# Using the 3SAQS 2008 Modeling Platform for NEPA Modeling

3SAQS\_NEPA08 Version 07/2014

1. **Introduction**

The Three State Data Warehouse (3SDW) under the Three State Air Quality Study (3SAQS) developed an air quality modeling platform for conducting air pollution environmental impact assessments pursuant to the National Environmental Policy Act (NEPA). Specifically designed to evaluate the air impacts of emissions from oil and gas (O&G) production, the modeling platform includes a complete set of data and software to conduct regional air quality modeling for the base and future years of 2008 and 2020. The NEPA modeling package includes both emissions and air quality modeling data and software. The emissions component of the NEPA package uses SMOKE (3SAQS\_SMOKE\_Base08b) to prepare oil and gas inventories for input to the CAMx modeling component of the NEPA package (3SAQS\_CAMx\_Base08a).

The emissions data in the 3SAQS 2008 NEPA modeling platform (3SAQS\_NEPA08) are configured for O&G emissions sensitivity modeling. The non-O&G emissions sectors (e.g. mobile, electricity generation, biogenic, fires) are merged together into a single daily emissions file. The O&G emissions sectors are separated into point and non-point sources. Users of these data should only modify the O&G sectors and then merge these files together with the non-O&G emissions to create CAMx-ready emissions input files.

The 3SAQS\_NEPA08 platform includes all of the data, software, and scripts required to prepare, simulate, and post-process estimates of 2008 air quality in the Western U.S.

Direct questions about the 3SAQS\_NEPA08 platform to Zac Adelman (zac@unc.edu).

1. **Overview of 3SAQS NEPA Modeling Process**

This document provides guidance on the following modeling steps:

* Install and benchmark CAMx
* Install SMOKE
* Prepare project O&G emissions using SMOKE
* Merge project O&G emissions with non-O&G emissions using SMOKE
* Download CAMx inputs and run CAMx with the new project O&G emissions
* Post process CAMx results and perform analysis of results

Start the installation process by creating a base working directory on your system. Unless specified otherwise, install everything described here in the first tier of the base directory. For example, you could create a directory called /home/NEPA/Base08 and alias this directory using the environment variable $NEPA\_HOME. We will use this example alias throughout this document to reference a base installation directory.

1. **Installing and compiling CAMx**
	1. Notes on the 3SAQS Base 2008 version A CAMx (3SAQS\_CAMx\_Base08a) benchmark data provided by 3SDW

The University of North Carolina (UNC) generated benchmarking data using the following system and compilers

Operating System: 64-bit RHEL 5.6 (Tikanga)

Hardware: Dell C6100 12-core nodes with 2.93 GHz Intel processors, 12M L3 cache (Model X5670)

Fortran Compiler: Portland Group Fortran version 11.6-0 (pgf90 11.6-0 64-bit target on x86-64 Linux -tp Nehalem)

MPI: mvapich2-1.7

CAMx: v5.41 hybrid MPI-OpenMP

Reproducing this configuration will give the best chance of exactly reproducing the CAMx results from the 3SDW. Along with differences in compilers, hardware and processor differences will contribute to results that diverge from the 3SDW simulations. The differences in the CAMx results due to compiler and hardware variations should be less than 1% for both gas and aerosol species.

**NOTE: The CAMx v5.41 model benchmarking is done with 3SAQS Base 2008 version A emissions data. These data are slightly different from the version B emissions and should only be used for benchmarking the modeling software installation.**

* 1. Download the CAMx Benchmark source codes and data from the following ftp address:

**ftp://viking.cira.colostate.edu/3SDW/3SAQS/2008/Base08a/AQ/CAMx/Benchmark**

* 1. Untar the CAMx benchmark model archive and data in a working directory:

tar xvzf 3SAQS\_CAMx\_Base08a\_Benchmark\_MODEL.tar.gz

tar xvzf 3SAQS\_CAMx\_Base08a\_Benchmark\_INPUTDATA.tar.gz

tar xvzf 3SAQS\_CAMx\_Base08a\_Benchmark\_METDATA.tar.gz

tar xvzf 3SAQS\_CAMx\_Base08a\_Benchmark\_REFDATA.tar.gz

After untarring these files you will see the following directories:

