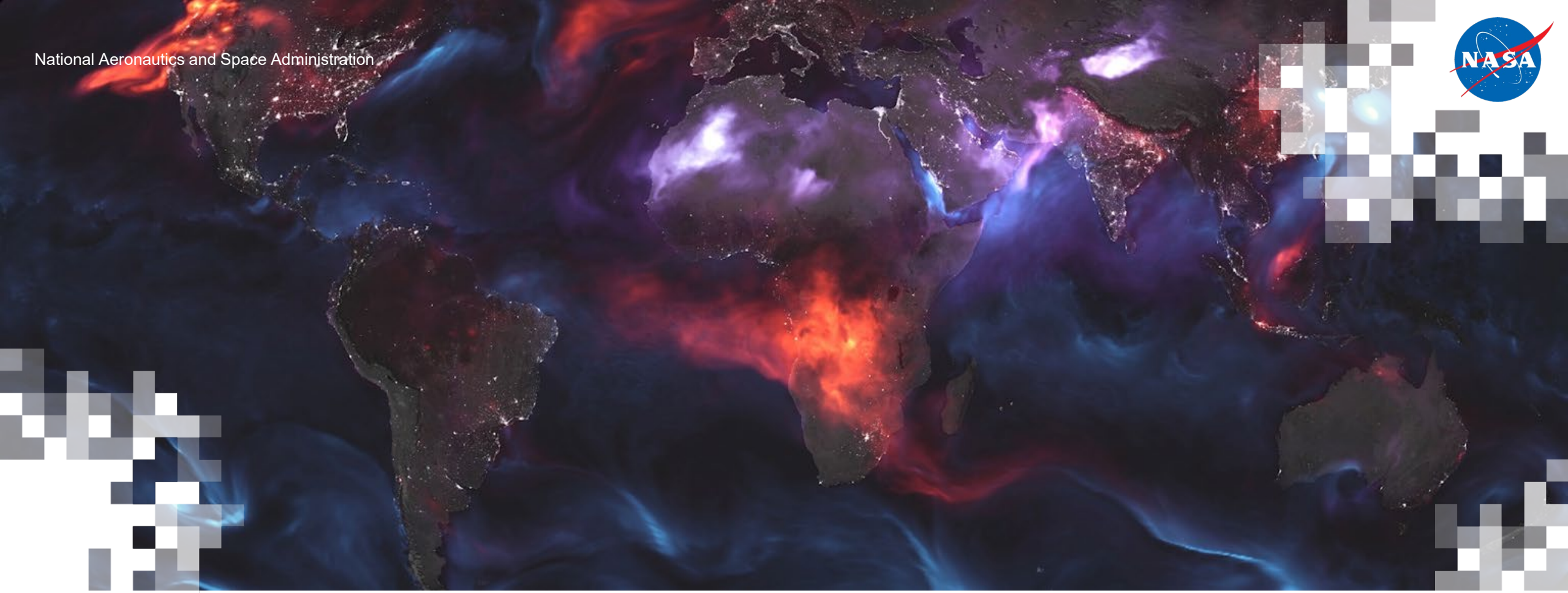


Tools for Analyzing NASA Air Quality Model Output

Part 1: Review of NASA Air Quality Forecasts and Reanalysis

Melanie Follette-Cook, Pawan Gupta, Sarah Strode, Xiaohua Pan, Suhung Shen, Binita KC, and Jennifer Wei

February 22, 24, and March 1, 2022

Course Information and Prerequisites

- Three 90-minute sessions on February 22, February 24, and March 1 from 10:00-12:00 EST (UTC-5)
- Webinar recordings and PowerPoint presentations can be found on the training webpage: https://appliedsciences.nasa.gov/join-mission/training/english/arset-tools-analyzing-nasa-air-quality-model-output

Prerequisites:

