

NASA Air Quality Forecasts and the MERRA-2 Reanalysis

Sarah Strode, Melanie Follette-Cook, Pawan Gupta, Carl Malings

NASA Air Quality Remote Sensing Training for EPA, March 21-23, 2023

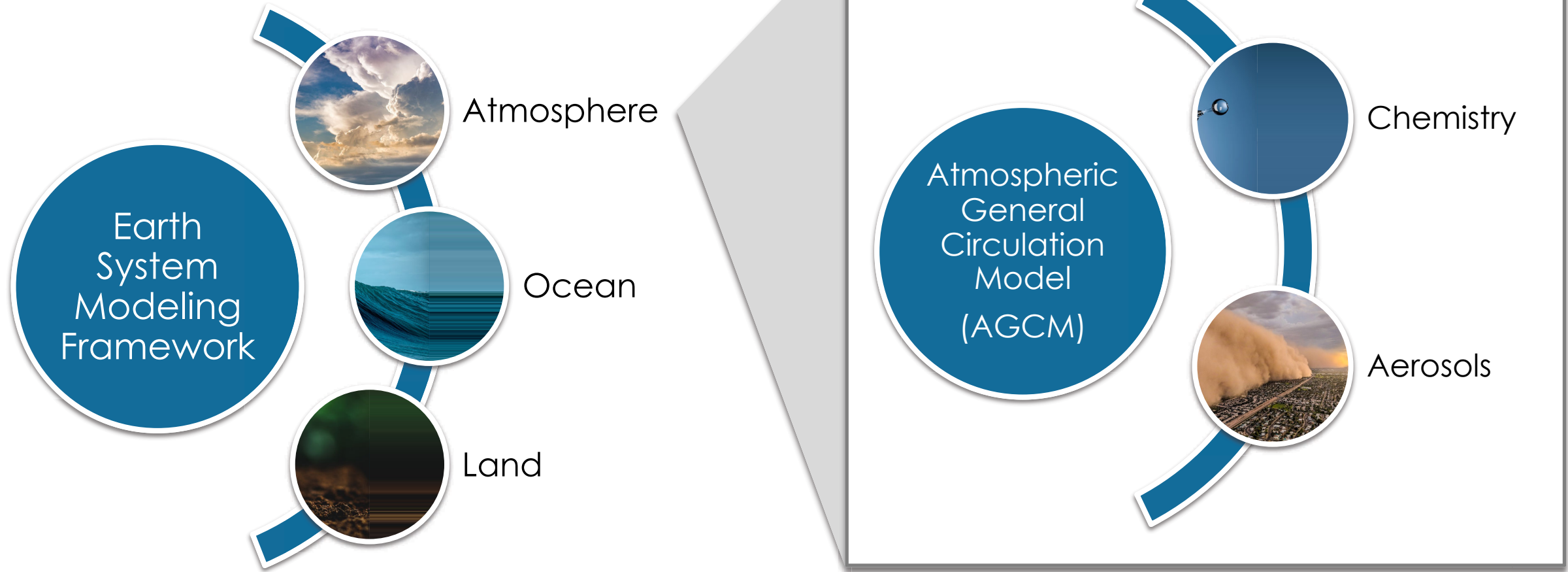
Learning Objectives

- Identify the different Air Quality (AQ) relevant model outputs available from Goddard Earth Observing System (GEOS) Earth System Model
- Understand the difference between analysis, reanalysis, and forecasting
- Understand the different ways satellite observations are used for forecasting, reanalysis, and evaluation
- Discover how to subset and visualize reanalysis and forecast outputs

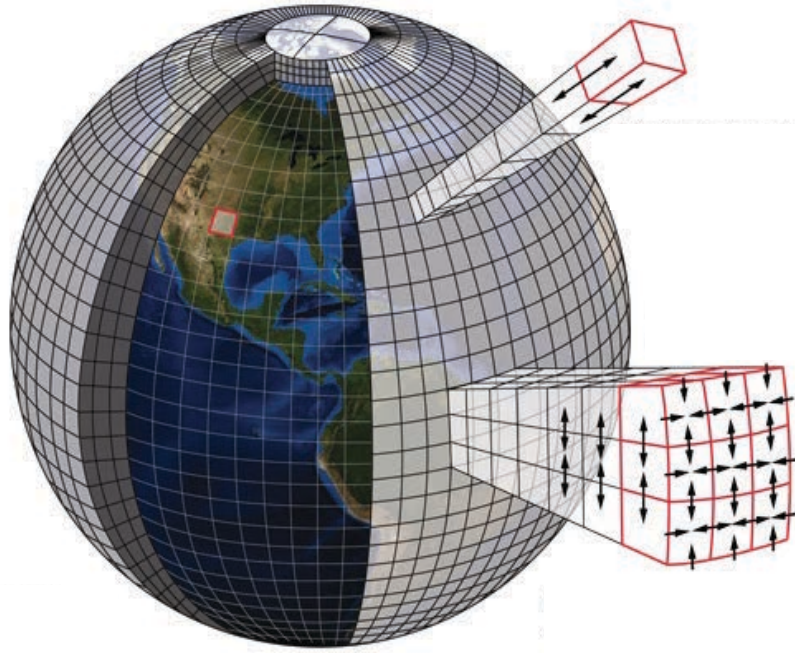


NASA GEOS Earth System Model

<https://gmao.gsfc.nasa.gov/>



Modeling the Atmosphere



Three-dimensional (3D) atmospheric chemistry models divide the atmosphere into a set of 3D grid cells.



The GEOS AGCM is run on a *cubed sphere* grid.

- Ensures uniform spatial grid
- Better for scalability

Models solve equations for physics, transport, and chemistry within each grid cell.

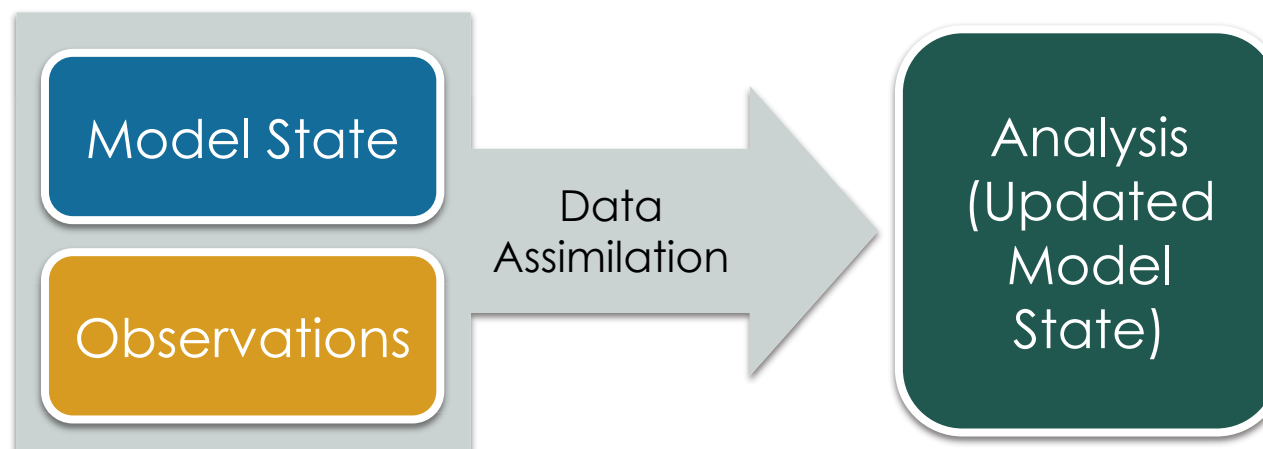


Forecast, Analysis, Reanalysis, Data Assimilation

What are the differences between these?

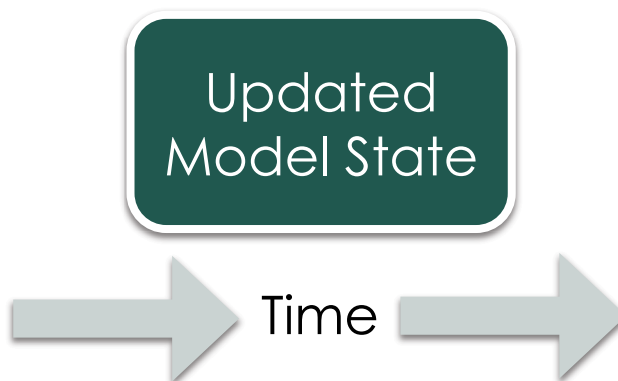
Data assimilation

describes the process of assimilating, or incorporating, observations into a model state to produce the best estimate of the atmosphere, land, and ocean conditions.



An **analysis** is the blend of the model and observations.

A **reanalysis** blends observations with model simulations of the past using a single model version.



A **forecast** is a model simulation run forward in time to predict a future state.



GEOS Forecast and Reanalysis Products

GEOS Forward Processing
(GEOS FP)
NRT Analysis and Forecast

GEOS-Composition
Forecast
(GEOS-CF)
NRT Forecast

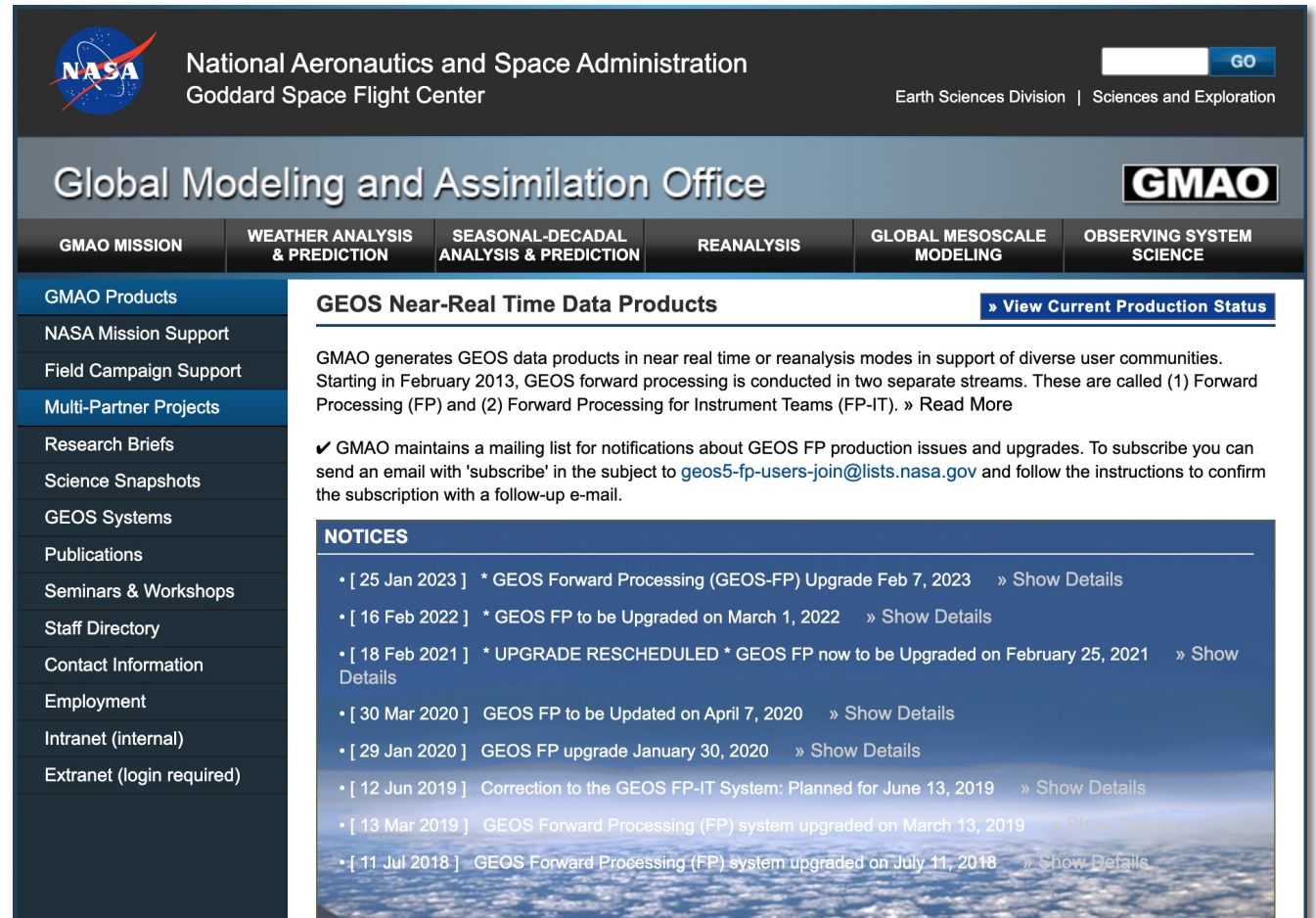
Modern-Era Retrospective
analysis for Research and
Applications, Version 2
(MERRA-2) Reanalysis



Global Modeling and Assimilation Office (GMAO)

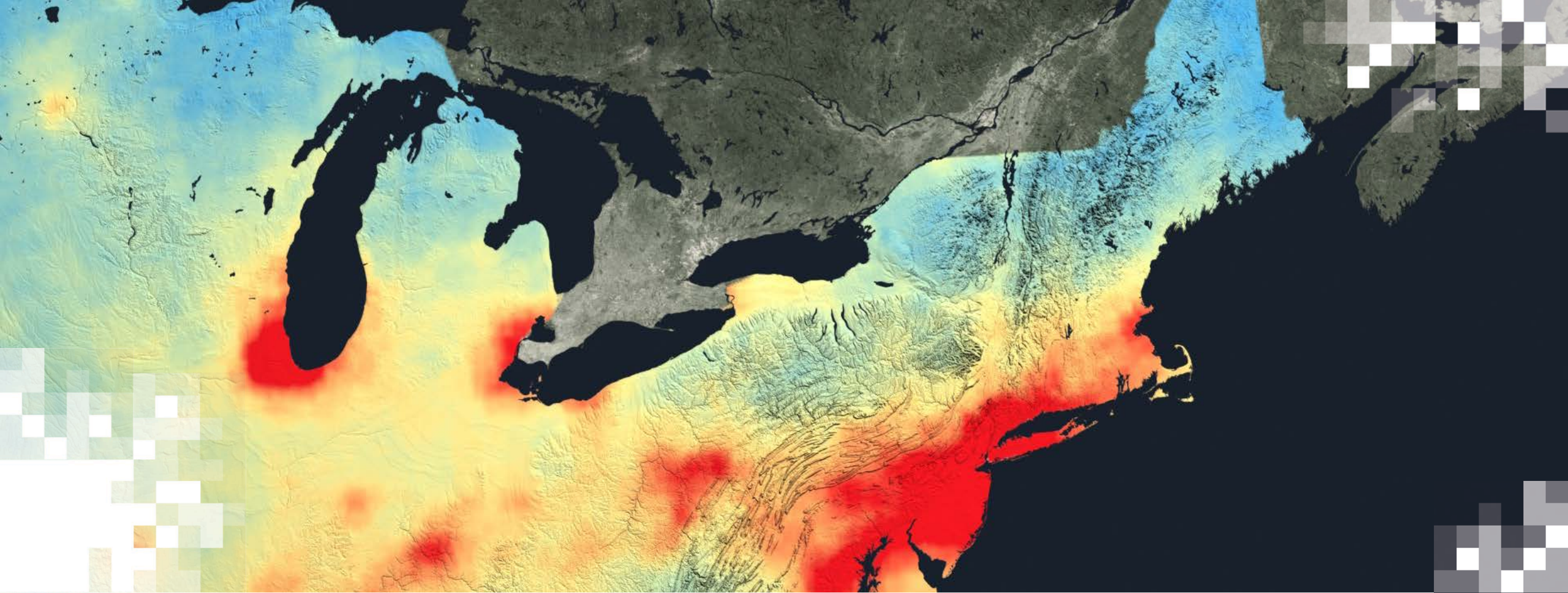
https://gmao.gsfc.nasa.gov/GMAO_products/NRT_products.php

- The GEOS forecasts are a dynamic system.
- For the most current information about GEOS NRT products, visit the GMAO website.
- The most current version of file specification documents are covered here:
https://gmao.gsfc.nasa.gov/pubs/office_notes.php



The screenshot displays the NASA Goddard Space Flight Center website for the Global Modeling and Assimilation Office (GMAO). The header includes the NASA logo, the text 'National Aeronautics and Space Administration Goddard Space Flight Center', a search bar with a 'GO' button, and the text 'Earth Sciences Division | Sciences and Exploration'. Below the header is the 'Global Modeling and Assimilation Office' title with the 'GMAO' logo. A navigation menu contains links for 'GMAO MISSION', 'WEATHER ANALYSIS & PREDICTION', 'SEASONAL-DECADAL ANALYSIS & PREDICTION', 'REANALYSIS', 'GLOBAL MESOSCALE MODELING', and 'OBSERVING SYSTEM SCIENCE'. A left sidebar lists various resources such as 'GMAO Products', 'NASA Mission Support', 'Field Campaign Support', 'Multi-Partner Projects', 'Research Briefs', 'Science Snapshots', 'GEOS Systems', 'Publications', 'Seminars & Workshops', 'Staff Directory', 'Contact Information', 'Employment', 'Intranet (internal)', and 'Extranet (login required)'. The main content area features a 'GEOS Near-Real Time Data Products' section with a 'View Current Production Status' link. Below this is a paragraph explaining that GMAO generates GEOS data products in near real time or reanalysis modes. A checkmark icon indicates that GMAO maintains a mailing list for notifications about GEOS FP production issues and upgrades, with instructions on how to subscribe. A 'NOTICES' section follows, listing several updates with dates and links to 'Show Details', such as 'GEOS Forward Processing (GEOS-FP) Upgrade Feb 7, 2023' and 'GEOS FP to be Upgraded on March 1, 2022'.