**3SAQS\_B08a\_Benchmark**

* **inputs**
	+ ahomap
	+ chem
	+ emis
	+ low
		- * 36km
* 12km
	+ ptsr
	+ icbc
	+ landuse
	+ met
	+ westjump.12km.25lay
	+ cr
	+ kv
	+ tp
	+ uv
	+ wv
	+ zp
	+ westjump.36km.25lay
	+ cr
	+ kv
	+ tp
	+ uv
	+ wv
	+ zp
	+ tuv
* **outputs**
* **post-proc**
* m3diff
* camx2ioapi
* **ref**
* **runfiles**
* **src**
	1. Change directories to the source code directory and build CAMx for your system configuration

cd 3SAQS\_B08a\_Benchmark/src

Set the Makefile environment variable MPI\_INST to the installation location of the MPI library and include files on you system.

Save the Makefile and run the make command to the build the CAMx executable:

 make pg\_linuxomp MPI=true |& tee CAMx\_build.log

This example shows the command used to build the CAMx executable used for the 3SAQS\_CAMx\_Base08a simulation. If you use a different compiler or MPI configuration see the syntax notes in the CAMx Makefile for guidance.

* 1. Configure the CAMx run script run.3SAQS.3612K.25L.Base08a.Q1.job. From the CAMx home directory, change to the runfiles directory and open the CAMx run script in a text editor. Set the HOME variable to the location of the top level CAMx installation directory. Set the MPIBIN variable directory to the installation location of the MPI library and include files on your system.

**NOTE: Do not change the file input or output naming conventions in this script as the downstream post-processing programs depend on these names.**

* 1. Run CAMx for the 3SAQS benchmark simulation. The CAMx 3SAQS benchmark case is a 5-day period from January 1-5, 2008 on 36 and 12-km nested grids. The run script is configured to initialize the simulation from a CAMx restart file and to use boundary conditions for the outer domain downscaled from a MOZART GCM simulation.

The provided run script is configured for the UNC computing cluster. You will now need to configure the CAMx run script for a parallel simulation on your system. For the 3SAQS Base08a simulation, including this benchmark simulation, UNC used 24 processors in hybrid MPI-OpenMP mode. The simulation was run on 6 nodes with 4 processors per node.

* 1. Post-process the CAMx output data by converting it to I/O API-netCDF format.

Source code, executable, and scripts are located in:

**3SAQS\_B08a\_Benchmark/post-proc/camx2ioapi**

You should be able to use the precompiled executable for this test. If you would like to compile the program from source, configure the Makefile to point to the I/O API and netCDF installations on your system. The distributed Makefile is configured for the Portland Group Fortran compiler. This can be changed to work for other compilers. Successful compilation of the program will result in the executable file camx2ioapi\_v6.

To configure the CAMx2IOAPI run script camx2ioapi\_v6\_3SAQS.job set the CAMXDIR variable to the root directory path of the 3SAQS Benchmark installation on your system. Save and run this script to create netCDF layer 1 concentration, wet deposition, and dry deposition files.

* 1. Compare the benchmark simulation against the reference data results.

Scripts and executable are located in:

**3SAQS\_B08a\_Benchmark/post-proc/m3diff**

The precompiled executable for the I/O API utility program m3diff will be used to calculate daily max, min, and average differences between the benchmark simulation and reference data. The m3diff script compares the daily average differences and reports a message if the difference is greater than a defined threshold value.

**Note: The benchmark simulation is successful (i.e. differences less than the threshold value) if you DO NOT see the message *“Performance threshold exceeded for [CO|O3|NH3|PEC] on day [1-5] for hour [1-24].”***

By default the m3diff script uses a threshold of 0.1% (environment variable THRESH). If the daily average difference between the benchmark and reference data is > THRESH x reference data daily average, the script writes a message to the screen.

The m3diff script is configured to check average concentrations, wet deposition, and dry deposition for the pollutants CO, O3, NH3, and PEC for layer 1 only.

