**“Data Dictionary” for files associated with research effort “*Extending the Community Multiscale Air Quality (CMAQ) Modeling System to Hemispheric Scales: Overview of Process Considerations and Initial Applications”***

**Nomenclature**: All files are named as following: FigureXX\_<description>, where Figure XX corresponds to the figure number described in the research effort manuscript. The files are in 3 formats: text, netcdf or Excel. The data sets used in construction of each of the figures in the research effort is further described below along with which components are available on ScienceHub and those that cannot be uploaded to files size restrictions.

**Figure1\_tracer\_netcdf.zip**: this zip files consists of 4 netcdf files. Each files contains information on 4 tracer species nominally representing the lateral boundary condition specification at different altitudes as follows:

O3\_BC1: Layers 1-21 (surface to ~750 mb)

O3\_BC2: Layers 22 – 31 (~750 mb - ~250 mb)

O3\_BC3: Layers 32 – 35 (~250 mb - top)

O3\_BC4: O3 all layers

**Figure2\_FullLayerHt\_ZF\_35\_44L**: text files with layer-top heights

**Figure3a\_35\_44L\_zero\_NA\_TrinidadHead, Figure3b\_35\_44L\_zero\_NA\_Boulder, Figure3c\_35\_44L\_zero\_NA\_Huntsville, Figure3d\_35\_44L\_zero\_NA\_Narragansett**: text files with data used on constructing Figure 3. Header on first row describes the variables in the columns.

**Figure4a\_o3sonde\_march\_mean\_hemi\_trinidad, Figure4b\_o3sonde\_august\_mean\_hemi\_trinidad:** text files with data used on constructing Figure 4. Header on first row describes the variables in the columns.

**Figure5a\_428\_dc8\_model\_obs\_1s, Figure5b\_501\_dc8\_model\_obs\_1s, Figure5c\_509\_dc8\_model\_obs\_1s, Figure5d\_512\_dc8\_model\_obs\_1s, Figure5e\_dc8\_modobs\_mean\_bins\_417\_501, Figure5f\_dc8\_modobs\_mean\_bins\_501\_515:** text files with data used on constructing Figure 5. Header on first row describes the variables in the columns.

**Figure6c\_421\_c130\_model\_obs\_so4\_halogen\_func**: text files with data used on constructing Figure 6c. Header on first row describes the variables in the columns.

**Figure7a\_dc8\_modobs\_so4\_halogen\_func\_mean\_417\_515, Figure7b\_C130\_modobs\_so4\_halogen\_func\_mean\_421\_515:** text files with data used on constructing Figure 7. Header on first row describes the variables in the columns

**Figure8\_LAY1\_CONC\_netcdf.zip**: this zip files contains 2 netcdf files each with 24-hr average PM concentrations for simulations with and without wind-blown dust emissions:

24hravg\_COMBINE\_CONC\_LAY1\_DUST\_PM\_715\_806\_2006 (with dust emissions)

24hravg\_COMBINE\_CONC\_LAY1\_NODUST\_PM\_715\_806\_2006 (without dust emissions)

Figure 8 of the research effort show the difference between these two data files.

**Figure9a\_Bias\_Diff\_SE\_states\_jul29-aug3**: This Excel file contains data that is plotted on the map in Figure9a

**Figure9b\_FL\_PM25\_timeseries\_compare**: text files with data used on constructing Figure 9b. Header on first row describes the variables in the columns.

**Figure10**: Due to file size restrictions, data for this figure cannot be uploaded to EPA ScienceHub. Interested users should contact the research effort lead for the data.

/work/MOD3DEV/xjxing/hemisphere/wrfcmaq/output\_hemi\_nf\_rrtmg\_20\_5\_1\_v3450\_2006\_halogen\_func/avg\_combine\_S\_CCTM\_CONC\_2006.nc

/work/MOD3DEV/xjxing/hemisphere/wrfcmaq/output\_hemi\_nf\_rrtmg\_20\_5\_1\_v3450\_2006\_halogen\_nopv/avg\_combine\_S\_CCTM\_CONC\_2006.nc

/work/MOD3DEV/xjxing/hemisphere/wrfcmaq/output\_hemi\_nf\_rrtmg\_20\_5\_1\_v3450\_2006\_halogen\_pv20/avg\_combine\_S\_CCTM\_CONC\_2006.nc

**Figure11\_CASTNET\_PV\_STEInfluence\_vs\_BiasChange\_seasonal\_O3gt40ppb**: text files with data used on constructing Figure 11. Header on first row describes the variables in the columns.

**Figures 12 & 13**: Due to file size restrictions, data for this figure cannot be uploaded to EPA ScienceHub. Interested users should contact the research effort lead for the data.

Script: /work/MOD3DEV/jae/hemisphere/wrfcmaq/column/hemiout/plot.HEMI.108km.ncl

/work/MOD3DEV/jae/hemisphere/wrfcmaq/column/hemiout/calc\_region\_no2.ncl

Data: /asm2/MOD3DEV/jae/hemisphere/wrfcmaq/column/sciamachy.tar

**Figure14\_CMAQ\_SCIAMACHY\_GOME**: Excel file containing regionally average monthly mean values for NO2 vertical column density for CMAQ, GOME and SCIAMACHY

**Figure15a\_O3\_dm8hr\_JJA\_avg\_trend\_obs\_hemis\_36km, Figure15b\_SO2\_SO4\_JJA\_trendcompare\_CASTNET\_hemis\_36km**: text files with data used on constructing Figure 15. Header on first row describes the variables in the columns

**Figure16a\_EChina\_AOD\_SWR\_JJAmonthly, Figure16b\_Europe\_AOD\_SWR\_JJAmonthly, Figure16c\_EUS\_AOD\_SWR\_JJAmonthly**: text files with data used on constructing Figure 15. Header on first row describes the variables in the columns