CMAQ Data Available on the Cloud through Amazon's Open Data Program
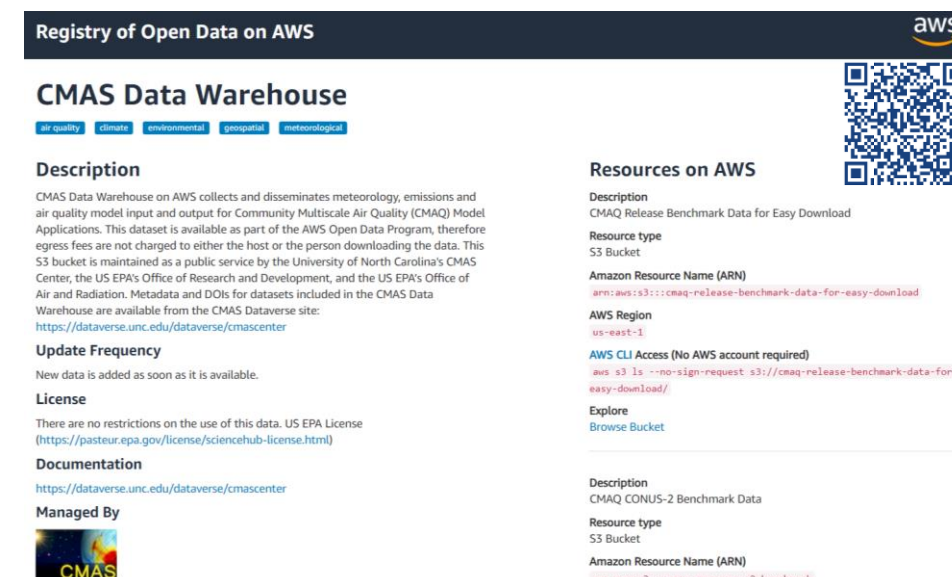
Fahim Sidi[1], Kristen Foley[1], George Pouliot[1], Alison Eyth[2], Liz Adams[3], Saravanan Arunachalam[3]

[1]U.S. EPA Office of Research and Development, Research Triangle Park, NC | [2]U.S. EPA Office of Air Quality Planning and Standards, Research Triangle Park, NC | [3]Institute for the Environment, The University of North Carolina at Chapel Hill, NC

United States Environmental Protection Agency

Celebrating 25 Years of Modeling Excellence — CMAQ

Fahim Sidi | sidi.fahim@epa.gov

## Advantages of Data Storage on the Cloud

- CMAS hosts a Data Warehouse for the larger user community containing popular model inputs and outputs
- Some of these data are on Google Drive and are being transitioned to the Amazon Open Data Program, part of Amazon Sustainability Data Initiative
- Advantages of Data Storage on the Cloud:
  - No egress fees to download/upload and host data as part of Amazon's Open Data Program
  - Fast transfer speeds (upload/download)
  - No limit or throttle for obtaining data
  - Ability to run CMAQ directly on the cloud since inputs are already on the cloud
  - Ability to share data widely and create reproducible workflows for running CMAQ on the cloud

Accessing AWS Data for Cloud Computing

- Download Data from AWS Registry (https://registry.opendata.aws/cmas-data-warehouse/) using AWS command line interface
- Tutorials have been developed to run CMAQ on AWS and Microsoft Azure which includes instructions to access data directly on the cloud:
  - AWS: https://cyclecloud-cmaq.readthedocs.io/en/latest/
  - Microsoft Azure: https://cyclecloud-cmaq.readthedocs.io/en/latest/

See talk & workshop by Liz Adams during Cloud Computing session to learn more!