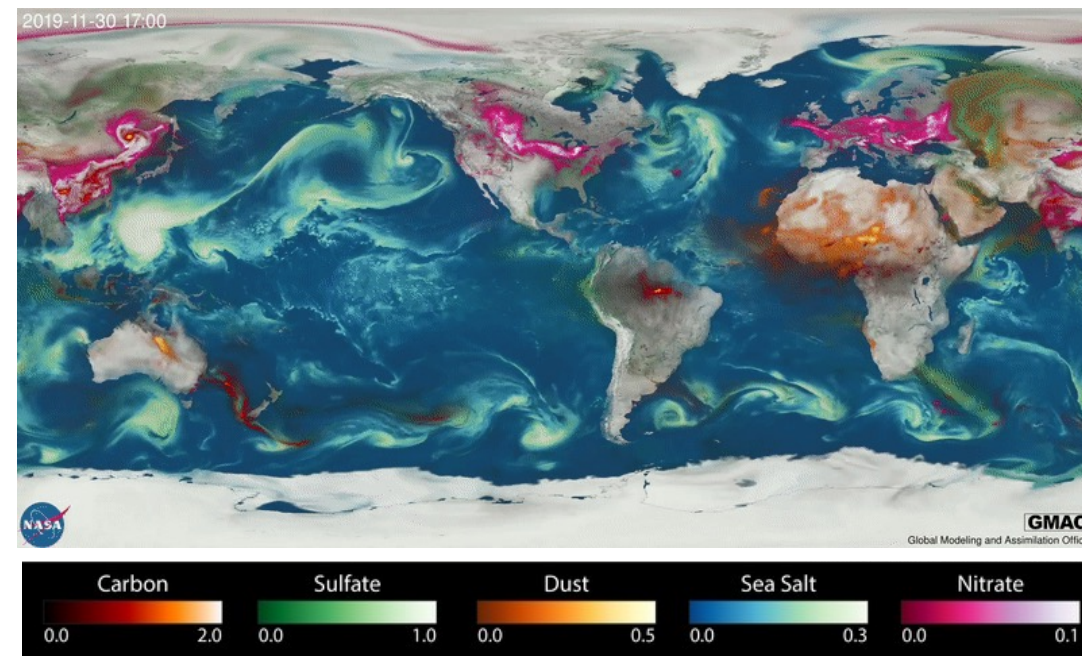
GEOS FP

GEOS FP

https://gmao.gsfc.nasa.gov/weather_prediction/

- GEOS FP analyses and forecasts support NASA field campaigns and provide a testbed for assimilation and forecast development.
- Publicly available
- Includes weather, aerosols, and carbon monoxide (CO) on the same spatial scale
- State of the science forecast system – model physics or observing system updated every 6-12 months
 - Not suitable for trend analysis
- Meteorology Used to Drive Chemistry Models:
 - GEOSChem, Whole Atmosphere Community Climate Model (WACCM)
- When using FP meteorology fields to drive another model, must ensure your simulation does not span an update
 - [GMAO NRT Product Page](https://gmao.gsfc.nasa.gov/nrt/) has updated details and dates

AOD (550 nm)



<https://svs.gsfc.nasa.gov/31100>

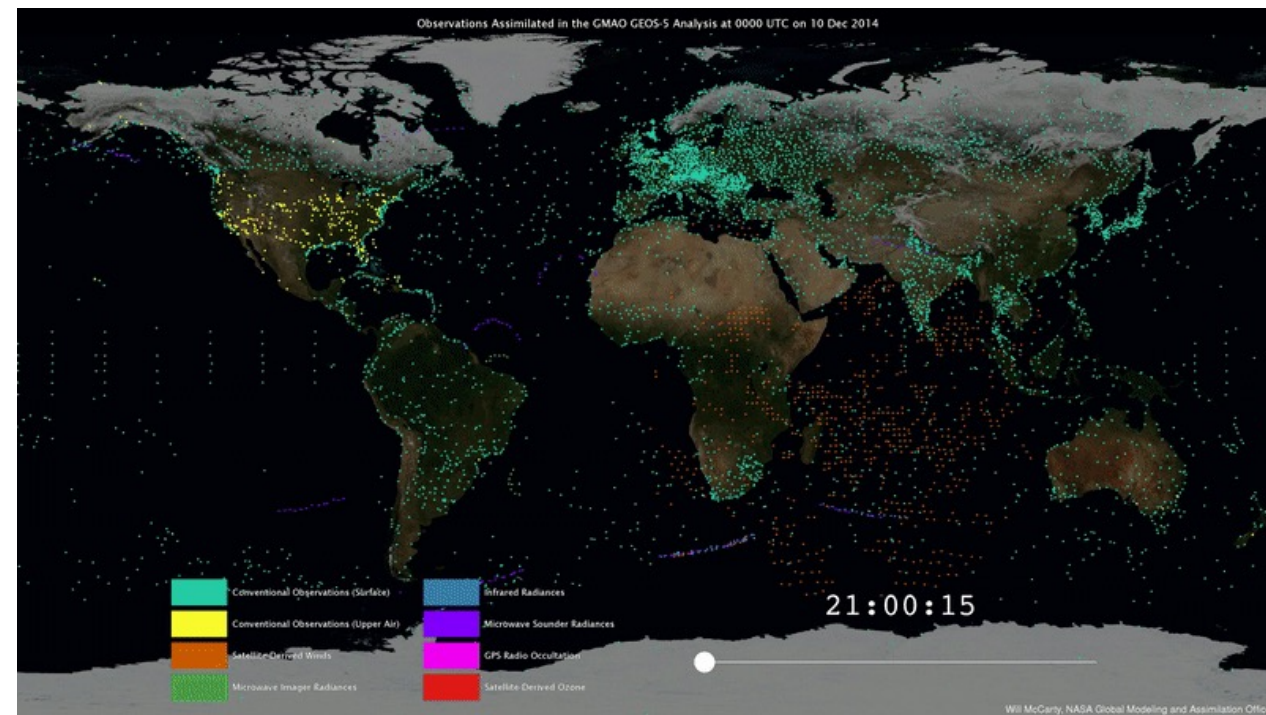
GEOS Output Quick Guide

	GEOS FP
Type	Analysis + Forecast
Domain	Global
Spatial Resolution	Simulation: ~12 km Output: ~25 km (0.25°x0.312°)
Temporal Resolution	2-D Data: Hourly 3-D Data: Every 3 h
Vertical Levels	72 (near surface-0.1 hPa)
Output Available	Analysis: 2014 – Present Forecast: ~21 days
Initialization	Daily 10-Day Forecast at 00Z Daily 5-Day Forecast at 12Z
Data Assimilation	Yes
File Specification Doc	https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf *



Data Assimilation in GEOS FP

- **Data assimilation** describes the process of assimilating, or incorporating, observations into a model state to produce the best estimate of the atmosphere, land, and ocean conditions.
- GEOS uses a Hybrid 4D-Ensemble Variational (Hyb-4D-EnVar) approach.
- Analyses are created every 6 hours using over 5 million observations.
- GEOS assimilates AOD from MODIS (Terra and Aqua).
- Important: AOD is the total extinction of all aerosol species, so when AOD is increased or decreased as a result of the assimilation, assumptions are made about how that is distributed among the species.

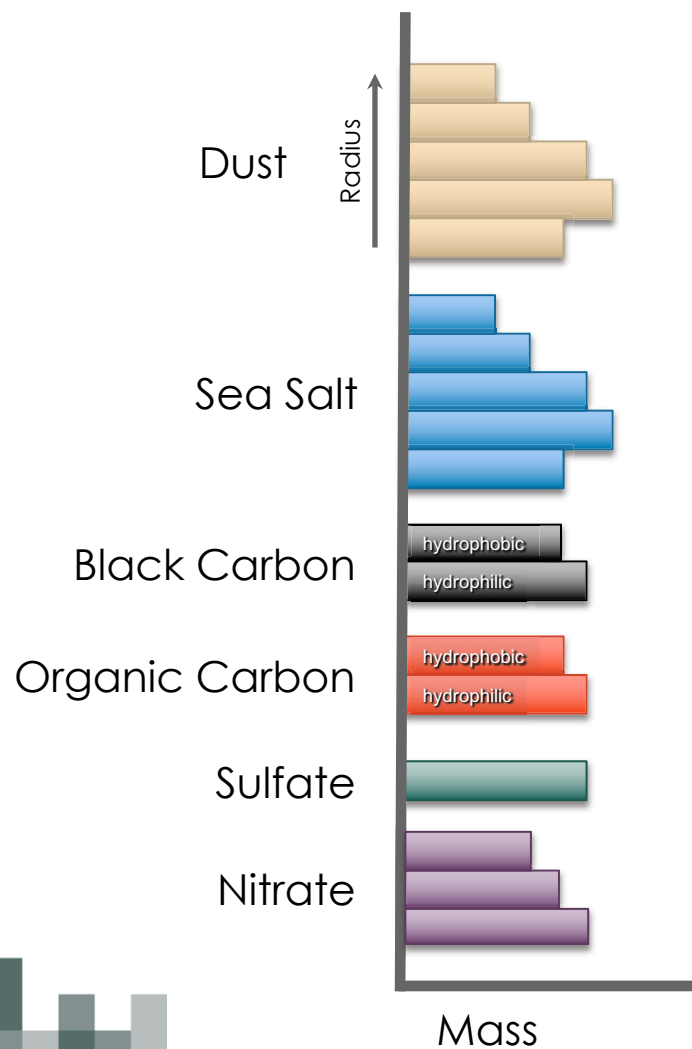


<https://svs.gsfc.nasa.gov/30590>



GOCART in GEOS

- Goddard Chemistry, Aerosol, Radiation and Transport Model (GOCART, Chin et al. 2002, Colarco et al. 2010)
- Sources and sinks for 6 non-interactive species
- Radiatively active



Wind and Topographic Sources, 5 Mass Bins

Wind-Driven Source, 5 Mass Bins

Anthropogenic and Wildfire Sources, Mass Hydrophobic & Hydrophilic

Anthropogenic, Biogenic, and Fire Sources, Mass Hydrophobic and Hydrophilic

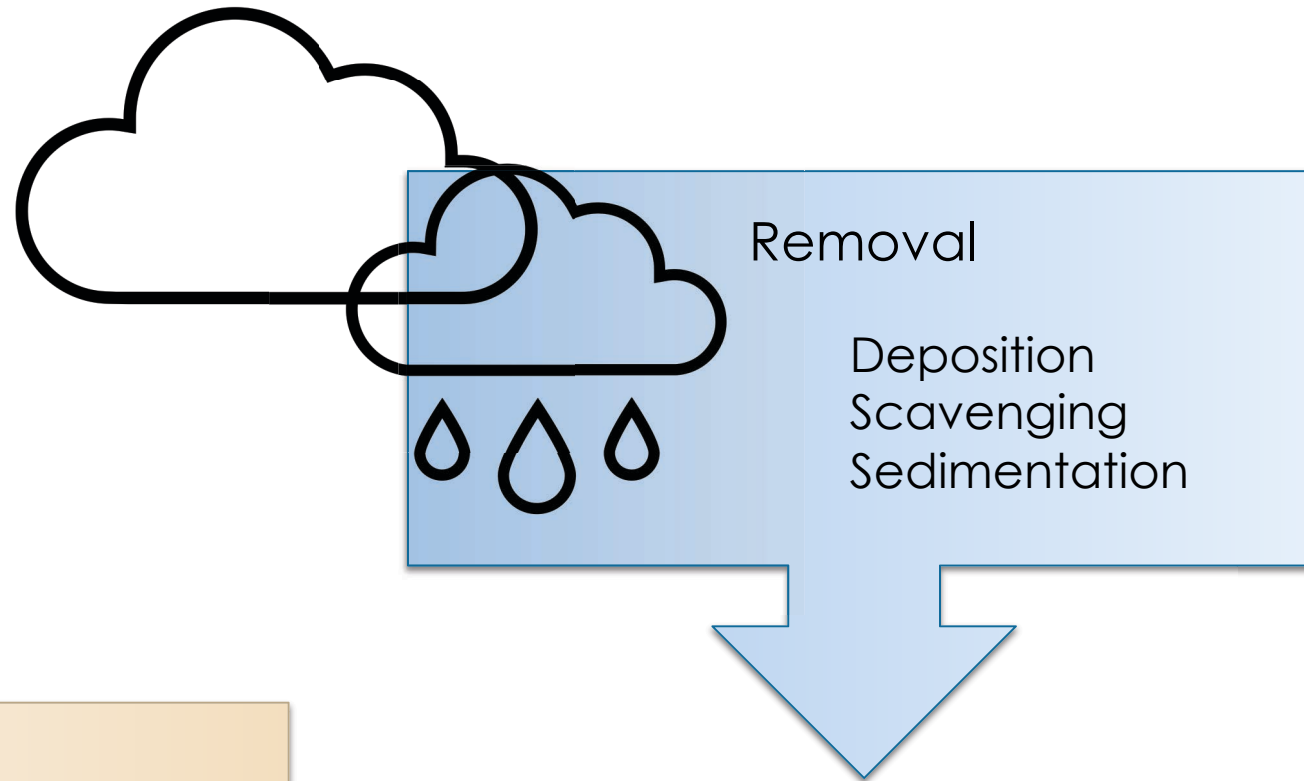
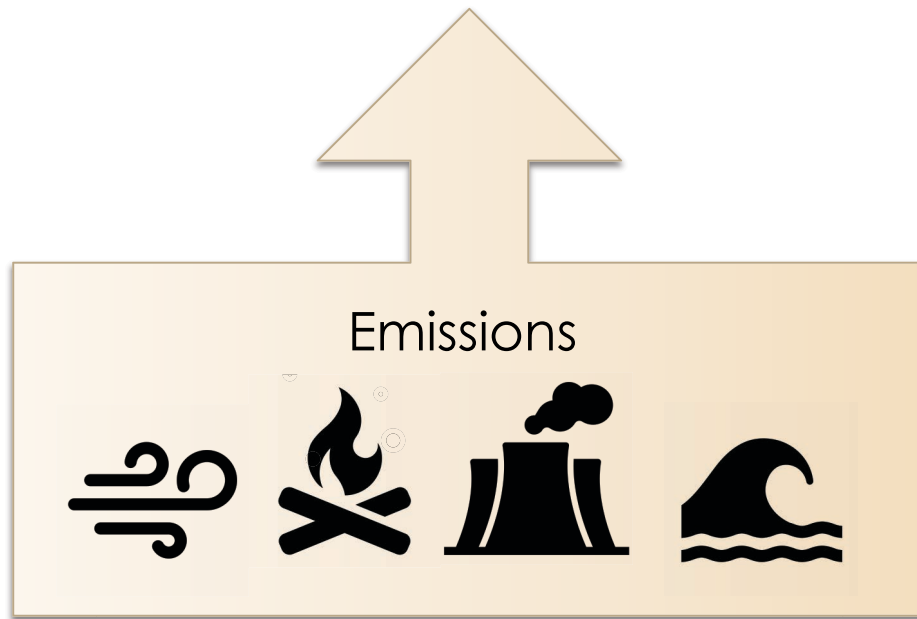
Anthropogenic, Wildfire, and Volcanic

Anthropogenic and Wildfire Sources



Aerosol Processes

GOCART simulates the emission, chemistry, and removal of aerosols and CO within each grid cell.