1. **Install SMOKE and Emissions Input Data**
	1. Download the SMOKE software and CAMx emissions utilities from the following ftp address:

**ftp://viking.cira.colostate.edu/3SDW/3SAQS/2008/Base08b/EMIS/software/**

* 1. Untar the SMOKE and CAMx utility software packages in a working directory:

cd $NEPA\_HOME

tar xvzf 3SAQS\_Base08b\_SMOKE\_17July2014.tar.gz

tar xvzf 3SAQS\_Base08b\_camxutil\_17July2014.tar.gz

* 1. Configure SMOKE for your system

Navigate to the SMOKE Assigns directory and change the Assigns file environment variable SMK\_HOME to the location of the SMOKE installation on your system.

cd 3SAQS\_SMOKE\_Base08b/SMOKE/subsys/smoke31/assigns

Edit the file ASSIGNS.3SAQS\_base08b.camx.cb05 and change the setting for the variable SMK\_HOME on line 41 to the full path of the SMOKE installation directory on your system.

Source the Assigns file:

source ASSIGNS.3SAQS\_base08b.camx.cb05

* 1. Compile the CAMx emissions utility programs

Navigate to the directory with the CAMx merge programs and compile the software. Note that the default Makefiles are configured to use the Intel Fortran compiler. You’ll need to change the compiler and compiler flags in these scripts if you are using something other than Intel.

cd $NEPA\_HOME/3SAQS\_SMOKE\_Base08b/merge\_camx\_files

 cd src/low

make

cd ../ptsmrg

make

* 1. Install the SMOKE input data and logs

**ftp://viking.cira.colostate.edu:/3SDW/3SAQS/2008/Base08b/EMIS/inputs/3SAQS\_Base08b\_OG\_EmissionsData\_17July2014.tar.gz**

**ftp://viking.cira.colostate.edu/3SDW/3SAQS/2008/Base08b/EMIS/inputs/3SAQS\_Base08b\_SMOKE\_Logs.tar.gz**

Untar the emissions data in your $NEPA\_HOME directory. This archive contains ancillary data for all sources and inventory data for O&G sources only. By untarring this archive into the base project directory (e.g. $NEPA\_HOME) it will install the data in the correct directories relative to your SMOKE installation.

The logs are a record of the SMOKE runs completed by the 3SAQS team and could be used as a reference for the SMOKE runs conducted on your system. Untar the SMOKE logs in an area separate from your working SMOKE directory for this project.

* 1. Install the SMOKE non-O&G processed emissions files

The SMOKE non O&G emissions files are SMOKE-output data in CAMx binary format. The daily files include hourly emissions for all non-O&G sources merged together in a single file. We merged all of the non-O&G sources together to reduce the number of files that we needed to distribute for the NEPA modeling package. As described in the introduction, users should process their project oil and gas data through SMOKE (see Section 5 below) and then merge these files with the non-O&G emissions (see Section 6 below).

If you have not already done so, source the SMOKE Assigns file for your system and then navigate to the SMOKE output directory to install the merged non-O&G emissions files.

cd $NEPA\_HOME 3SAQS\_SMOKE\_Base08b/SMOKE/subsys/smoke31/assigns

source ASSIGNS.3SAQS\_base08b.camx.cb05

cd $OUTPUT

Download the low level and elevated 36-km non-O&G emissions files into the SMOKE 36-km $OUTPUT directory:

**ftp://viking.cira.colostate.edu:/3SDW/3SAQS/2008/Base08b/EMIS/CAMx/no\_OG/low\_36km**

**ftp://viking.cira.colostate.edu:/3SDW/3SAQS/2008/Base08b/EMIS/CAMx/no\_OG/elev**

Download the low-level 12-km non-O&G emissions files into the SMOKE 12-km $OUTPUT directory:

cd $OUTPUT/../camx.cb5tx.12km

**ftp://viking.cira.colostate.edu:/3SDW/3SAQS/2008/Base08b/EMIS/CAMx/no\_OG/low\_12km**

NOTE: The elevated emissions files are grid-independent and are only needed for the 36-km modeling grid.