U.S. Environmental Protection Agency
Office of Research and Development

CMAQ Input Data Available on AWS

CMAQ WWLLNs Lightning Data

- CMAQ-ready gridded hourly observed lightning flash density data for the 12 km CONUS domain from the World Wide Lightning Location Network (WWLLN) with scaling (WWLLNs)
- Used by CMAQ to estimate lightning emissions inline; see chapter 6 of CMAQ Users Guide for additional details on how to configure CMAQ to use this dataset
- Temporal Coverage: Jan 1, 2016 – Dec 31, 2022
- Link to Data: https://cmas-wwlln-lightning.s3.amazonaws.com/index.html

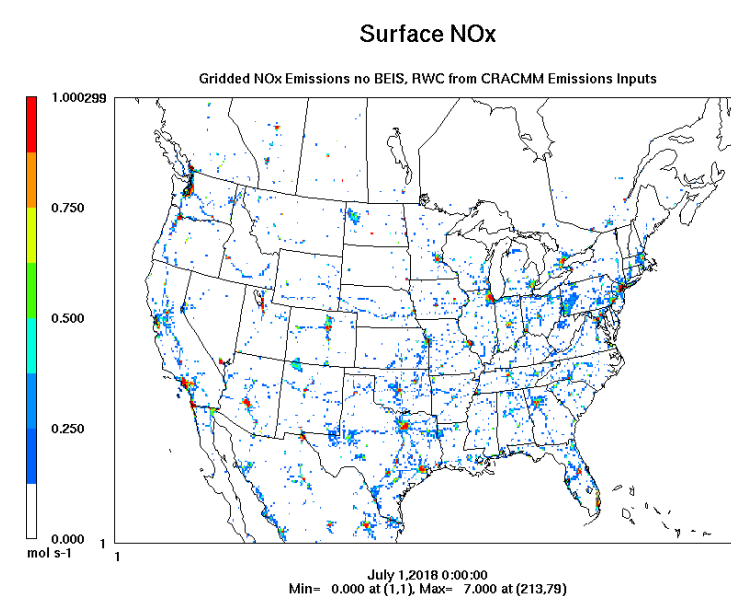Annual Total Lightning Flashes in 2016
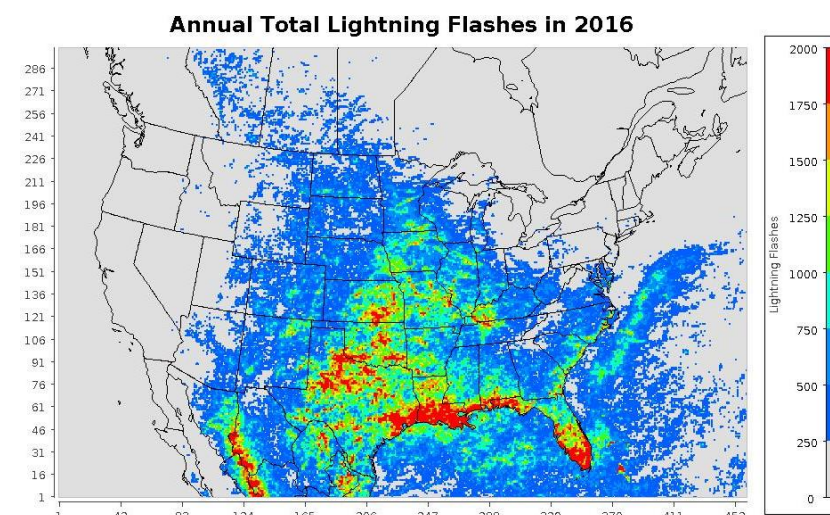
EPA's Air Quality TimE Series (EQUATES) Project Data

- CMAQ-ready gridded hourly emissions, meteorology and air quality modeling inputs for the 12 km CONUS domain & 108 North Hemispheric Domain
- Additional Documentation: www.epa.gov/cmaq/equates
- Temporal Coverage: Jan 1, 2017 – Dec 31, 2017
- Link to data: https://cmas-equates.s3.amazonaws.com/index.html