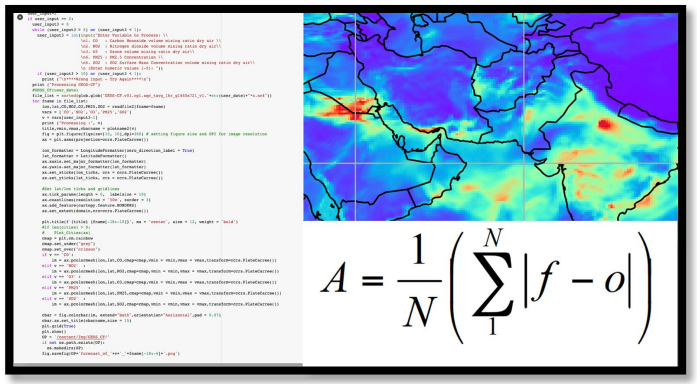
- Introduction and Access to Global Air Quality Forecasting Data and Tools https://appliedsciences.nasa.gov/join-mission/training/english/arset-introduction-and-access-global-air-quality-forecasting-data-and
- Create a NASA Earthdata Account and link GES DISC to your account
  - https://disc.gsfc.nasa.gov/earthdata-login
- Copy Jupyter Notebooks into Google Colab

Training Outline

Session 1: February 22, 2022

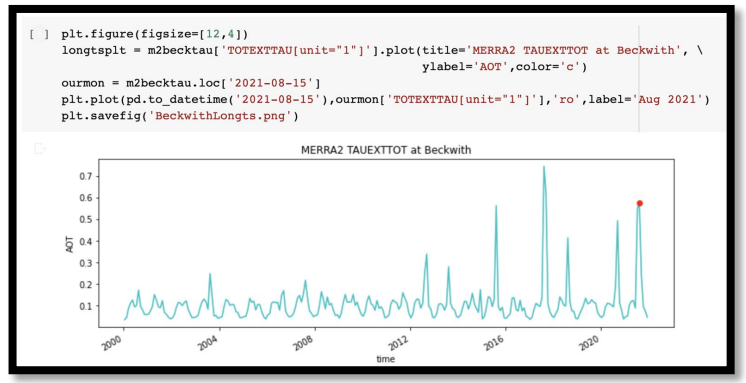
Review of NASA Air Quality Forecasts and Reanalysis