Aerosol Emissions: Dust, Biomass Burning, Anthropogenic Sources, Volcanic Emissions



Example GEOS FP File Name

GEOS.fp.fcst.tavg3_2d_aer_Nx.20210901_00+20210902_1330.V01.nc4



GEOS FP File Collections

- GEOS FP output is organized into file *collections* that contain related variables.
- These have the form:

Frequency_Dimensions_Group_HV

Frequency

Frequency or Averaging Interval

- const = time-independent
- inst = instantaneous
- tavg = time-average

Dimensions

Dimensions of Variables

- 2d = only 2d fields
- 3d = can have 2d and 3d

Group

Three letter abbreviation for the type of variables

- Also used in the short name
- Ex. aer = Aerosol fields
- See documentation for full list

HV

Horizontal and Vertical Grid

- H = typically N, for nominal grid
- V = x, horizontal only
- V = p, pressure level
- V = v, model level
- V = e, model layer edges



GEOS FP Output File Names

- Each GEOS FP file has the form:

GEOS.fp.mode.collection.time.file_ver.nc4

Mode

- asm = Assimilation
- fcst = Forecast

Collection

See previous slide

File_ver

File Version (usually V01)

Time

Date and Time of Data File

- For Assimilation:
yyymmdd_hhmm
- For Forecast:
yyymmdd_hhmm+yyymmdd_hhmm

Assimilation Cycle

Forecast Time

All GEOS FP output files are in NetCDF-4 format.



Example GEOS FP File Name

GEOS.fp.fcst.tavg3_2d_aer_Nx.20210901_00+20210902_1330.V01.nc4

- fp – forward processing
- fcst – forecast product
- tavg3_2d_aer_Nx : 2D time-averaged aerosol diagnostics
- 20210901_00+20210902_1330 : Forecast initialized at 2021-09-01 00 Z. The valid time for the data in this file is 2021-09-02 1330 Z, which represents the center point of a 3-hour time-averaging period between 1200 and 1500 Z.



AQ-Relevant Collections and Variables

<https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf> *

Collection Name	Description
tavg3_2d_aer_Nx	2D time-averaged aerosol diagnostics <i>Optical properties (Extinction AOT, Scattering AOT, Angstrom parameter) Surface concentration (kg/m³), Column Density (kg/m²)</i>
tavg3_2d_adg_Nx	2D time-averaged aerosol diagnostics (extended) <i>Emissions and removal processes (deposition, sedimentation, and scavenging)</i>
tavg3_2d_chm_Nx	2D time-averaged chemistry diagnostics <i>Surface CO, column CO, emissions, chemical loss, chemical production</i>
inst3_3d_aer_Nv	3D instantaneous aerosol diagnostics <i>Mass mixing ratios (kg/kg) of aerosol species in each size bin</i>
inst3_3d_chm_Nv	3D instantaneous chemistry diagnostics <i>CO molar mixing ratio (mol/mol)</i>

To calculate PM_{2.5}, use the formula:

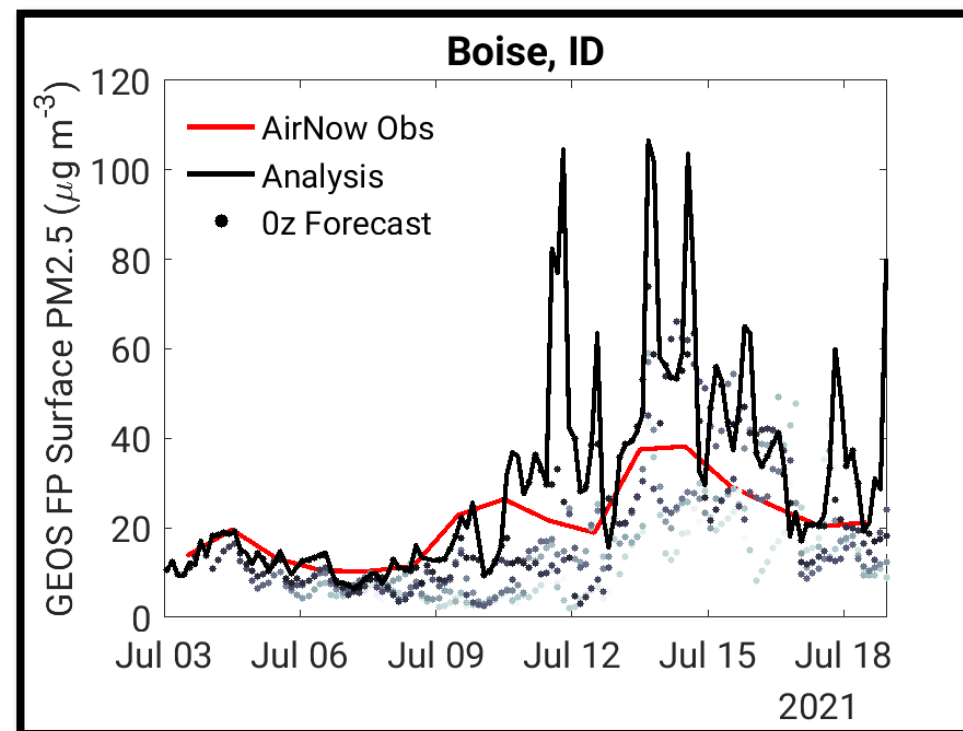
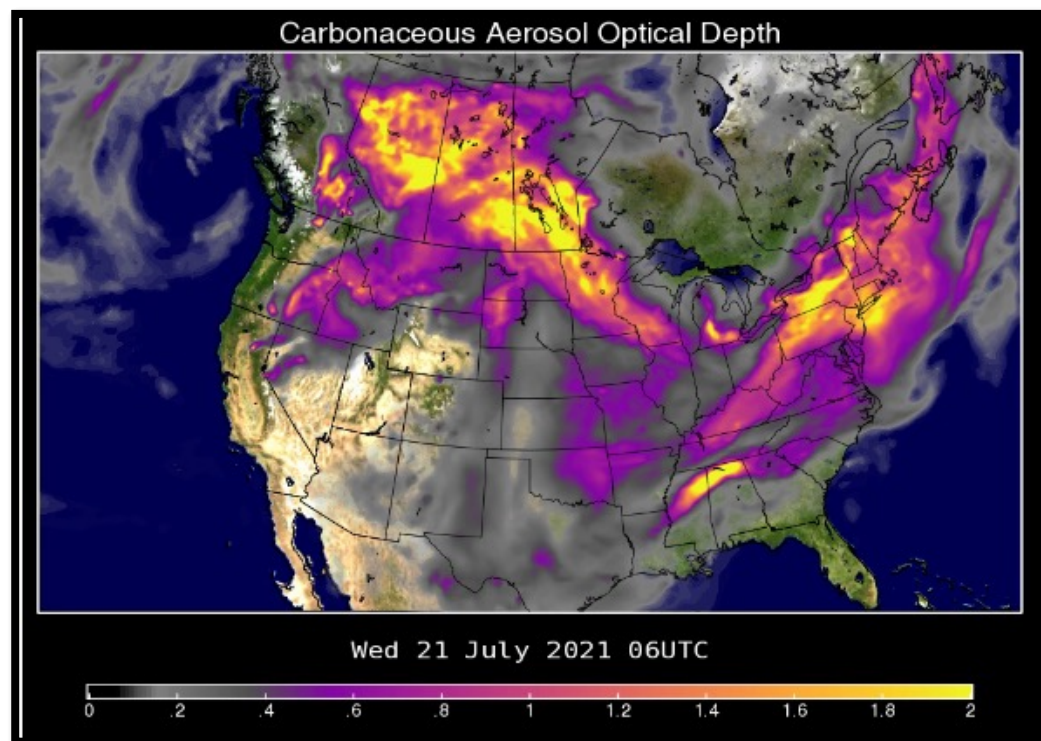
$$PM_{2.5} = [DUSMASS25] + [SSSMASS25] + [BCSMASS] + [OCSSMASS] + 1.375 \times [SO4SSMASS] + 1.29 \times [NISSMASS25]$$

These variables are contained in the *tavg3_2d_aer_Nx* collection.

In the near future, GEOS FP output will include a PM_{2.5} variable.



Case Study: 2021 Biomass Burning Season



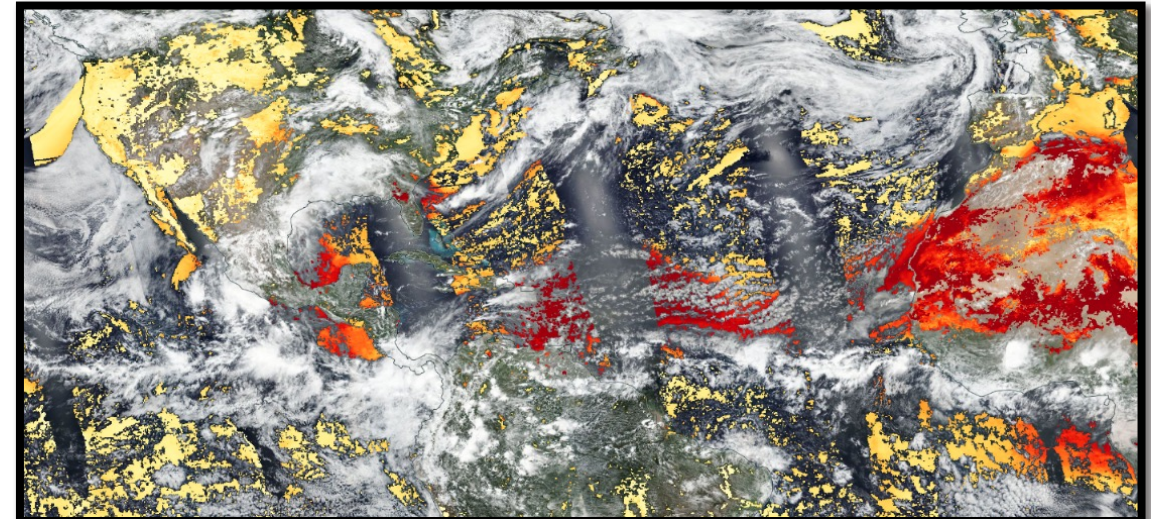
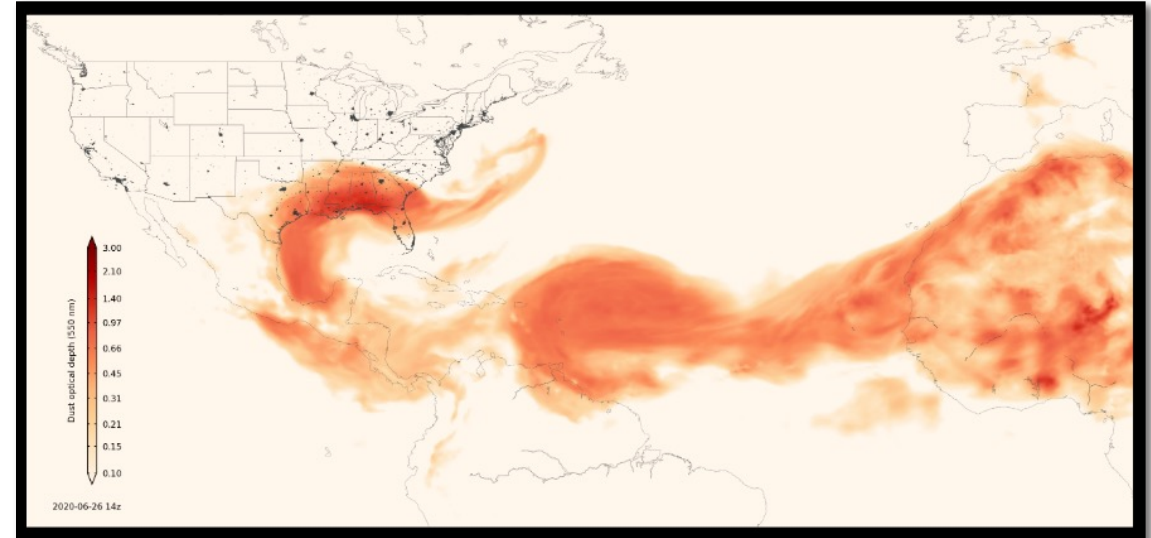
https://gmao.gsfc.nasa.gov/research/science_snapshots/2021/na_biomass_burning_2021.php

At the beginning of July 2021, GEOS FP was able to accurately simulate current conditions and forecast the amount of PM_{2.5} as a result from the smoke from the wildfires in the region. However, beginning on July 10th, GEOS FP overestimated PM_{2.5} at the surface near Boise, Idaho.



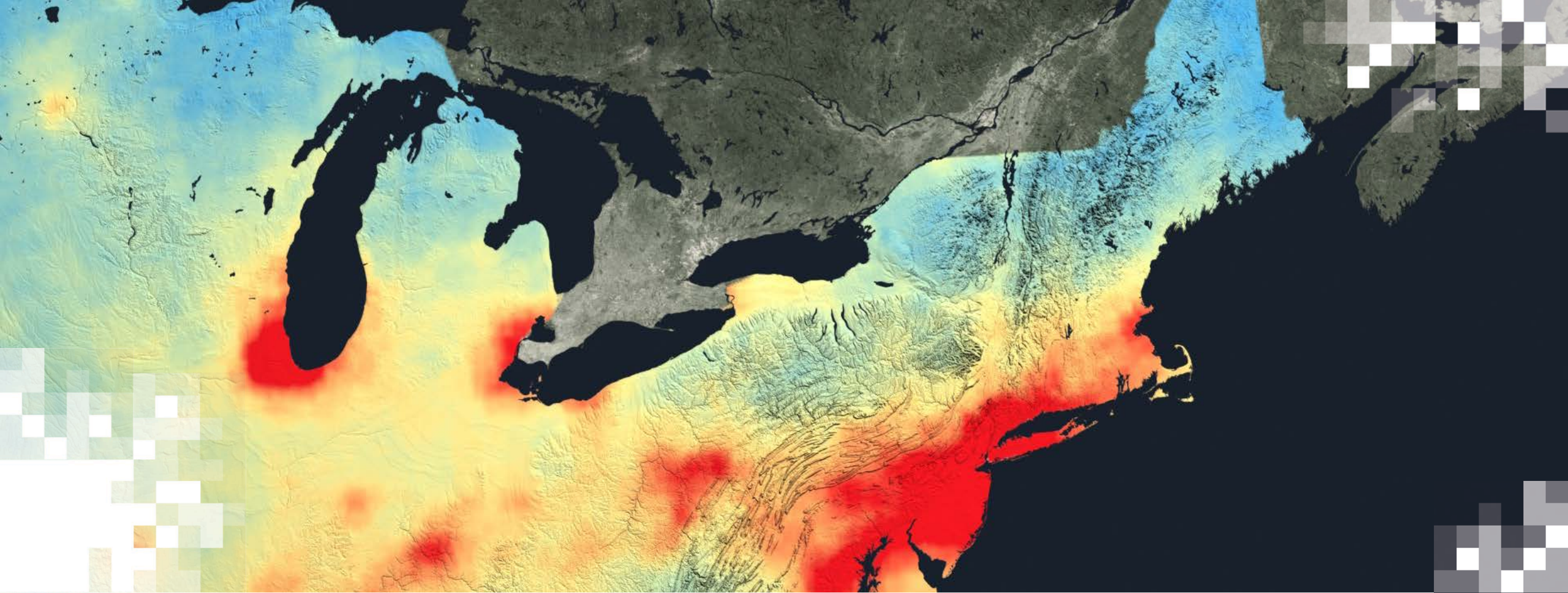
Case Study: June 2020 Dust Storm

- During June 2020, dust was transported across the Atlantic Ocean.
 - Highest Caribbean aerosol loading in the last 20-year period
 - Surface concentrations almost three times higher than the 24-hour EPA standard
- GEOS FP analysis (top) shows the Saharan dust on June 26, 2020, at 14z
- VIIRS AOD (Deep Blue, Land and Ocean) and true color imagery shows the satellite observed plume on the same day (bottom, [Worldview](#)).



