time-series comparing model predictions to observations. Graphics may be developed using a mix of several plotting applications, including GIS, PAVE, NCAR/NCL, IDL, R, Python, and others. Statistical methods will include computation of metrics for bias and error between predictions and observations for ozone and precursors. Standard statistical metrics as described in EPA air quality modeling guidance (EPA, 2007) will be calculated (Table 3). Linear regression analysis (e.g., coefficient of determination, r^2) will be utilized to examine the model's ability to capture observed variability.

A member of the research team who did not develop the model input datasets or conduct model simulations will review at least 10% of the input data and model output for quality assurance purposes.

Data generated for this project, including model inputs, various air quality observational data and statistical performance calculations, will be securely archived during the project on portable hard drives and stored for a period of at least three years following the completion of the project. Although model output files can be reproduced from the model inputs and model software, we will also archive the optimized model output files for each of the three test episodes. All data obtained for this project will be stored in electronic format. Our teams' experience has been that 100+ GB hard drives provide an accessible and portable system for storing data files of the size routinely encountered in the type of modeling activities for this effort. If data are provided on paper, the paper documents will be scanned to electronic PDF files for storage. The University of Texas will receive an electronic copy of all data sets.

Table 3. Definition of air quality model performance metrics (EPA, 2007).

Metric	Definition[1]
Mean Bias (MB)	$\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)$
Mean Error (ME)	$\frac{1}{N}\sum_{i=1}^{N}\lvert P_i - O_i\rvert$
Mean Normalized Bias (MNB) (-100% to +∞)	$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{P_i - O_i}{O_i}\right)$
Mean Normalized Error (MNE) (0% to +∞)	$\frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{P_i - O_i}{O_i}\right\rvert$
Normalized Mean Bias (NMB) (-100% to +∞)	$\frac{\sum_{i=1}^{N}(P_i - O_i)}{\sum_{i=1}^{N}O_i}$
Normalized Mean Error (NME) (0% to +∞)	$\frac{\sum_{i=1}^{N}\lvert P_i - O_i\rvert}{\sum_{i=1}^{N}O_i}$
Fractional Bias (FB) (-200% to +200%)	$\frac{2}{N}\sum_{i=1}^{N}\left(\frac{P_i - O_i}{P_i + O_i}\right)$
Fractional Error (FE) (0% to +200%)	$\frac{2}{N}\sum_{i=1}^{N}\left\lvert\frac{P_i - O_i}{P_i + O_i}\right\rvert$
Coefficient of Determination (r^2) (0 to 1)	$\left(\frac{\sum_{i=1}^{N}(P_i - \overline{P})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{N}(P_i - \overline{P})^2\sum_{i=1}^{N}(O_i - \overline{O})^2}}\right)^2$

[1] P_i and O_i are prediction and observation at the *i*-th site, respectively; $\overline{P}$ and $\overline{O}$ are mean prediction and observation, respectively.

6. REPORTING

A final Technical Work Plan (statement of work, schedule, key personnel) and Quality Assurance Project Plan (this document) will be submitted to AQRP for review and approval. Monthly financial and technical progress reports will be submitted throughout the duration of the project.

A project report will be developed during the course of the work that fully documents the methodology, assumptions, and evaluation results from Tasks 1 through 4. It will also include documentation of the specific quality assurance steps, associated findings, and any necessary corrective actions taken to rectify data quality issues related to the software updates described above. Conclusions will include recommendations to address near-term issues associated with the new CAMx cloud treatment, as well as recommendations for longer-term research. A draft will be submitted to AQRP and TCEQ for review by May 18, 2015. A final report addressing review comments received on the draft will be subsequently submitted to AQRP by June 30, 2015. The final Project Report will be delivered electronically and will meet State of Texas Accessibility requirements in 1 TAC 213. The project approach and results will also be summarized in presentation format for use at an AQRP meeting to be held in June 2015.

Electronic copies of all text, graphic, spreadsheet files and models used in the preparation of any documents related to the project reports, to document results and conclusions (e.g. sampling data, work files, etc.) or developed as work products under this Contract, shall be supplied the conclusion of the project. All copies of deliverable documents and other work products will be provided in Microsoft Word and PDF format. The WRF, WRFCAMx, CAMx and other pre-processor codes, associated scripts, and all input/output data and analysis products generated under Task 4 will be delivered to the AQRP and TCEQ at the completion of the project via high-volume disk media. The data will be submitted in a format that will allow AQRP or TCEQ or other outside parties to utilize the information.

After completion of the project, if warranted on the basis of our modeling results, the project team will collaborate on developing a manuscript from the project report for submission to an appropriate and mutually-agreed scientific journal. A draft will first be submitted to AQRP and TCEQ for review. A final draft will be submitted to an appropriate and journal for peer review. Mr. Emery and Dr. Nielsen-Gammon will supervise the completion of all reports, presentations, and manuscripts, which will be collaborative efforts between the ENVIRON and TAMU team.