1. **Prepare project O&G emissions using SMOKE**

Source the SMOKE Assigns file (see Section 4.3) and navigate to the SMOKE inventory data directories

cd $INVDIR

 The four directories with oil and gas inventory data include:

 **ptog** – 3SAQS basin point O&G inventories and list file

 **arog** – 3SAQS basin non-point O&G inventories and list file

 **usptog** – non-3SAQS basin point O&G inventories and list file

 **usarog** – non-3SAQS basin non-point O&G inventories and list file

The list files in these directories include the directory paths and file names that SMOKE will use to process the inventory data. You will need to replace the string “INSTALLDIR” with the base SMOKE installation directory on your system before processing these data with SMOKE. Along with adding the correct directory paths of the inventories on your system, configure the list files to include any new O&G project inventories.

These processing sectors (i.e. ptog, arog, usptog, usarog) were used for the 3SAQS simulations. They can either be augmented with new project data or used as templates when adding a completely new processing sector to the SMOKE installation.

The SMOKE general data (GE\_DAT) directory includes the temporal, spatial, and chemical speciation data needed for processing the emissions. Users must configure the 3SAQS ancillary emissions data to apply the proper profiles to any new O&G inventories. As with the inventory list files, the existing 3SAQS ancillary data may be adapted for use with new project inventories or used as templates for adding now profiles that describe the temporal/spatial/chemical characteristics of any new inventory sources.

Once the inventory and ancillary data are configured for input to SMOKE, use the run scripts under $SCRIPTS/run/all\_grids/base08b to process the O&G data with SMOKE. The O&G scripts for the four sectors listed above can be used as templates for processing new data with SMOKE. The following SMOKE programs must be run when processing new emissions sources:

* Smkinven (RUN\_SMKINVEN = Y) – import new inventories into SMOKE
* Grdmat (RUN\_GRDMAT = Y) – calculate gridding coefficients for allocating emissions to model grid cells
* Spcmat (RUN\_SPCMAT = Y) – calculate NOx, VOC, and PM chemical speciation factors for converting inventory pollutants to chemistry-transport model species
* Temporal (RUN\_TEMPORAL = Y) – convert annual/daily inventory data to hourly emissions
* Elevpoint (RUN\_ELEVPOINT = Y) – select low-level and elevated point sources; for point source inventories only
* Smkmerge (RUN\_SMKMERGE = Y) – merge all of the intermediate steps together to create gridded/temporalized/speciated emissions estimates
* Smk2emis (RUN\_SMK2EMIS = Y ) – convert output low-level emissions to CAMx format and units
* Mrgelev (RUN\_MRGELEV = Y ) – convert output elevated emissions to CAMx format and units

After processing the O&G emissions with SMOKE they will need to be merged with the non-O&G data to finalize the emissions for input to CAMx.

1. **Merge project O&G emissions with non-O&G emissions using CAMx utilities**

Once the O&G emissions are processed with SMOKE use the CAMx emissions merge utilities to combine the low level and elevated emissions into CAMx emission input files. Navigate back to the directory with the CAMx merge utilities and configure the run scripts to point to the emissions files on your system.

cd $NEPA\_HOME/3SAQS\_SMOKE\_Base08b/merge\_camx\_files/run\_scripts

The low level and elevated sources need to be merged separately. As the elevated files are grid-independent they only need to be merged for the outer, 36-km modeling domain. Configure the following variables in the elevated source run script **mrgpts\_OG.3saqs\_camx\_36km.job**:

* ASSIGNS – SMOKE Assigns file directory path
* noOG\_DIR – directory path with the merged non-O&G emissions files downloaded from the 3SDW
* OUTDIR – directory path for the final O&G + non-O&G emissions

The default script is set up to merge two O&G elevated point file sectors in with the non-O&G data: ptog and usptog. Configure the script to work with the elevated point O&G files that you created for the new project inventories.