https://gmao.gsfc.nasa.gov/research/science_snapshots/2020/Saharan_dust_2020.php



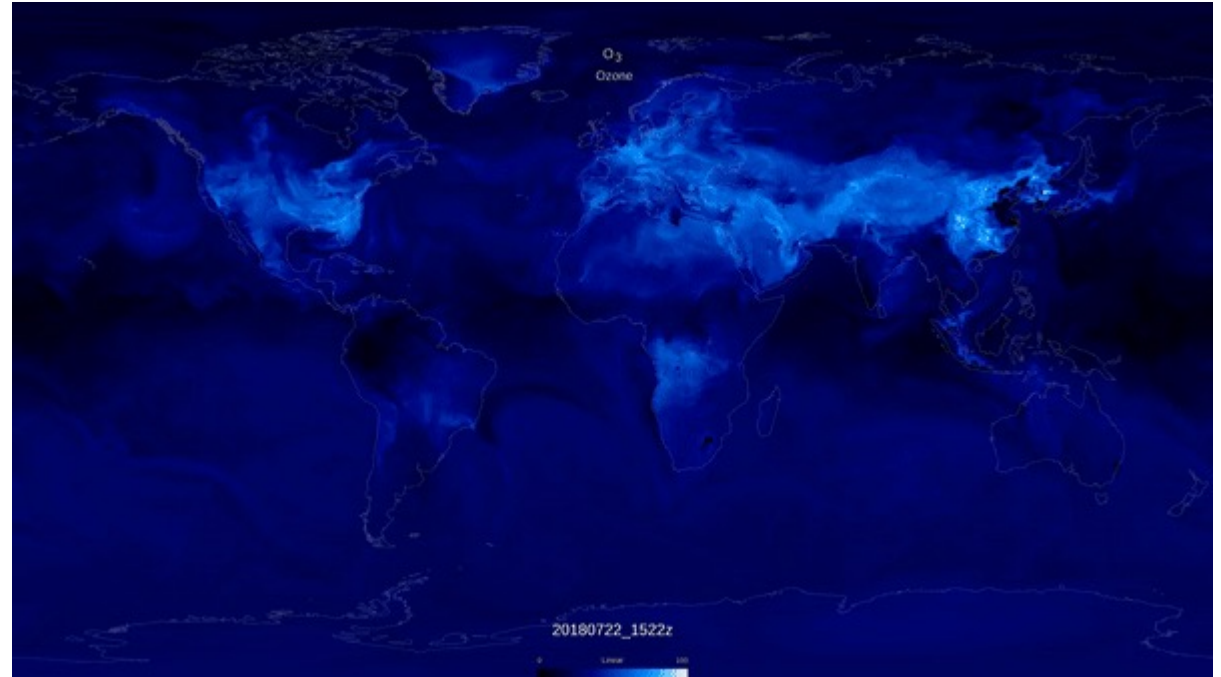


GEOS-CF

NASA Composition Forecasts (GEOS-CF)

https://gmao.gsfc.nasa.gov/weather_prediction/GEOS-CF/

- The GEOS-Composition Forecast (CF) system forecasts trace gas and aerosol fields using constrained meteorology from GEOS and the GEOS-Chem chemical mechanism.
- Publicly available
- GEOS-Chem is a community-developed global 3-D model of atmospheric chemistry.
 - 250 chemical species
 - 725 chemical reactions
- Questions about GEOS-CF can be sent to geos-cf@lists.nasa.gov



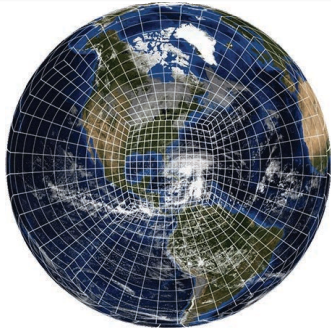
<https://svs.gsfc.nasa.gov/4754>



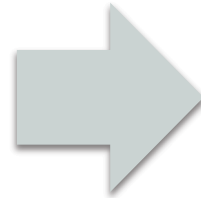
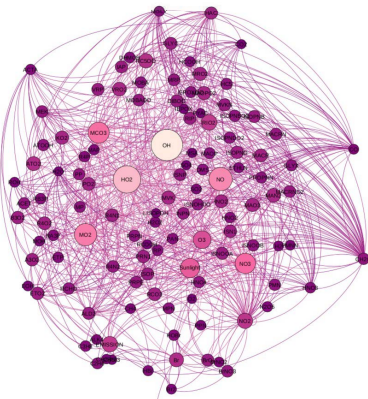
NASA Composition Forecasts (GEOS-CF)



GEOS
Meteorology



GEOS-Chem
Chemistry



GEOS-CF

- [GEOS-Chem](#): Global chemistry transport model driven by GEOS meteorology
- 1-day simulation of the previous day using the analysis from FP-IT
 - Uses a **replay** technique to force the meteorology towards the FP-IT analysis
 - FP-IT is a ‘frozen’ version of FP used for satellite retrievals, similar to the version used to make MERRA-2.
- 5-day forecast
- Two Aerosol Schemes:
 - GOCART – Radiatively coupled to AGCM
 - GEOS-Chem – No feedbacks to model physics
- Full description in [Keller et al., 2021](#)



GEOS Output Quick Guide

	GEOS FP	GEOS-CF
Type	Analysis + Forecast	Replay + Forecast
Domain	Global	Global
Spatial Resolution	Simulation: ~12 km Output: ~25 km (0.25°x0.312°)	~25 km (0.25°x0.312°)
Temporal Resolution	2-D Data: Hourly 3-D Data: Every 3 h	15 min, Hourly
Vertical Levels	72 (near surface-0.1 hPa)	72 (near surface-0.1 hPa)
Output available	Analysis: 2014 – Present Forecast: ~20 days	Replay: 2018 – Present Forecast: 2019 – Present (aqc collection) ~14 days (all collections)
Initialization	Daily 10-Day Forecast at 00Z Daily 5-Day Forecast at 12Z	Daily 5-Day Forecast at 12Z
Data Assimilation	Yes	No
File Specification Doc	https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf *	https://gmao.gsfc.nasa.gov/pubs/docs/Knowland1204.pdf *



GEOS-CF File Collections

- GEOS-CF output is organized into file *collections* that contain related variables.
- These have the form:

Group_Time_#Frequency_H_V

Group

Three letter abbreviation for the type of variables

- aqc = AQ Relevant
- chm = Chemistry
- htf = High-Temporal Frequency
- met = Meteorology
- xgc = Extra Chem Fields

Time

- inst = Instantaneous
- tavg = Time-Average

Frequency

Frequency or Averaging Time Interval

- mn = Minute
- hr = Hour

H

Horizontal Grid, hIxJ

- h = Horizontal Domain
- g = (Global) or
- r = (Regional)
- IxJ = Horizontal Resolution (# lon points, # lat points)

V

Vertical Resolution, vL

- v = x, 2d fields
- v = p, pressure levels
- v = v, model levels
- L = # vertical levels



GEOS-CF File Names

- Each GEOS-CF file has the form:

GEOS-CF.**version**.**mode**.**collection**.timestamp.nc4

Version

File Version (usually V01)

Mode

- rpl = Replay
- fcst = Forecast

Collection

See previous slide

Timestamp

Date and Time of Data File

- For assimilation:
yyymmdd_hhz
- For forecast:
yyymmdd_hhz+yyymmdd_hhmmz

Assimilation Cycle

Forecast Time

All GEOS-CF output files are in NetCDF-4 format.



Example GEOS-CF File Name

GEOS-CF.v01.fcst.chm_tavg_1hr_g1440x721_v1.20190309_12z+20190314_0730z.nc4

- GEOS-CF.v01.fcst - GEOS-CF forecast filename
- chem_tavg_1hr_g1440x721_v1 – Chemical species collection (“chm”), 1-hour time-averaged (“tavg_1hr”) at the global $\sim 0.25^\circ$ horizontal resolution (“g1440x721”) for single model layer data (“v1”)
- 20190309_12z+20190314_0730z – Forecast initialized at 20190309 at 12 Z. The valid time for the data in this file is 20190314 at 0730 Z, which represents the center point of a one-hour time-averaging period between 0700 and 0800 Z.



AQ-Relevant Collections and Variables

<https://gmao.gsfc.nasa.gov/pubs/docs/Knowland1204.pdf> *

Collection Name	Description
htf_inst_15mn_g1440x721_x1	High Temporal Frequency Chemistry and Meteorology <i>Surface CO, NO₂, O₃, SO₂, PM_{2.5} (GCC & GOCART), and meteorology (RH, T, P, etc)</i>
aqc_tavg_1hr_g1440x721_v1	Air Quality Concentrations <i>One-hour time-averaged surface CO, NO₂, O₃, PM_{2.5}(GCC), SO₂</i>
chm_tavg_1hr_g1440x721_v1	Chemistry Fields <i>One-hour time-averaged surface mixing ratios of many chemical species and speciated PM_{2.5} (GCC)</i>
xgc_tavg_1hr_g1440x721_x1	Extra GEOS-Chem Fields <i>One-hour time-averaged AOD, column quantities, and removal processes (deposition)</i>
chm_inst_1hr_g1440x721_p23	Chemistry Fields <i>3D (23 pressure levels, 1000 to 10 hPa) instantaneous CO, NO₂, O₃, PM_{2.5}(GCC, speciated), SO₂</i>

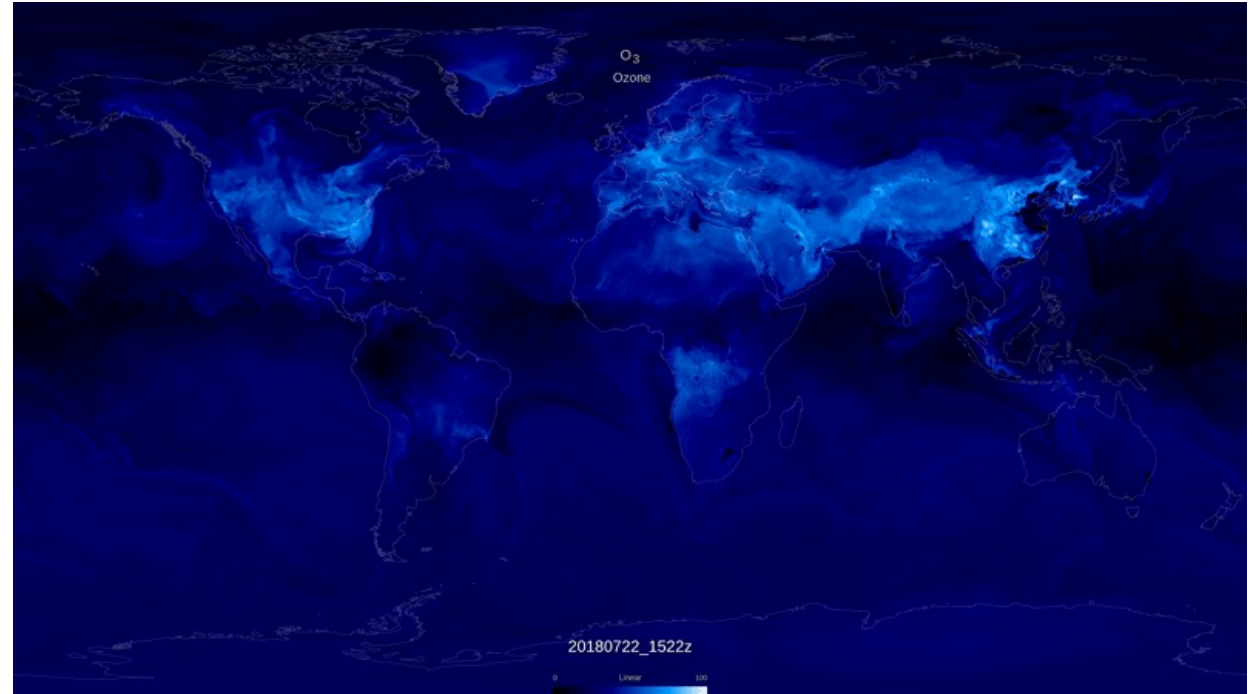
There are Two PM_{2.5} Variables:

- PM25_RH35_GCC: PM_{2.5} at 35% RH from GEOS-Chem (GCC)
- PM25_RH35_GOCART: PM_{2.5} at 35% RH from GOCART (only available in htf)



Case Study: High Resolution Simulation of Ozone

- Surface O₃ from GEOS-CF during the summer of 2018
- O₃ is a pollutant produced and destroyed through interactions of various chemical species such as nitrogen oxides (NO₂, NO) and volatile organic compounds (VOCs).
- Forecasted concentrations of pollutants like O₃, NO₂, and PM_{2.5} can be combined to calculate air quality indices.

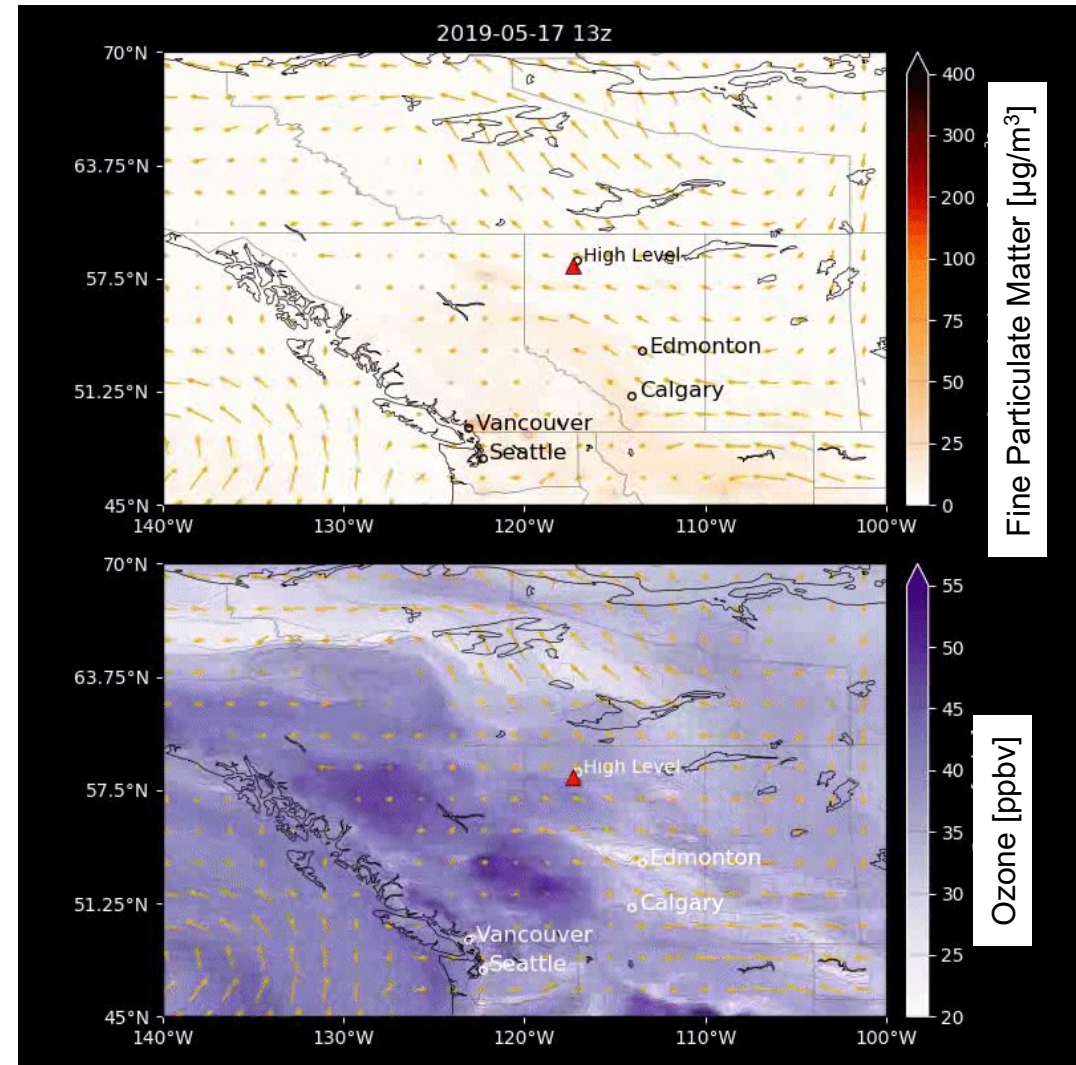


<https://svs.gsfc.nasa.gov/4764>



Case Study: O₃ and PM_{2.5} Forecasts During 2019 Canadian Wildfires

- Smoke from fires led to poor air quality in the Northwest US.
- Within the smoke plumes, high concentrations of NO consume O₃, leading to lower O₃ levels (lighter colors) near the source.
- As the plume mixes with surrounding air, O₃ is produced, leading to increased concentrations near the plume edge.
- O₃ produced in wildfire plumes can be comparable to urban pollution levels.