EQUATES all_onroad 2017: NOX
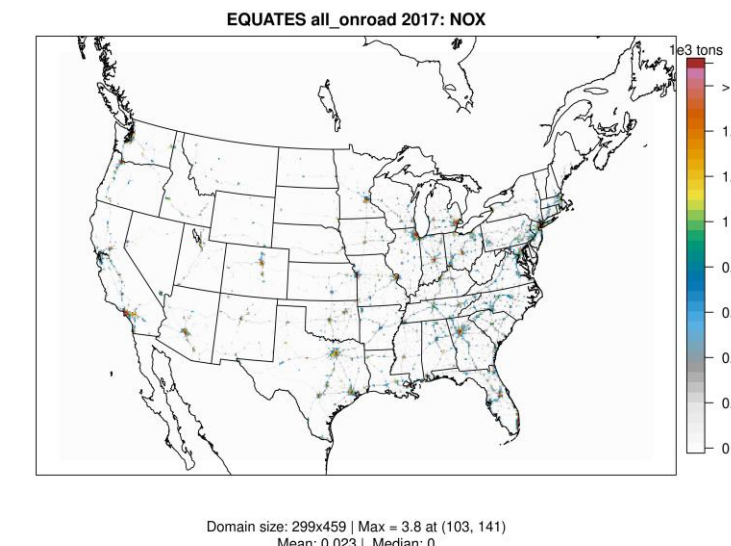
CMAQ 2018 Modeling Platform

- CMAQ-ready gridded hourly emissions, meteorology and air quality modeling inputs for the 12 km CONUS domain
- Inputs available to run a CMAQ simulation with CB6R5 or the Community Regional Atmospheric Chemistry Multiphase Mechanism (CRACMM) version 1.0 mechanism
- Includes inputs to an inline version of Model of Emissions of Gases and Aerosols from Nature (MEGAN) in CMAQ to estimate BVOC and SOIL NO emissions inline
- Temporal Coverage: Jan 1, 2018 – Dec 31, 2018
- Link to data: https://cmas-cmaq-modeling-platform-2018.s3.amazonaws.com/index.html

Surface NOx
Gridded NOx Emissions on BEIS, RWC from CRACMM Emissions Inputs
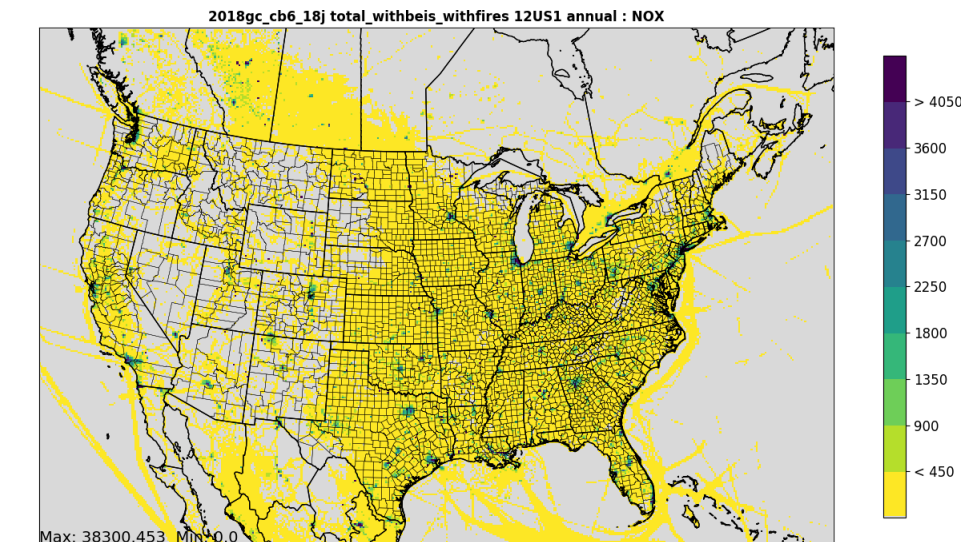
CMAQ Benchmark Data