As the low-level emissions files are grid-dependent, there are two scripts for merging the new project O&G emissions with the non-O&G data. Use the script **mrglow\_OG.3saqs\_36km\_camx.job** for the 36-km modeling domain and the script **mrglow\_OG.3saqs\_12km\_camx.job f**or the 12-km modeling domain. Like the elevated source script, configure the script variables for the Assigns directory and the input/output directories. These scripts are set up to merge one O&G low-level emissions sector in with the non-O&G data. They assume that you ran the SMOKE program Mrggrid to combine all of the low-level O&G emissions together into a single file before merging with the non-O&G emissions. An alternative to this approach is to merge each individual O&G sector in with the non-O&G emissions in sequence using the CAMx merge utility.

1. **Run CAMx with the new project O&G emissions**

Download and install the 3SAQS Base08b CAMx modeling package in your NEPA\_HOME directory:

**ftp://viking.cira.colostate.edu/3SDW/3SAQS/2008/Base08b/AQ/CAMx/software/3SAQS\_CAMxv5.41\_B08b.tar.gz**

 **I**nstall the CAMx input files in the directory $NEPA\_HOME/3SAQS\_CAMx\_B08b/inputs:

**ftp://viking.cira.colostate.edu:/3SDW/3SAQS/2008/Base08b/AQ/CAMx/inputs/ahomap**

**ftp://viking.cira.colostate.edu:/3SDW/3SAQS/2008/Base08b/AQ/CAMx/inputs/icbc**

**ftp://viking.cira.colostate.edu:/3SDW/3SAQS/2008/Base08b/AQ/CAMx/inputs/tuv**

**ftp://viking.cira.colostate.edu:/3SDW/3SAQS/2008/Base08b/AQ/CAMx/inputs/landuse**

**ftp://viking.cira.colostate.edu:/3SDW/3SAQS/2008/Base08b/AQ/CAMx/inputs/chem**

**ftp://viking.cira.colostate.edu:/3SDW/WestJumpAQMS/MET/WRFCAMx/36km**

**ftp://viking.cira.colostate.edu:/3SDW/WestJumpAQMS/MET/WRFCAMx/12km**

Navigate to the run script directory and configure the CAMx run script for your system.

cd $NEPA\_HOME/3SAQS\_CAMx\_B08b/runfiles

The CAMx run script **run.3SAQS.3612K.25L.Base08b.Q1.job** is set up for running a two-way nested 36-km and 12-km simulation. Other key run script variables include:

* Configure the **HOME** variable in the script to point to the base CAMx installation directory on your system.
* Set the variable **EXEC** to point to the CAMx executable that you compiled for the benchmarking exercise in Section 3.4.
* Set the **START\_DATE** and **END\_DATE** variables for the simulation start and end dates.
* Set **MPIBIN** to the MPI installation directory on your system if you’re using MPI for parallelization
* Set **NPROCS** to specify the number of processors for a multi-processor simulation

After configuring the script for your system, run it and redirect the output to a log file.

1. **Post process CAMx results and perform analysis of results**

A CAMx post-processing package includes two programs:

* **camx2ioap**i – converts the CAMx outputs to I/O API-netCDF format
* **combine** – extracts layer 1 CAMx variables and combines them into chemical species for comparison to surface observations

Download and install these post processing programs in the $NEPA\_HOME directory

**ftp://viking.cira.colostate.edu/3SDW/3SAQS/2008/Base08b/AQ/CAMx/software/3SAQS\_postproc\_B08b.tar.gz**

First run the program camx2ioapi to create netCDF versions of the CAMx average concentration and deposition output files. Then run the program Combine to create monthly netCDF CAMx output files.

The outputs from Combine can then be used with the programs sitecmp, cmp\_airs, and cmp\_castnet to pair the model output with observations from various surface networks. The site compare programs are available on the 3SDW:

**ftp://viking.cira.colostate.edu/3SDW/3SAQS/2008/Base08b/AQ/CAMx/software/site\_compare.tar.gz**

An alternative post-processing package is available from Environ called CAMxpost. This program works directly with the CAMx output files and also pairs the model output with observations.

[**http://www.camx.com/camx/media/camx/Files/camxpost-02oct09.tgz**](http://www.camx.com/camx/media/camx/Files/camxpost-02oct09.tgz)