 Fire Detection



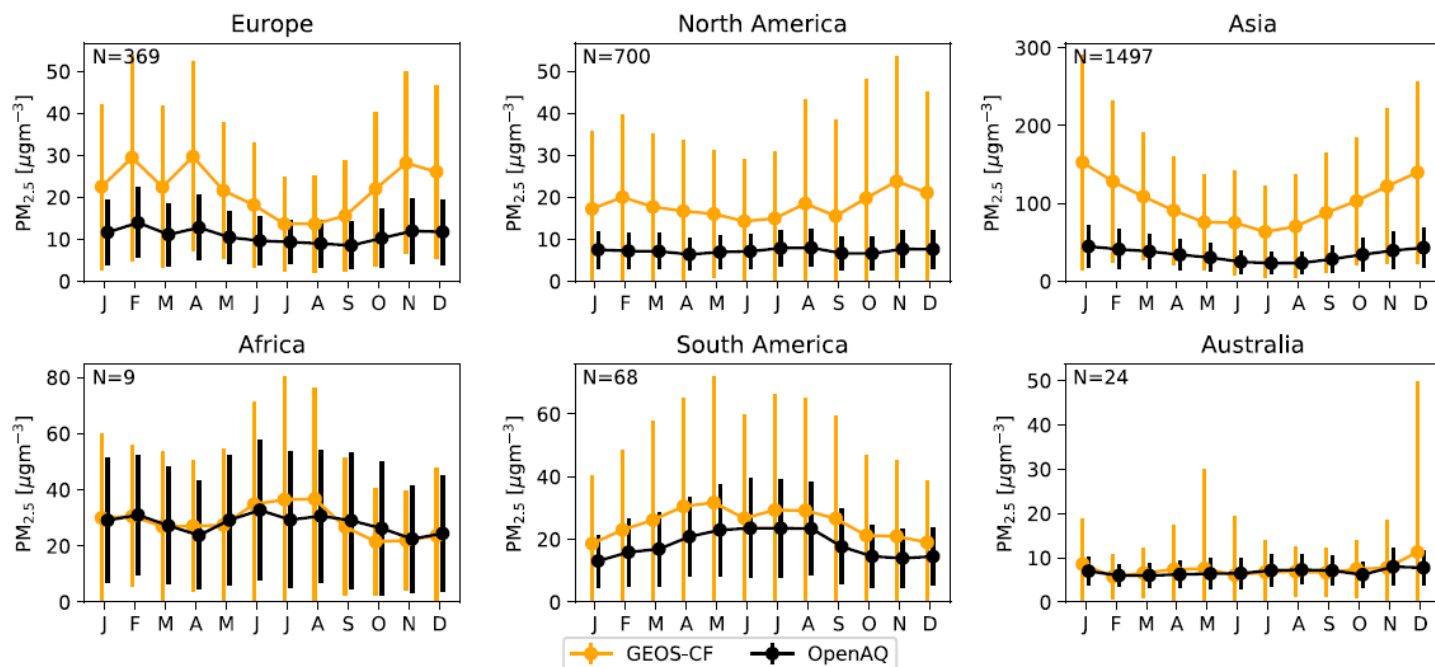
GEOS-CF Evaluation

- Keller et al. (2021) includes a detailed evaluation of GEOS-CF
- Variety of observations used for evaluation:
 - Sondes
 - Ground-Based In-Situ
 - Ground-Based Remote Sensing (Aeronet AOD)
 - Satellite (OMI NO₂, MOPITT CO Column, MODIS AOD)
- Evaluation of both replay and forecast skill included
- Aerosol evaluation of GEOS-Chem aerosols



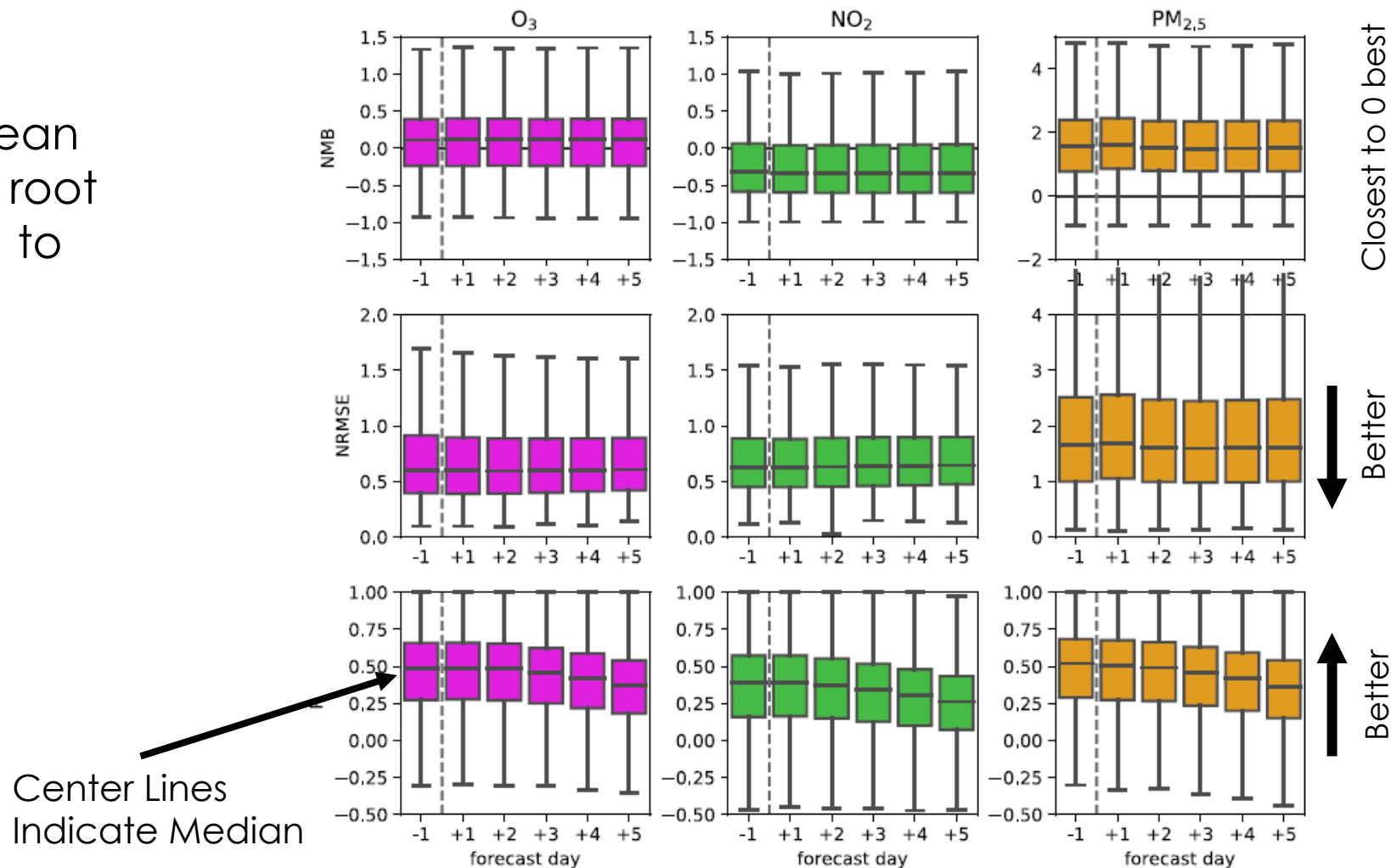
GEOS-CF Replay Evaluation – PM_{2.5}

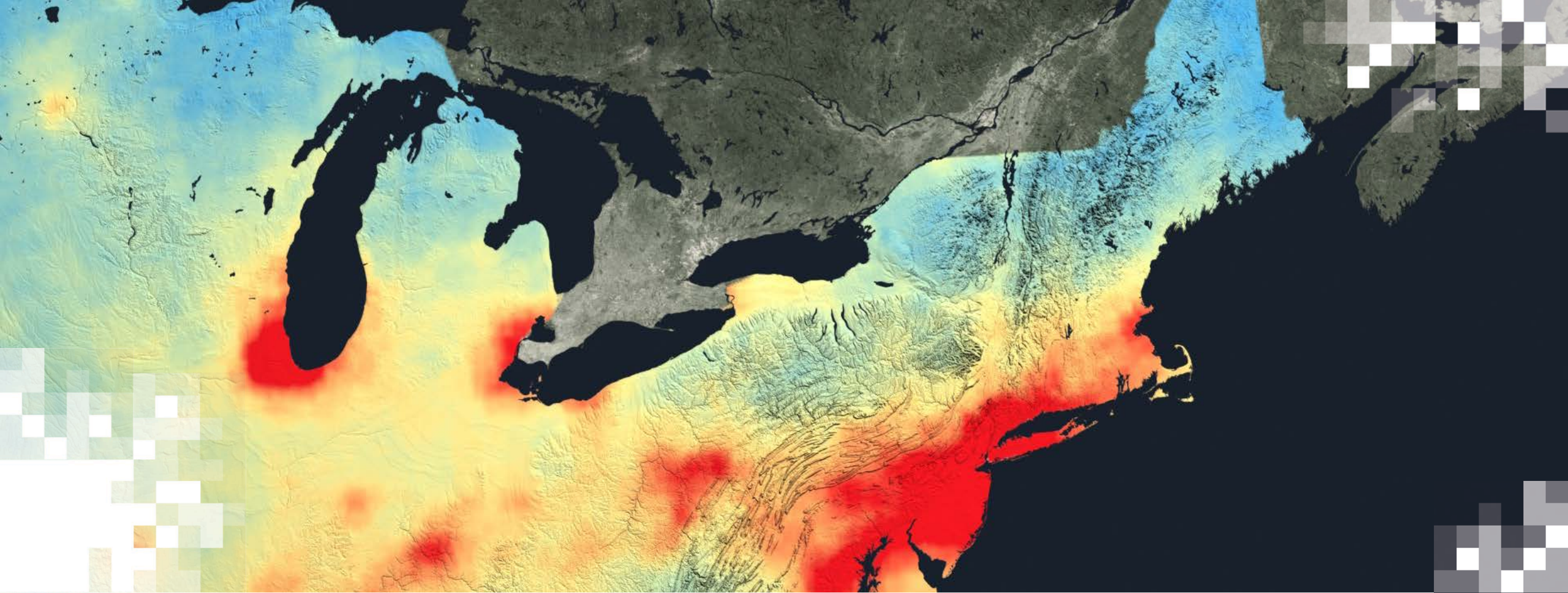
- Comparison with ~2600 surface PM_{2.5} observations from [OpenAQ](#)
- Similar results as comparison with AOD
- Even though GEOS-CF has high resolution for a global forecast model, there are still biases when comparing with station-level observations, especially in urban areas.



GEOS-CF Forecast Evaluation

- 5-day forecasts have comparable normalized mean bias (NMB) and normalized root mean square error (NRMSE) to the 1-day replay.





MERRA-2

What is reanalysis, and why do we do it?

What:

- A consistent reprocessing of Earth system observations using a modern, unchanging data assimilation system
- Relies on models to interpret, relate, and combine different observations from multiple sources
- Successful reanalysis **requires** a good forecast model combined with bias-corrected and quality-controlled observations

Why:

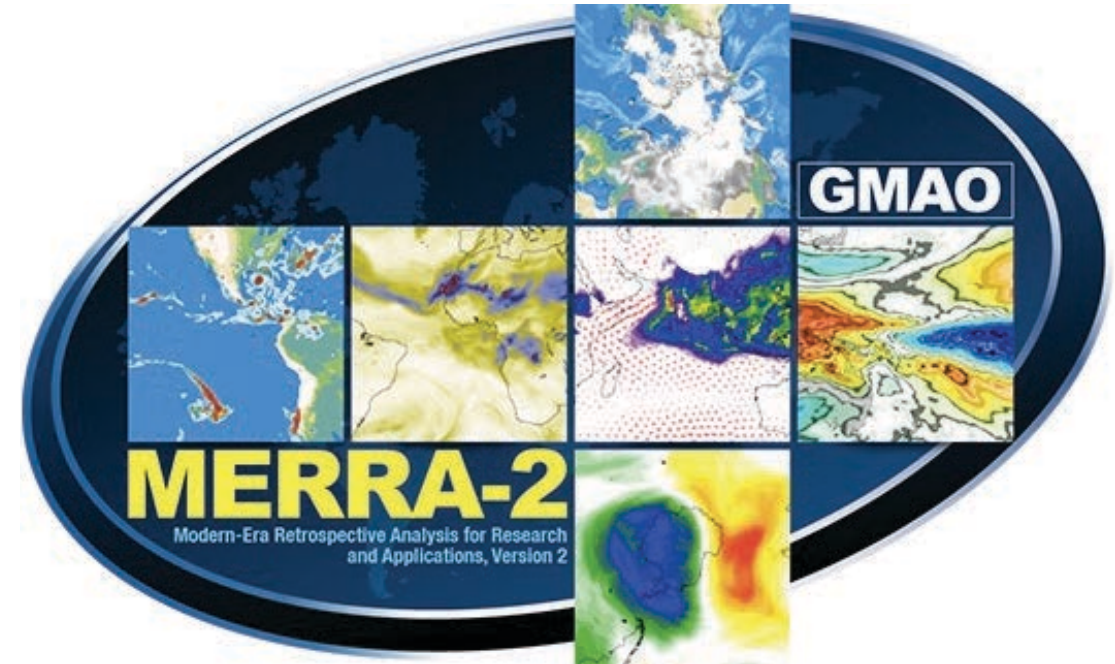
- Produces multi-decadal, gridded datasets that estimate a large variety of Earth system variables, including ones that are not directly observed
- Has become fundamental to research and education in the Earth sciences



MERRA-2 Reanalysis

<https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/>

- The **M**odern-**E**ra **R**etrospective analysis for **R**esearch and **A**pplications version **2** (MERRA-2) provides data beginning in 1980 and runs a few weeks behind real-time.
- Long-term, model-based analyses of multiple datasets using a fixed assimilation system
- Includes meteorology, stratospheric ozone, and aerosols at the spatial resolution of a $0.5^\circ \times 0.66^\circ$ (~50 km) grid