- CMAQ-ready gridded hourly emissions, meteorology and air quality modeling inputs for the 12 km Northeast Domain
- Benchmark dataset used to verify successful installation of CMAQv5.4
- Temporal Coverage: July 1, 2018 – July 2, 2018
- Link to data: https://cmas-cmaq.s3.amazonaws.com/index.html

Surface Gridded NOx Emissions no BEIS, RWC
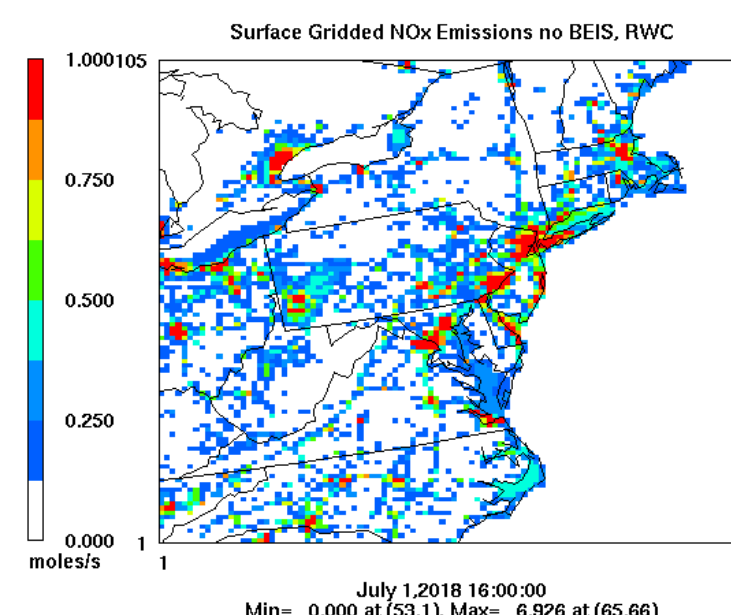
SMOKE Emissions Platform Data Available on AWS

National Emissions Inventory (NEI) 2019 Modeling Platform

- CMAQ-ready gridded hourly emissions from SMOKE for the 12 km CONUS domain
- 2019 emissions based on 2017 NEI released in 2020 including sector specific updates to represent emissions in 2019
- Temporal Coverage: Dec 22, 2018 – Dec 31, 2019
- Supporting Documentation: https://www.epa.gov/air-emissions-modeling/2019-emissions-modeling-platform
- Link to data: https://2019platform.s3.amazonaws.com/index.html

2019gc_cb6_19j total_withbeis 12US1 annual : NOX

NEI 2016v3 Modeling Platform

- CMAQ & CAMx-ready gridded hourly emissions from SMOKE for the CONUS domain at various resolutions
- 2016 emissions version 3 (2016v3), an update to 2016 version 2 which incorporates updates made in response to comments by stakeholders, improved methods, some corrections, and refinements to projection factors
- Temporal Coverage: 2016, 2023, 2026
- Supporting Documentation: https://www.epa.gov/air-emissions-modeling/2016v3-platformLink to data
- Link to data: https://2016v3platform.s3.amazonaws.com/index.html

AMET Observation Data Available on AWS

- Observational data formatted for the Atmospheric Model Evaluation Tool (AMET) available from various networks including AQS, CASTNET, IMPROVE, MDN, NADP, SEARCH, CSN, NAPS, ESRL, TOAR
- Temporal Coverage: 2000 – 2019
- Link to data: https://cmas-amet.s3.amazonaws.com/index.html

Datasets coming to AWS Soon

- MPAS-CMAQ 120 km uniform global mesh inputs for 2014-2018 CB6R5M & 2017-2018 CRACMM
- MPAS-CMAQ 92-25 km variable resolution mesh inputs for 2016 CB6R5M
- 2020 Emissions Modeling Platform based on 2020 NEI