**Data sets used in the analysis presented in the manuscript** “Description and evaluation of the Community Multiscale Air Quality (CMAQ) model version 5.1”

The datasets used for the analysis presented in this manuscript were provided by researchers in EPA/ORD/NERL/CED. In particular, the EPA/ORD/NERL/CED researchers provided surface mixing ratios for daily maximum 8-hr average ozone observations and WRF/CMAQ simulations for the 1990 – 2010 time period. The data were provided in 12 monthly files for each of the 21 years, i.e. 252 individual files. Each file is in comma-separated value (csv) format and contains paired observed and simulated daily maximum 8-hr average ozone mixing ratios at all stations for which monitoring data were available during that month. The first six lines of each file contain header information.

The data column in the files are as follows:

Column 1: “state” – two-character ID of the state containing the monitor

Column 2: “county” – three-character ID of the county containing the monitor

Column 3: “state” – four-character ID of the site at which the monitor is located

Column 4: “Longitude” – longitude (in degrees) of the monitor site location

Column 5: “Latitude” – latitude (in degrees) of the monitor site location

Column 6: “column” – integer indicating the model grid cell column containing the monitor

Column 7: “row” – integer indicating the model grid cell row containing the monitor

Column 8: “date” – date (mm/dd/yyyy) of the observed and modeled values

Column 9: “O3\_1hrmax” – observed daily maximum 1-hr ozone in ppb

Column 10: “O3\_1hrmax\_time” – hour when the observed daily maximum 1-hr ozone occurred

Column 11: “O3\_1hrmax” – WRF/CMAQ daily maximum 1-hr ozone in ppb

Column 12: “O3\_1hrmax\_9cell” – WRF/CMAQ daily maximum 1-hr ozone in ppb, considering the grid cell containing the monitor as well as the surrounding eight grid cells

Column 13: “O3\_8hrmax” – observed daily maximum 8-hr average ozone in ppb

Column 14: “O3\_8hrmax\_time” – hour when the observed daily maximum 8-hr average ozone occurred

Column 15: “O3\_8hrmax” – WRF/CMAQ simulated daily maximum 8-hr average ozone in ppb

Column 16: “O3\_8hrmax\_9cell” – WRF/CMAQ daily maximum 8-hr average ozone in ppb, considering the grid cell containing the monitor as well as the surrounding eight grid cells

Column 17: “W126” – cumulative W126 exposure index calculated from observations in ppm-hours

Column 18: “W126” – cumulative W126 exposure index calculated from WRF/CMAQ simulations in ppm-hours

Only the data contained in columns 4, 5, 13 and 15 was used in the analysis presented in this manuscript.

The individual files are named “AQS\_Daily\_\*\*\*\*.csv”, where \*\*\*\* is the simulation name. Addition .csv files were also utilized in this analysis and follow a similar structure and naming convention as described above. Specifically, they contain metadata in the first several columns followed by paired observed and simulated values for various species.

**Original WRF/CMAQ Model Data**

The data from the original WRF/CMAQ model simulations from which the attached data were extracted are very large (several terabytes) and cannot be uploaded to ScienceHub due to size restrictions. These original WRF/CMAQ model simulations are stored on the /asm archival system accessible through the sol high-performance computing system. Due to data management policies, files on /asm are subject to expiry depending on the template of the project. Files not requested for extension after the expiry date are deleted permanently from the system. The format of the files used in this analysis and listed below is ioapi/netcdf. Documentation of this format, including definitions of the geographical projection attributes contained in the file headers, are available at <https://www.cmascenter.org/ioapi/documentation/3.1/html/AA.html>

Seasonal average ioapi files are archived that contain ozone and PM2.5 and were used to create the model/model difference plots included in the manuscript. These seasonal files are provided in zip files with the names:

CMAQv502\_Base\_aconc.zip

CMAQv51\_Base\_NEIv1\_aconc.zip

CMAQv51\_Base\_NEIv2\_aconc.zip