Source: <https://gmao.gsfc.nasa.gov/reanalysis/>

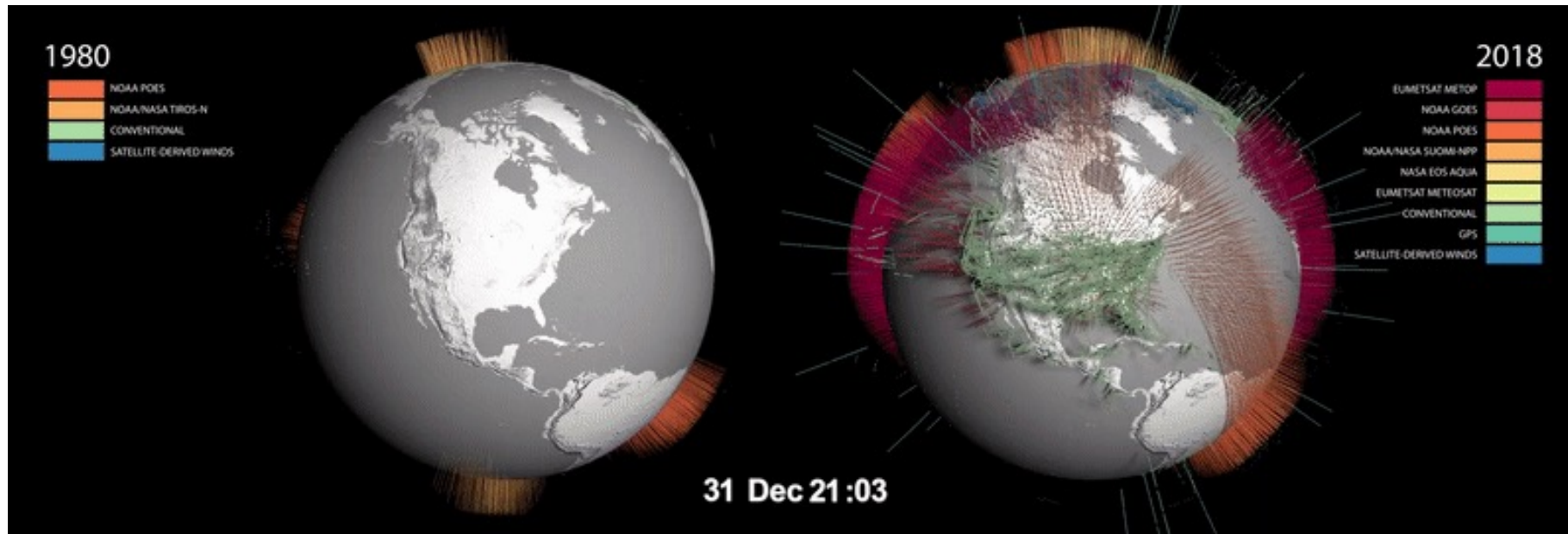


GEOS Output Quick Guide

	GEOS FP	GEOS-CF	MERRA-2
Type	Analysis + Forecast	Replay + Forecast	Reanalysis
Domain	Global	Global	Global
Spatial Resolution	Simulation: ~12 km Output: ~25 km (0.25°x0.312°)	~25 km (0.25°x0.312°)	~50km (0.5°x0.625°)
Temporal Resolution	2-D Data: Hourly 3-D Data: Every 3 h	15 min, Hourly	Hourly, Daily, Monthly
Vertical Levels	72 (near surface-0.1 hPa)	72 (near surface-0.1 hPa)	72 (near surface-0.1 hPa)
Output available	Analysis: 2014 – Present Forecast: ~20 days	Replay: 2018 – Present Forecast: 2019 – Present (aac collection) ~14 days (all collections)	1980-Present
Initialization	Daily 10-Day Forecast at 00Z Daily 5-Day Forecast at 12Z	Daily 5-Day Forecast at 12Z	~1-2 months behind real time
Data Assimilation	Yes	No	Yes
File Specification Doc	https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf *	https://gmao.gsfc.nasa.gov/pubs/docs/Knowland1204.pdf *	https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf *



Observing System in MERRA-2



In 1980, there were few satellites providing observations. These satellites, with global surface and upper-air observations were the first observations used for the beginning of MERRA-2 in 1980. Every 6 hours, a median number of 175,000 observations were assimilated.

Today, our observing system has advanced significantly, and MERRA-2 assimilates a median number of 5 million observations every 6 hours. When using MERRA-2 products, one must take care to consider the changing observing system over time.



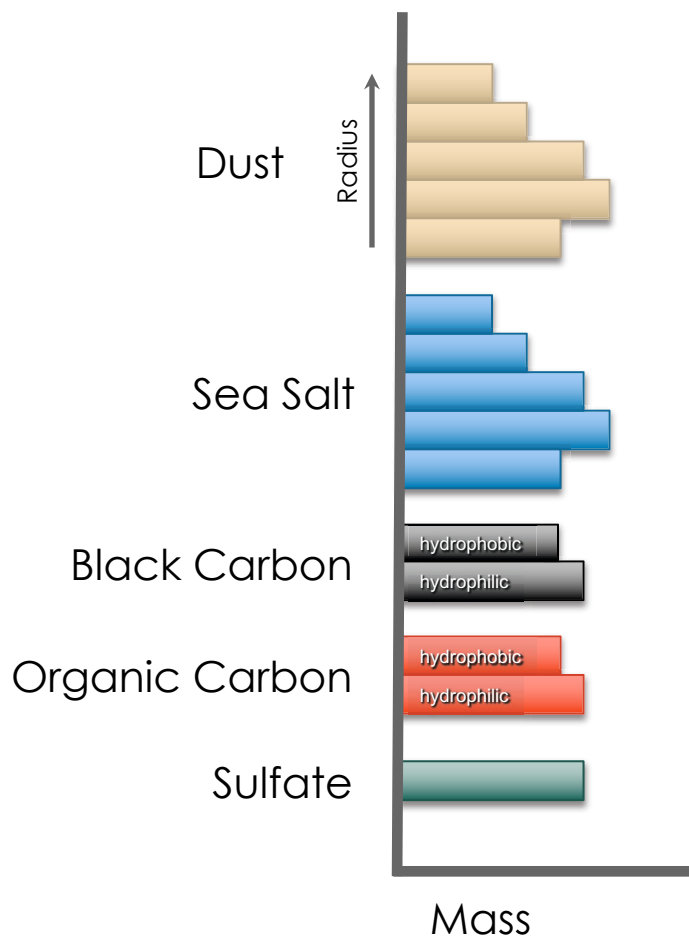
MERRA-2 Aerosol Observations and Evaluation

- Aerosol emissions and assimilation are described in detail in [Randles et al. 2017](#) and <https://gmao.gsfc.nasa.gov/pubs/docs/Randles887.pdf>.
- In MERRA-2, AOD at 550 nm is assimilated.
- Some Notes:
 - No information on vertical structure or composition
 - Daylight observations only
 - Subject to meteorological conditions (e.g., clouds) and viewing geometry (e.g., sun glint)
 - When there are no observations, MERRA-2 draws towards the GEOS/GOCART simulation
- In order to evaluate the performance of a reanalysis, it is important to compare the output with *independent* sources of data (i.e., those not used for assimilation).
- A detailed evaluation of MERRA-2 aerosols can be found in [Buchard et al. \(2017\)](#).
 - Optical properties, vertical distribution, and surface PM_{2.5}



GOCART in MERRA-2

- Goddard Chemistry, Aerosol, Radiation and Transport Model (GOCART, Chin et al. 2002, Colarco et al. 2010)
- Sources and sinks for 5 non-interactive species
- Radiatively active



Wind and Topographic Sources, 5 Mass Bins

Wind-Driven Source, 5 Mass Bins

Anthropogenic and Wildfire Sources, Mass Hydrophobic & Hydrophilic

Anthropogenic, Biogenic, and Fire Sources, Mass Hydrophobic and Hydrophilic

Anthropogenic and Wildfire Sources of SO_2 , Oxidation to SO_4 Mass

There are no nitrate aerosols in MERRA-2.



MERRA-2 File Collections

- MERRA-2 outputs are organized into file *collections* that contain related variables.
- These have the form:

Frequency_Dimensions_Group_HV

Frequency

Frequency or Averaging Interval

- const = Time-Independent
- inst = Instantaneous
- tavg = Time-Average
- Stat = Statistics

Can be 1, 3, 6-Hourly, Daily (D), Monthly (M), or a Monthly-Diurnal Mean (U)

Group

Three letter abbreviation for the type of variables

- These are also used in the short name
- Ex. aer = Aerosol Fields
- See documentation for full list

Dimensions

Dimensions of Variables

- 2d = Only 2d Fields
- 3d = Can Have 2d and 3d

HV

Horizontal and Vertical Grid

- H = Typically N, for Nominal Grid
- V = x, horizontal only
- V = p, pressure level
- V = v, model level
- V = e, model layer edges



MERRA-2 File Names

- Each MERRA-2 file has the form:

MERRA2_**SVv**.**collection**.timestamp.nc4

Stream and Version

File Version (usually 100, 200, 300, or 400)

Collection

See previous slide

All MERRA-2 output files are in NetCDF-4 format.

Timestamp

Date and Time of Data File

- For instantaneous or time-averaged files: `yyyymmdd`
- For monthly files: `yyyymm`

For collections with instantaneous or time-averaging frequency < 1 day, the daily file will contain all of the timesteps



Example MERRA-2 File Name

MERRA2_400.tavgM_2d_aer_Nx.202106.nc4

- MERRA2_400 – MERRA-2 file from fourth assimilation stream
- tavgM_2d_aer_Nx – 2D monthly time-averaged (“tavgM_2d”) aerosol species collection (“aer”) on the horizontal grid (“Nx”)
- 202106 – This file contains monthly averages for June 2021



AQ-Relevant Collections and Variables

<https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf> *

Collection Name	Description
tavg1_2d_aer_Nx	Aerosol Diagnostics <i>Optical properties (Extinction AOT, Scattering AOT, Angstrom Parameter) Surface Concentration (kg/m³), Column Density (kg/m²)</i>
tavg1_2d_adg_Nx	Aerosol Diagnostics (Extended) <i>Emissions and Removal Processes (Deposition, Sedimentation, and Scavenging)</i>
tavg1_2d_chm_Nx	2D Time-Averaged Chemistry Diagnostics <i>Surface CO, Column CO, Emissions, Chemical Loss, Chemical Production, Total Column O₃</i>
inst3_3d_aer_Nv	3D Instantaneous Aerosol Diagnostics <i>Mass Mixing Ratios (kg/kg) of Aerosol Species in Each Size Bin</i>
inst3_3d_chm_Nv	3D Instantaneous Chemistry Diagnostics <i>CO Molar Mixing Ratio (mol/mol), O₃ (not for use in scientific analysis)</i>

To calculate PM_{2.5}, use the formula:

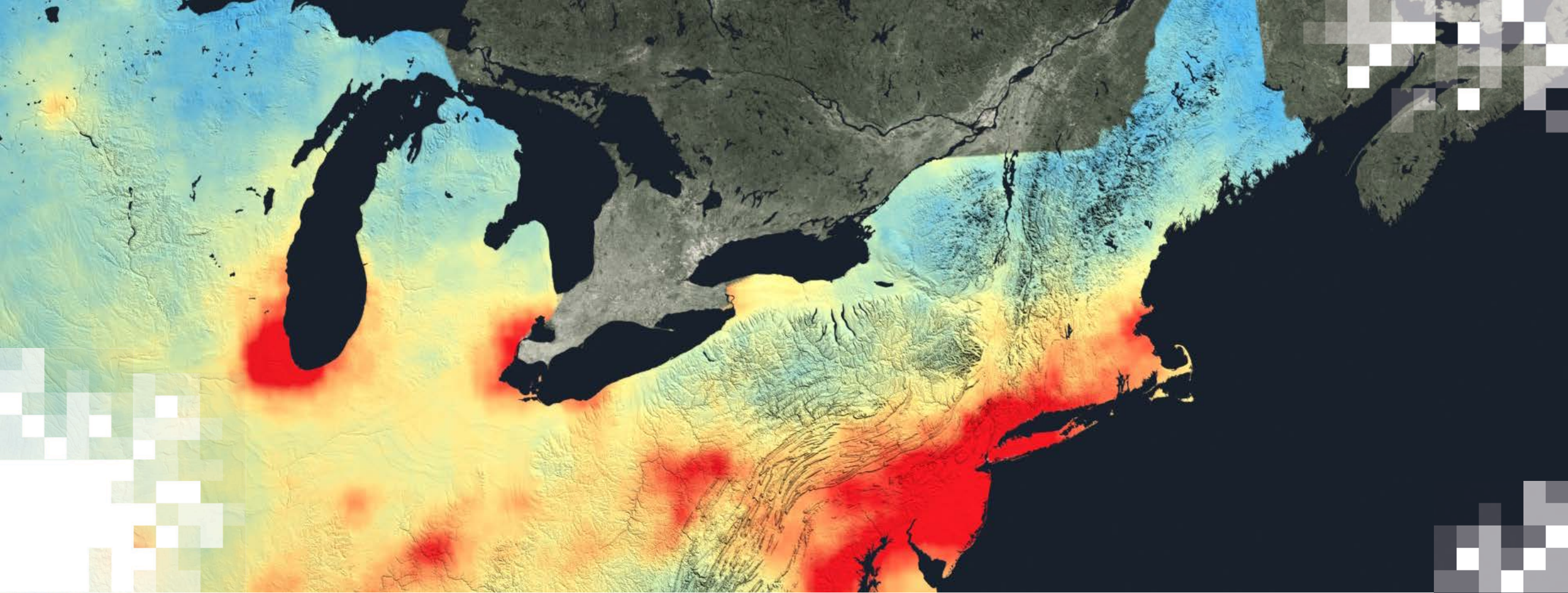
$$PM_{2.5} = [DUSMASS25] + [SSSMASS25] + [BCSMASS] + [OCMASS] + 1.375 \times [SO4MASS]$$

These variables are contained in the *tavg1_2d_aer_Nx* collection



References

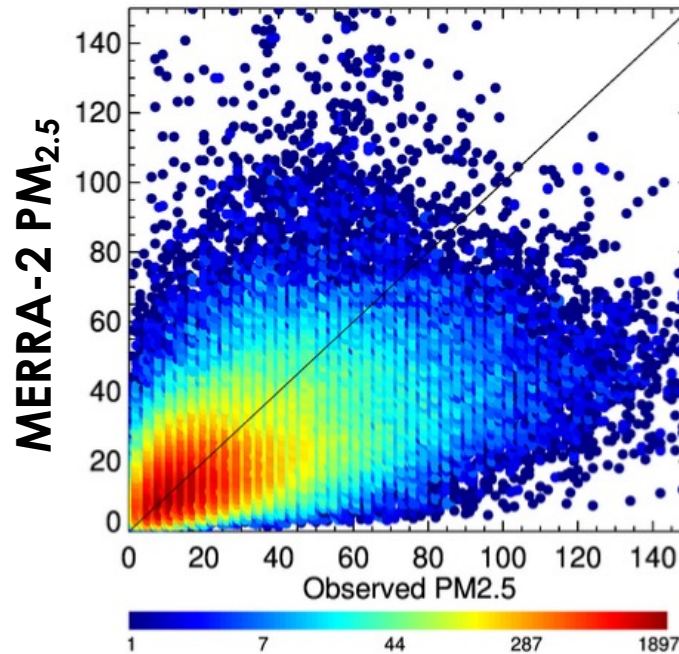
- GOCART
 - Chin, M., P. Ginoux, S. Kinne, O. Torres, B. Holben, B. Duncan, R. Martin, J. Logan, A. Higurashi, and T. Nakajima (2002), Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements, J. Atmos. Sci., 59(3), 461–483. [https://doi.org/10.1175/1520-0469\(2002\)059<0461:TAOTFT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2)
 - Colarco, P., A. Da Silva, M. Chin, and T. Diehl (2010), Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth, J. Geophys. Res.-Atmos., 115, –, doi:10.1029/2009JD012820. <https://doi.org/10.1029/2009JD012820>
 - Liu, F. et al. (2018). A new global anthropogenic SO₂ emission inventory for the last decade: A mosaic of satellite-derived and bottom-up emissions - <https://doi.org/10.5194/acp-18-16571-2018>
 - Janssens-Maenhout, G. et al. (2015). HTAP_v2.2: A mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution - <https://doi.org/10.5194/acp-15-11411-2015>
 - Ginoux, P. et al (2001) Sources and global distributions of dust aerosols simulated with the GOCART model, J. Geophys. Res., <https://doi.org/10.1029/2000JD000053>
 - Carn, S. (2019). Multi-satellite volcanic sulfur dioxide L4 long-term global database V3. <https://doi.org/10.5067/measures/so2/data404>
- GEOS-CF
 - Keller et al. (2021) Description of the NASA GEOS Composition Forecast Modeling System GEOS-CF v1.0, <https://doi.org/10.1029/2020MS002413>
 - File Specification Document – <https://gmao.gsfc.nasa.gov/pubs/docs/Knowland1204.pdf>
- GEOS FP
 - Reinecker et al. (2008), The GEOS-5 Data Assimilation System Documentation of Versions 5.0.1, 5.1.0, and 5.2.0,
 - File Specification Document - <https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf>
- MERRA-2 - <https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/>
 - File Specification Document - <https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf>
 - Aerosol Assimilation – Technical Document - <https://gmao.gsfc.nasa.gov/pubs/docs/Randles887.pdf>
 - Randles et al. (2017) The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part I: System Description and Data Assimilation Evaluation - <https://journals.ametsoc.org/view/journals/clim/30/17/jcli-d-16-0609.1.xml>
 - Buchard et al. (2017) The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies - <https://journals.ametsoc.org/view/journals/clim/30/17/jcli-d-16-0613.1.xml>



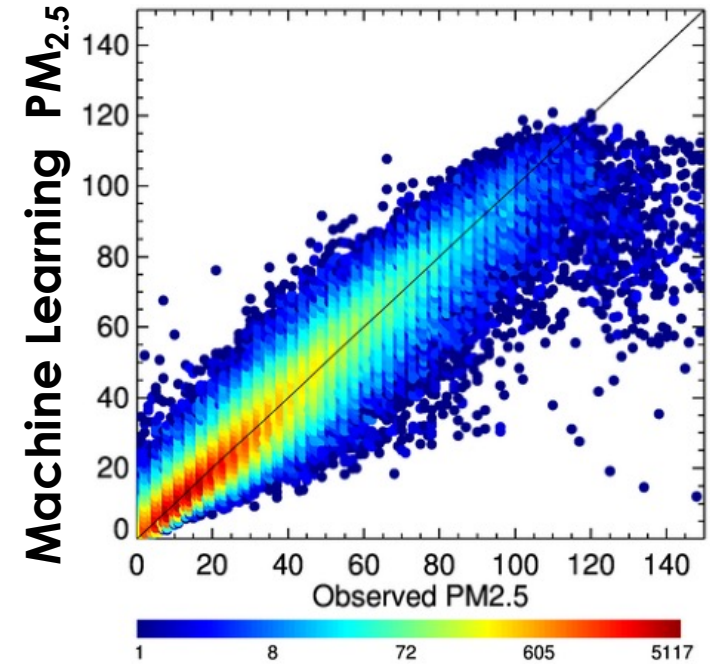
Machine Learning to Adjust Forecasts

Case Study: MERRA-2 as a Dataset for Machine Learning

- In a recent study, [Gupta et al. \(2021\)](#) used MERRA-2 output, along with ground observations of $PM_{2.5}$ to train a machine learning model to predict $PM_{2.5}$ in Thailand.
- The machine learning predicted $PM_{2.5}$ shows better correlation and reduced bias with respect to observations.
- This algorithm can be used to bias-correct the entire MERRA-2 time period, creating a more accurate long-term dataset for this region.