7. REFERENCES

1) Alapaty, K., J.A. Herwehe, T.L. Otte, C.G., Nolte, O.R. Bullock, M.S. Mallard, J.S. Kain, J. Dudhia, 2012. Introducing subgrid-scale cloud feedbacks to radiation for regional meteorological and climate modeling. *Geophys. Res. Lett.*, 39, L24809, doi:10.1029/ 2012GL054031.
2) Belikov, D.A., et al., 2013. Off-line algorithm for calculation of vertical tracer transport in the troposphere due to deep convection. *Atmos. Chem. Phys.*, 13, 1093-1114, doi:10.5194/acp-13-1093-2013.
3) Bowman, K.P., 1993. Large-scale isentropic mixing properties of the Antarctic polar vortex from analyzed winds. *J. Geophys. Res.*, 98, 23,013-23,027, doi: 10.1029/ 93JD02599.
4) Bowman, K.P. and G.R. Carrie, 2002. The mean-meridional transport circulation of the troposphere in an idealized GCM. *J. Atmos. Sci.*, 59, 1502-1514, doi: 10.1175/1520-0469(2002)059<1502:TMMTCO>2.0.CO;2.
5) EPA, 2007. Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, PM2.5 and Regional Haze. US Environmental Protection Agency, Research Triangle Park, NC. EPA-454/B-07-002. April. (http://www.epa.gov/ttn/scram/guidance/guide/final-03-pm-rh-guidance.pdf).
6) Emery, C., J. Jung, J Johnson, G. Yarwood, S. Madronich, G. Grell, 2010. Improving the Characterization of Clouds and Their Impact on Photolysis Rates Within the CAMx Photochemical Grid Model, Final Report: Work Order 582-7-84005-FY10-23. Prepared for the Texas Commission on Environmental Quality, Austin, TX, by ENVIRON International Corporation, Novato, CA (August).
7) Emery, C., B. Koo, T. Sakulyanontvittaya, G. Yarwood, 2013a. Improving CAMx GREASD PiG Efficiency, Final Report: Work Order 582-11-10365-FY13-09. Prepared for the Texas Commission on Environmental Quality, Austin, TX, by ENVIRON International Corporation, Novato, CA (August).
8) Emery, C., E. Tai, G. Yarwood, M. Lin, 2013b. Using Global and Regional Models to Represent Background Ozone Entering Texas, Final Report: AQRP Project 12-011. Prepared for the Texas Air Quality Research Program, University of Texas at Austin, and the Texas Commission on Environmental Quality, Austin, TX, by ENVIRON International Corporation, Novato, CA, and Princeton University/NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NY (November).
9) ENVIRON, 2014. User’s Guide: Comprehensive Air quality Model with extensions, v6.10. Prepared by ENVIRON International Corporation, May 2013 (www.camx.com).
10) Foley, K.M., Roselle, S.J., Appel, K.W., Bhave, P.V., Pleim, J.E., Otte, T.L., Mathur, R., Sarwar, G., Young, J.O., Gilliam, R.C., Nolte, C.G., Kelly, J.T., Gilliland, A.B., Bash, J.O. 2010. Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7. *Geosci. Model Dev.*, 3, 205-226, doi:10.5194/gmd-3-205-2010, 2010.
11) Johnson, J., E. Tai, P. Karamchandani, G. Wilson, G. Yarwood, 2013. TCEQ Ozone Forecasting System, Final Report: Work Order 582-11-10365-FY13-13. Prepared for the Texas Commission on Environmental Quality, Austin, TX; by ENVIRON International Corporation, Novato, CA, August.
12) Kain, J.S., 2004. The Kain-Fritsch convective parameterization: An update. *J. Appl. Meteor.*, 43, 170–181.

13) Kemball-Cook, S., J. Johnson, G. Yarwood, 2012. Evaluating TCEQ NOx Emission Inventories Using Satellite NO2 Data, Final Report: Work Order 582-11-10365-FY12-08. Prepared for the Texas Commission on Environmental Quality, Austin, TX; by ENVIRON International Corporation, Novato, CA, August.
14) Kemball-Cook, S., J. Johnson, G. Yarwood, 2013. Continuation on Use of Satellite Nitrogen Dioxide (NO2) Data, Final Report: Work Order 582-11-10365-FY13-10. Prepared for the Texas Commission on Environmental Quality, Austin, TX; by ENVIRON International Corporation, Novato, CA, August.
15) McDonald-Buller, E., Y. Kimura, C. Wiedinmyer, C. Emery, E. Tai, 2013. The Effects of Uncertainties in Fire Emissions Estimates on Predictions of Texas Air Quality, Final Report: AQRP Project 12-018. Prepared for the Texas Air Quality Research Program, University of Texas at Austin, and the Texas Commission on Environmental Quality, Austin, TX, by the University of Texas at Austin, the National Center for Atmospheric Research, Boulder, CO, and ENVIRON International Corporation, Novato, CA (November).
16) NASA, 2011. Earth Venture 1: Data Management Plan. Prepared by the National Aeronautics and Space Administration, DISCOVER-AQ Science Directorate, Langley Research Center (D-AQ-DMP-01, Version 1, 6/22/2011).
17) Pan, L.L., K.P. Bowman, E.L. Atlas, S.C. Wofsy, F. Zhang, J.F. Bresch, B.A. Ridley, J.V. Pittman, C.R. Homeyer, P. Romashkin, W.A. Cooper, 2010. The Stratosphere-Troposphere Analyses of Regional Transport 2008 (START08) Experiment. *Bull. Amer. Meteor. Soc.*, 91, 327-342, doi: 10.1175/2009BAMS2865.1.
18) Pickering, K., et al., 2013. DISCOVER-AQ Houston 2013 mission web site (http://www-air.larc.nasa.gov/missions/discover-aq/discover-aq.html).
19) Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, M.G. Duda, X.-Y. Huang, W. Wang, J.G. Powers, 2008. A Description of the Advanced Research WRF Version 3. Prepared by the National Center for Atmospheric Research, NCAR Technical Note NCAR/TN–475+STR (June 2008).
20) Zhao, C., Y. Wang, T. Zheng, 2009. Summertime impact of convective transport and lightning NOx production over North America: modeling dependence on meteorological simulations. *Atmos. Chem. Phys.*, 9, 4315-4327, www.atmos-chem-phys.net/9/4315/2009/.