R = 0.53
Mean Bias = -3.8
RMSE = 16.7



R = 0.95
Mean Bias = 0.03
RMSE = 5.9

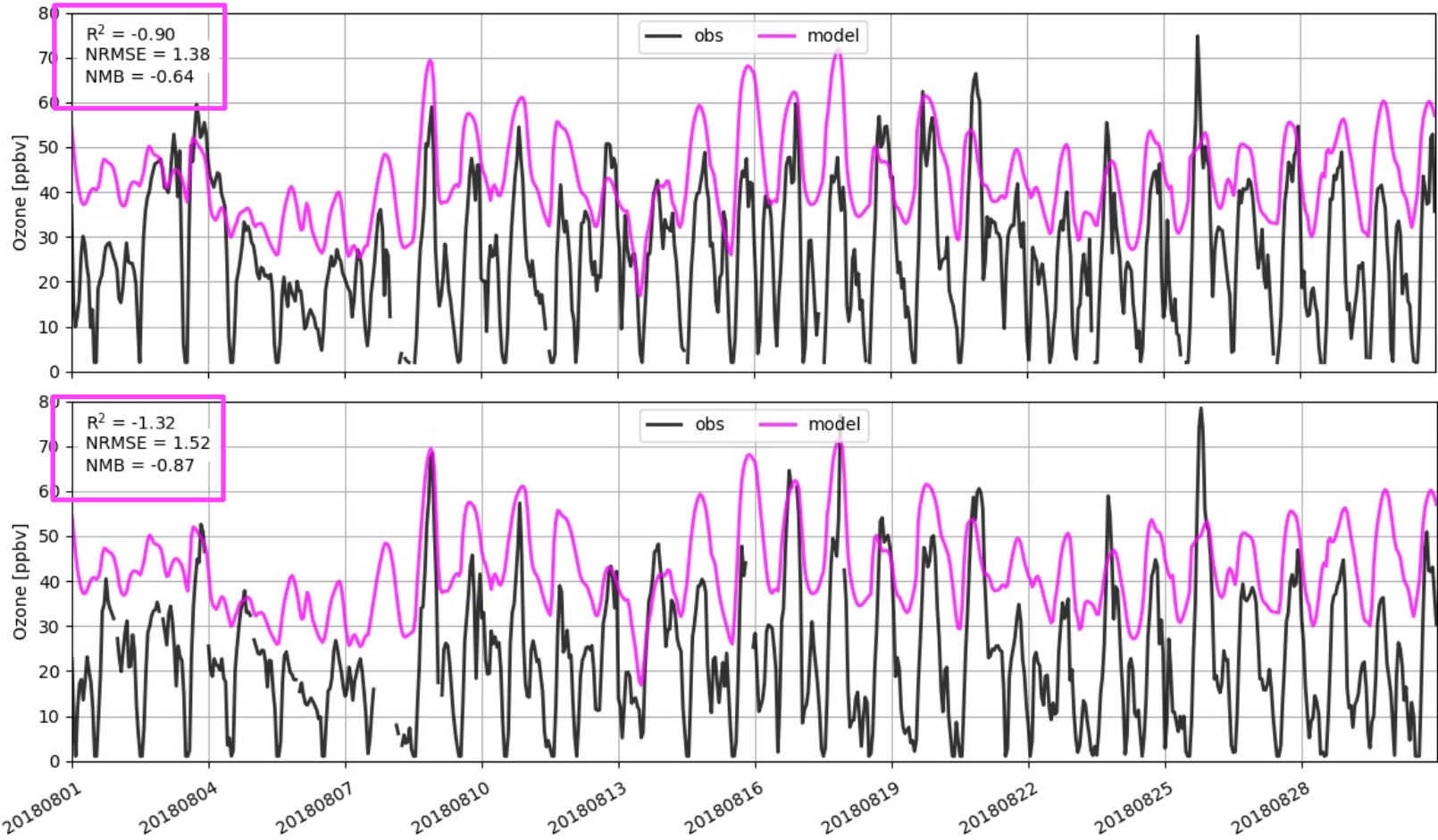


Improve Local Forecasts Using Statistical Bias Correction



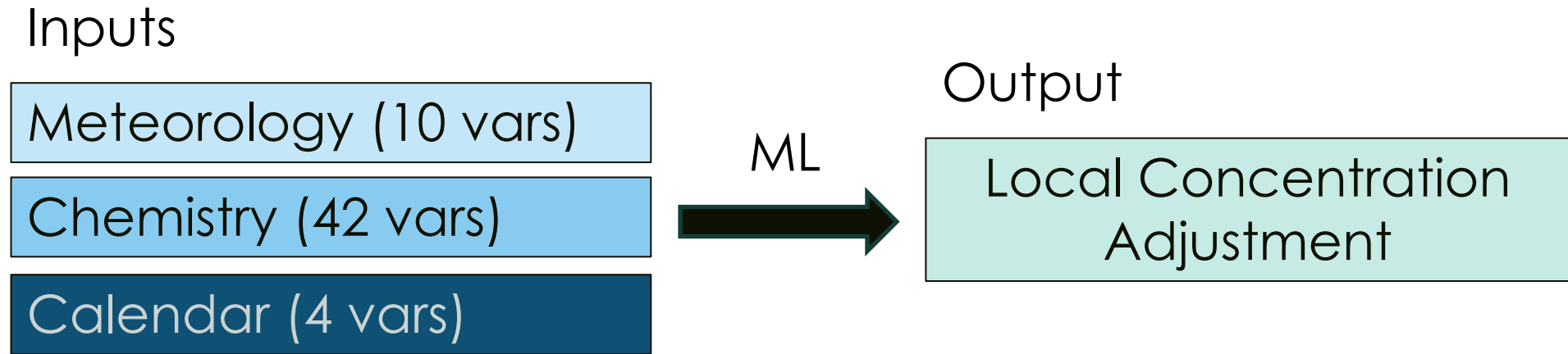
Two observation sites in the same grid box

➤ GEOS-CF generally overestimates



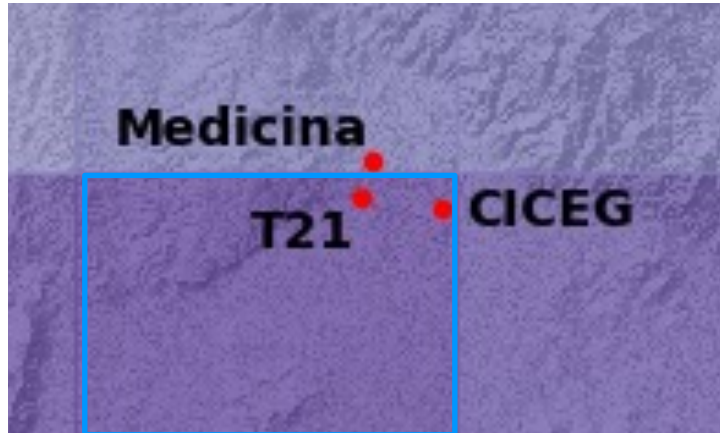
Observations Model

Use machine learning to correct for small scale variability and/or model biases.



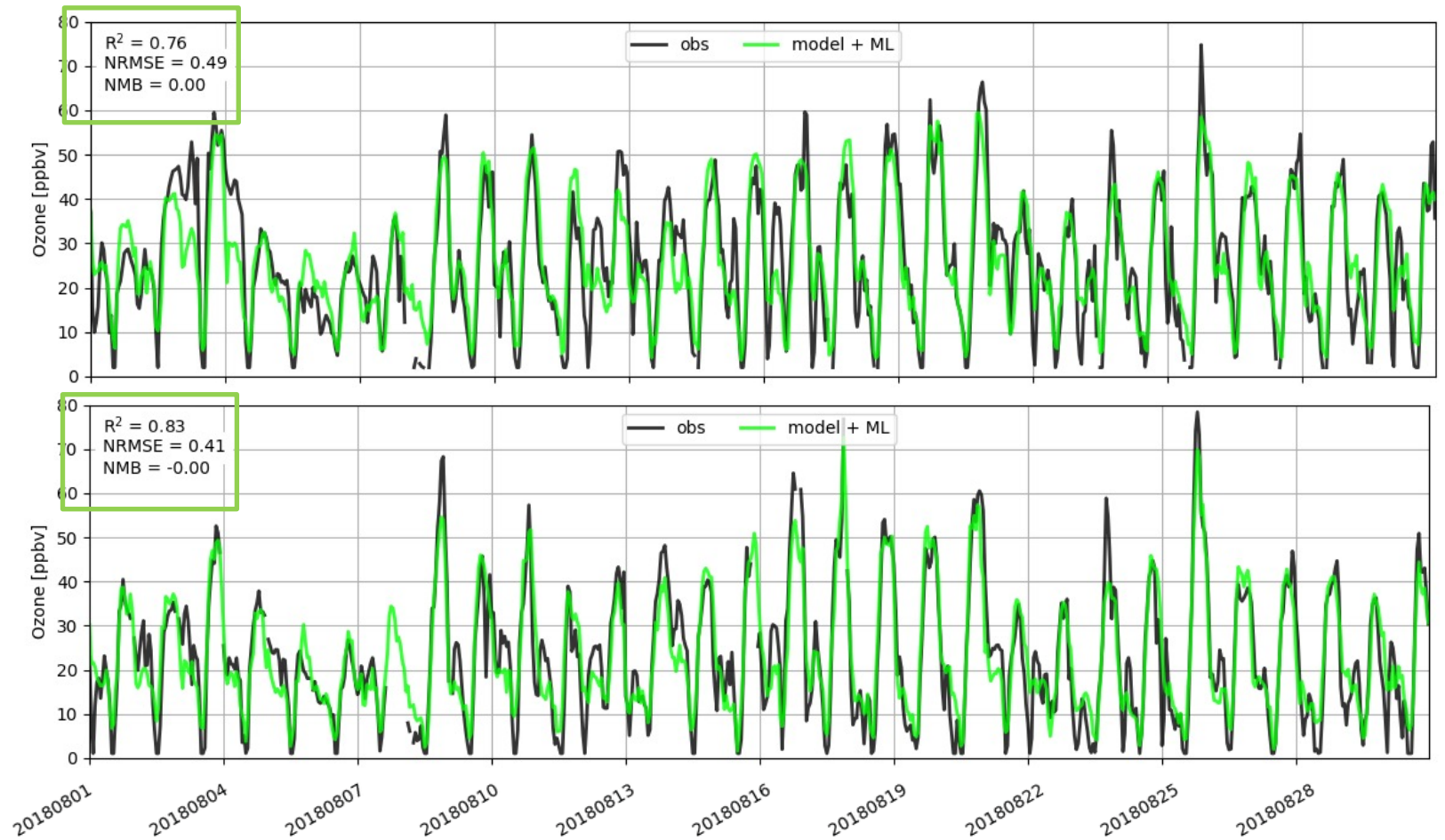
- Algorithm: Gradient Boosted Decision Trees (XGBoost)
- Train separate algorithm for each site
- See Keller et al., ACP 2021

Improve local forecasts using statistical bias correction.

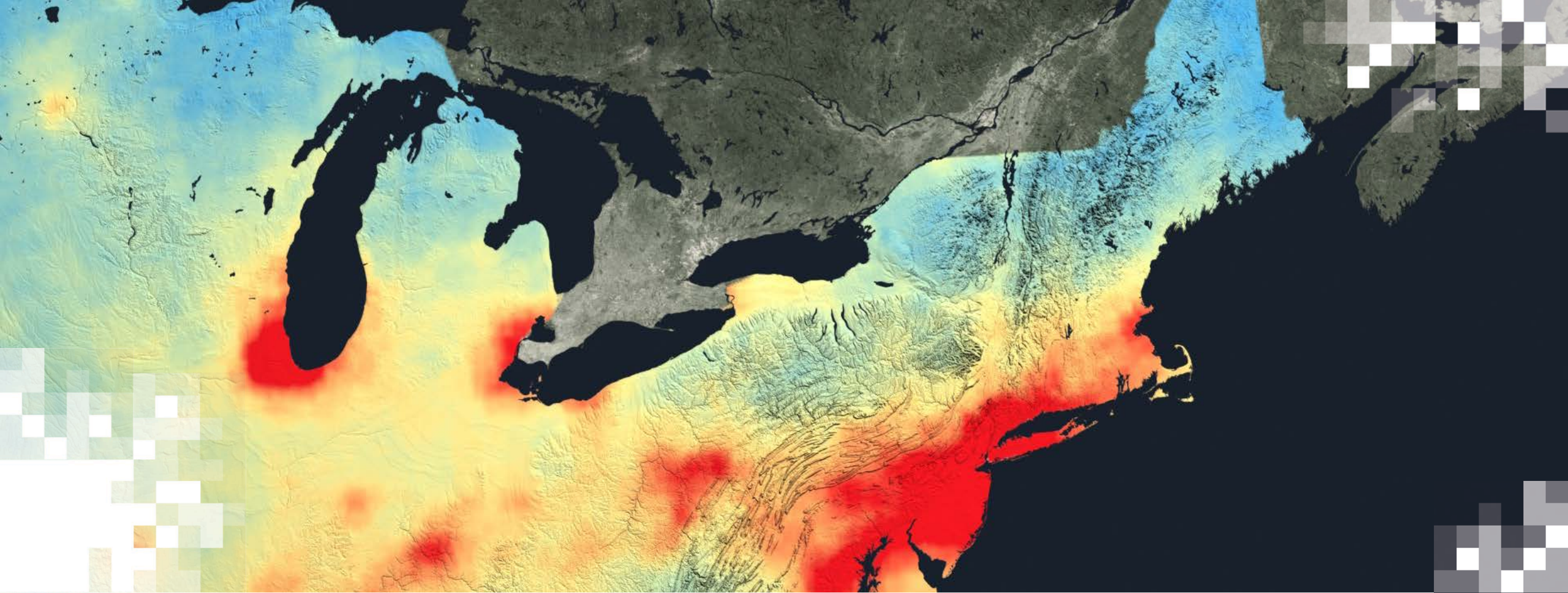


Two observation sites
in the same grid box

- GEOS-CF+ML captures diurnal variability at sub-grid scale



Observations **Model + ML**



Data Access and Visualization Tools

GMAO's Fluid

- Visualization tool for GEOS-FP, GEOS-CF, and MERRA-2
- Let's take a tour: <https://fluid.nccs.nasa.gov/weather/>



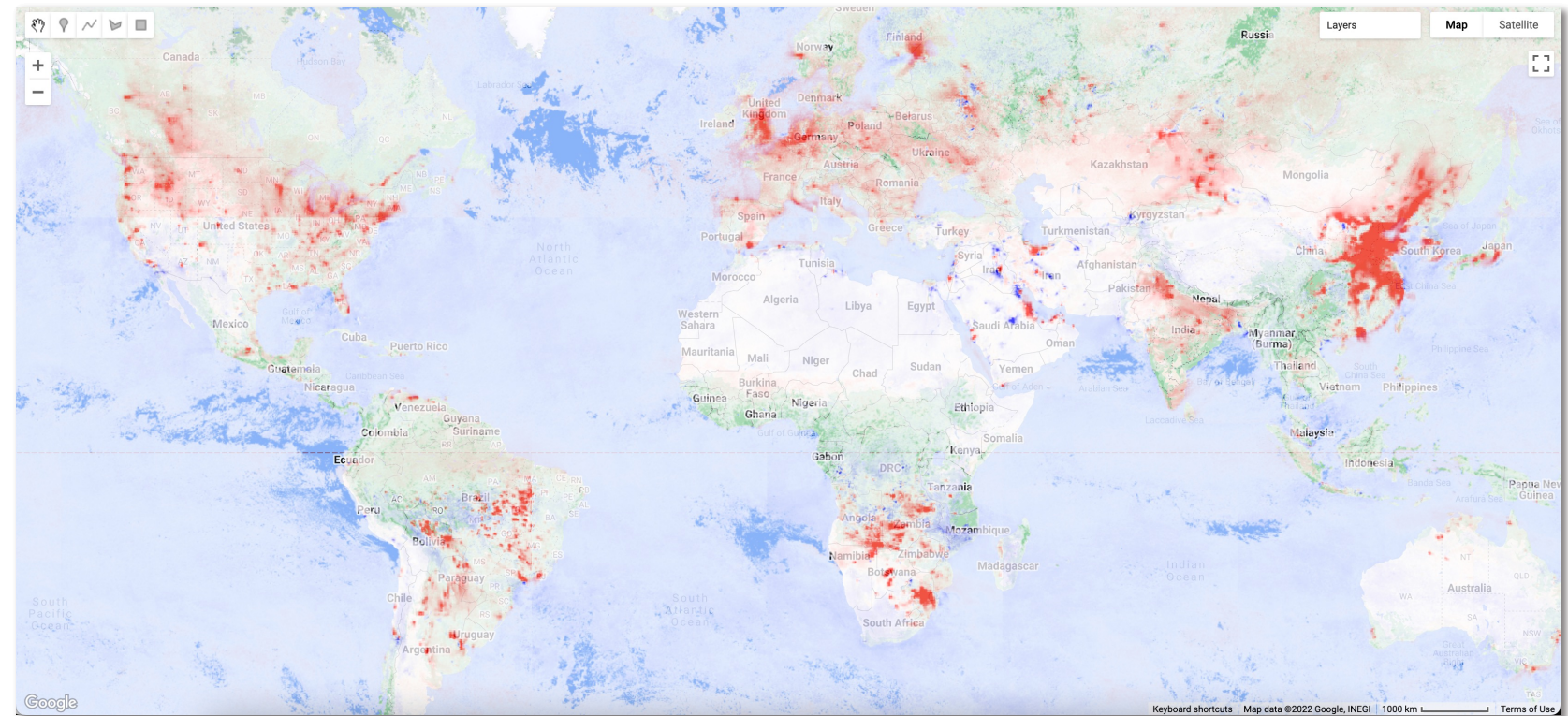
GEOS-CF is now available on Google Earth Engine.

- Increase data accessibility
- Allow easier intercomparison of diverse global datasets
- Implement data analysis with cloud computing resources

- Example (at right)

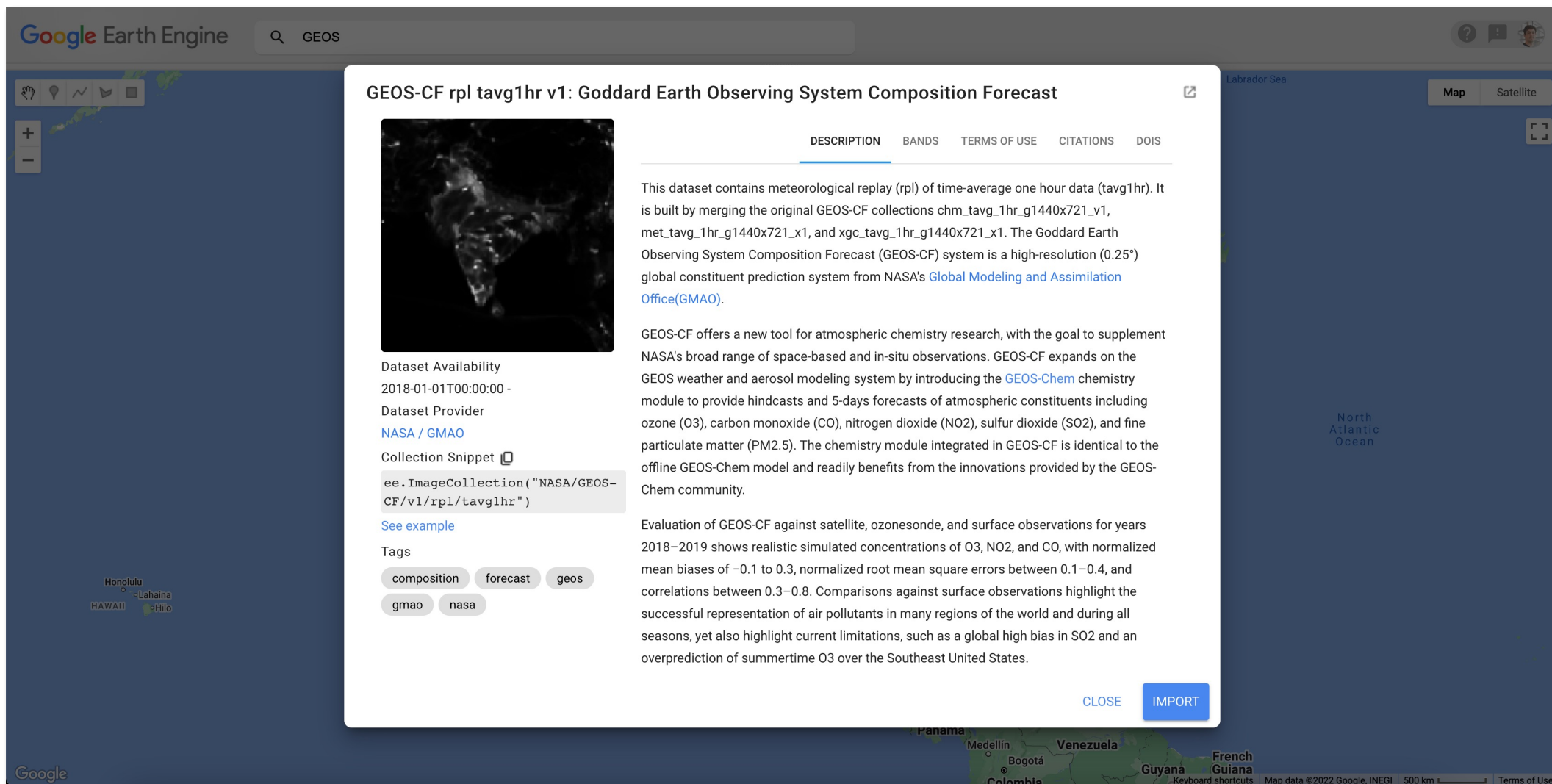
- Comparison of GEOS-CF tropospheric column NO₂ with TROPOMI (ESA Satellite) column NO₂ over a 1-week period

- Red: GEOS-CF higher
- Blue: TROPOMI higher



- Source: Google Earth Engine
<https://code.earthengine.google.com/>
- `ee.ImageCollection("NASA/GEOS-CF/v1/rpl/tavg1hr")`

NASA GMAO Datasets in GEE: GEOS-CF



The screenshot shows the Google Earth Engine (GEE) interface with a search bar containing "GEOS". A modal window is open for the dataset "GEOS-CF rpl tavg1hr v1: Goddard Earth Observing System Composition Forecast". The modal includes a thumbnail image of a dark, textured map, a description, dataset availability, provider information, a code snippet, tags, and buttons for "CLOSE" and "IMPORT".

GEOS-CF rpl tavg1hr v1: Goddard Earth Observing System Composition Forecast

DESCRIPTION BANDS TERMS OF USE CITATIONS DOIS

This dataset contains meteorological replay (rpl) of time-average one hour data (tavg1hr). It is built by merging the original GEOS-CF collections `chm_tavg_1hr_g1440x721_v1`, `met_tavg_1hr_g1440x721_x1`, and `xgc_tavg_1hr_g1440x721_x1`. The Goddard Earth Observing System Composition Forecast (GEOS-CF) system is a high-resolution (0.25°) global constituent prediction system from NASA's [Global Modeling and Assimilation Office\(GMAO\)](#).

GEOS-CF offers a new tool for atmospheric chemistry research, with the goal to supplement NASA's broad range of space-based and in-situ observations. GEOS-CF expands on the GEOS weather and aerosol modeling system by introducing the [GEOS-Chem](#) chemistry module to provide hindcasts and 5-days forecasts of atmospheric constituents including ozone (O3), carbon monoxide (CO), nitrogen dioxide (NO2), sulfur dioxide (SO2), and fine particulate matter (PM2.5). The chemistry module integrated in GEOS-CF is identical to the offline GEOS-Chem model and readily benefits from the innovations provided by the GEOS-Chem community.

Evaluation of GEOS-CF against satellite, ozonesonde, and surface observations for years 2018–2019 shows realistic simulated concentrations of O3, NO2, and CO, with normalized mean biases of -0.1 to 0.3, normalized root mean square errors between 0.1–0.4, and correlations between 0.3–0.8. Comparisons against surface observations highlight the successful representation of air pollutants in many regions of the world and during all seasons, yet also highlight current limitations, such as a global high bias in SO2 and an overprediction of summertime O3 over the Southeast United States.

Dataset Availability
2018-01-01T00:00:00 -

Dataset Provider
[NASA / GMAO](#)

Collection Snippet

```
ee.ImageCollection("NASA/GEOS-CF/v1/rpl/tavg1hr")
```

[See example](#)

Tags

composition forecast geos
gmao nasa

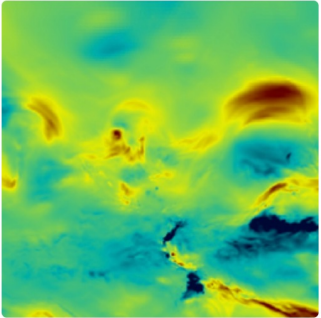
CLOSE **IMPORT**



NASA GMAO Datasets in GEE: MERRA-2

MERRA-2 M2T1NXAER: Aerosol Diagnostics V5.12.4

DESCRIPTION BANDS TERMS OF USE



M2T1NXAER (or `avg1_2d_aer_Nx`) is an hourly time-averaged 2-dimensional data collection in Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2). This collection consists of assimilated aerosol diagnostics, such as column mass density of aerosol components (black carbon, dust, sea-salt, sulfate, and organic carbon), surface mass concentration of aerosol components, and total extinction (and scattering) aerosol optical thickness (AOT) at 550 nm. The total PM1.0, PM2.5, and PM10 may be derived with the formula described in the [in the FAQ] (<https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/FAQ/>)

The data field is time-stamped with the central time of an hour starting from 00:30 UTC, e.g.: 00:30, 01:30, ..., 23:30 UTC.

MERRA-2 is the latest version of global atmospheric reanalysis for the satellite era produced by NASA Global Modeling and Assimilation Office (GMAO) using the Goddard Earth Observing System Model (GEOS) version 5.12.4. The dataset covers the period of 1980-present with the latency of ~3 weeks after the end of a month.

Dataset Availability
1980-01-01T00:00:00 -

Dataset Provider
[NASA/MERRA](#)

Collection Snippet

```
ee.ImageCollection("NASA/GSFC/MERRA/aer/2")
```

[See example](#)

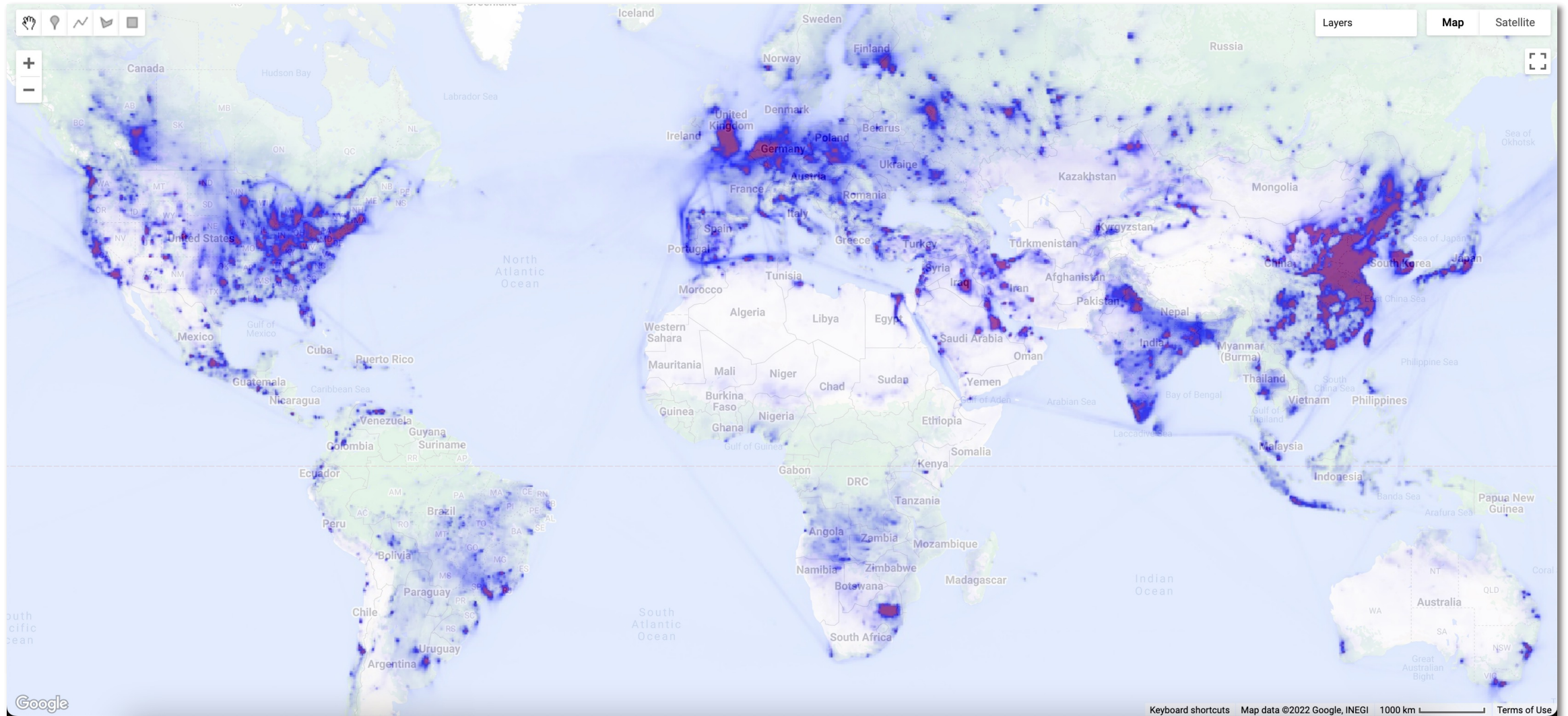
Tags

- aer
- aerosol
- carbon
- dust
- mass
- merra
- nasa
- sea-salt
- so2
- so4

CLOSE **IMPORT**



GEOS-CF Surface NO₂ in Google Earth Engine



Additional Resources

- MERRA-2 data is available at MDISC:
<https://disc.gsfc.nasa.gov/datasets?project=MERRA-2>
 - See also GMAO's MERRA2 data access website:
https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/data_access/
- GES DISC “how to” page has many useful tools:
<https://disc.gsfc.nasa.gov/information/howto>
- For more information on downloading, analyzing, and visualizing model output, see the ARSET webinar on “Tools for Analyzing NASA Air Quality Model Output”:
<https://appliedsciences.nasa.gov/join-mission/training/english/arset-tools-analyzing-nasa-air-quality-model-output>





Thank You!