Session 2: February 24, 2022



Introduction to Python Tools for Visualization and Analysis

Session 3: March 1, 2022



Interpreting Model Output for Air Quality Assessment

Webinar Agenda

Session 1: February 22, 2022

Review of NASA Air Quality Forecasts and Reanalysis

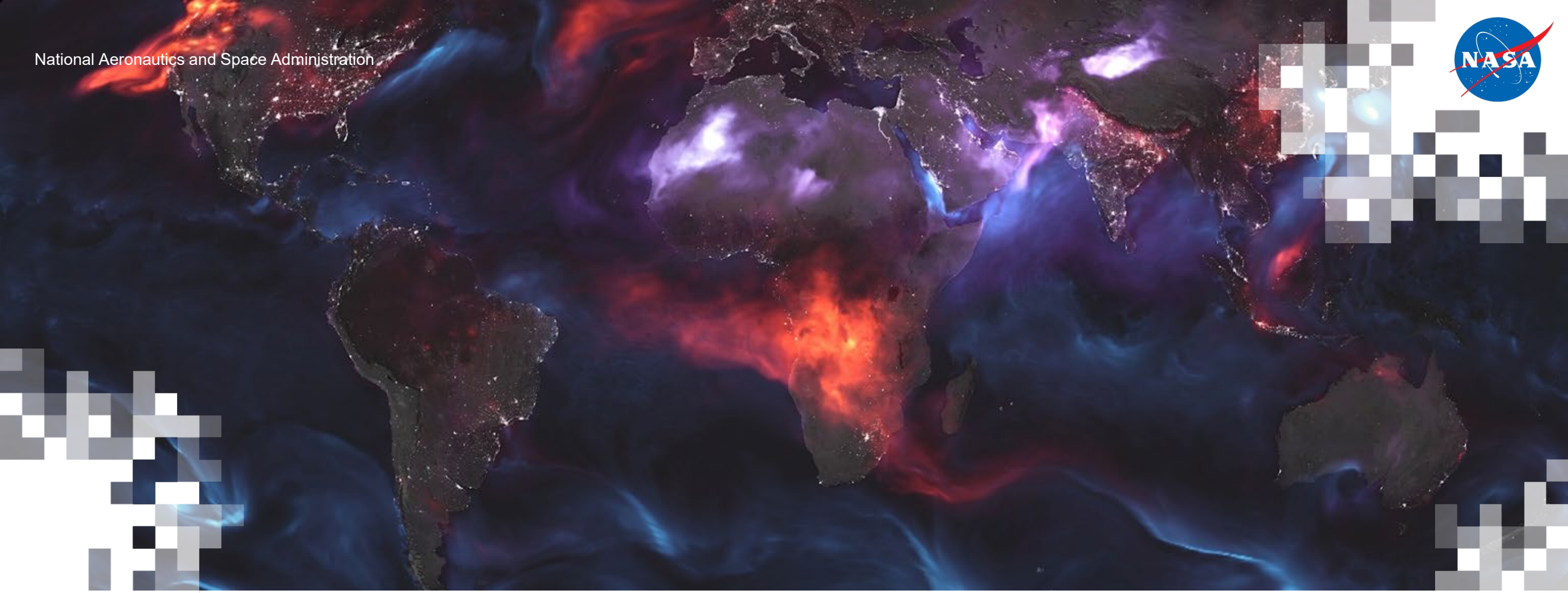
Melanie Follette-Cook

Pawan Gupta

Sarah Strode

Xiaohua Pan

National Aeronautics and Space Administration



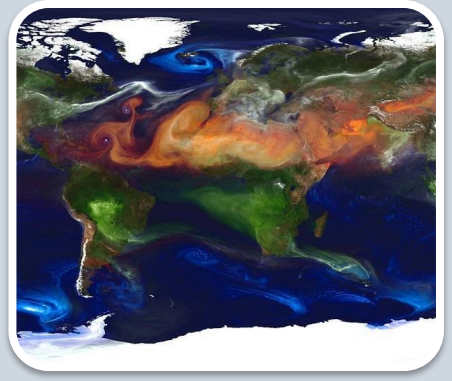
Review of NASA Air Quality Forecasts and Reanalysis

Learning Objectives

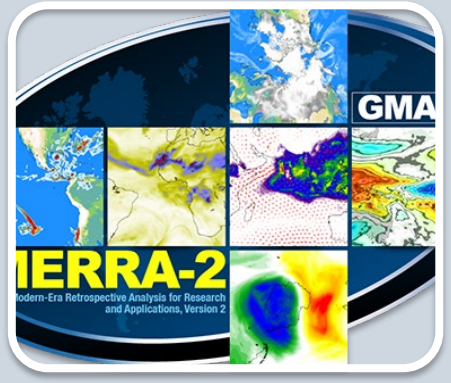
By the end of this session participants will be able to:

- Review and 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 forecasts
- Identify which MERRA-2 data are frequently used in air quality (AQ) studies
- Locate and access MERRA-2 AQ data and additional resources available from the GES DISC website
- Subset, download, and analyze the MERRA-2 AQ data using the web tool and Python API developed by the GES DISC

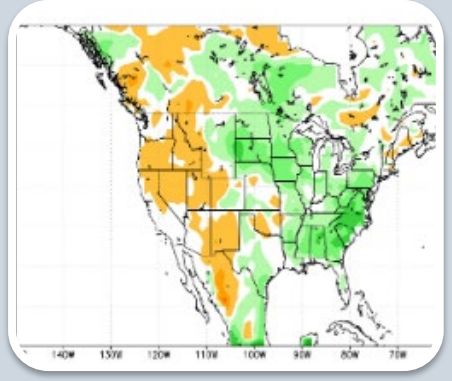
NASA GEOS Earth System Model



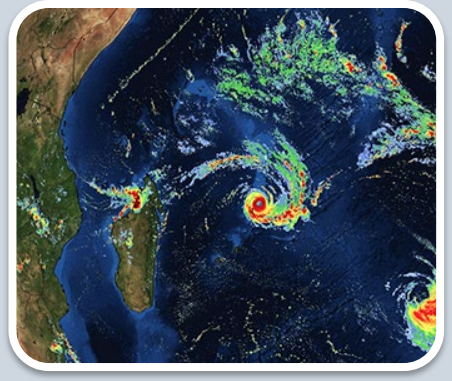
NRT Weather and Chemical Forecasts



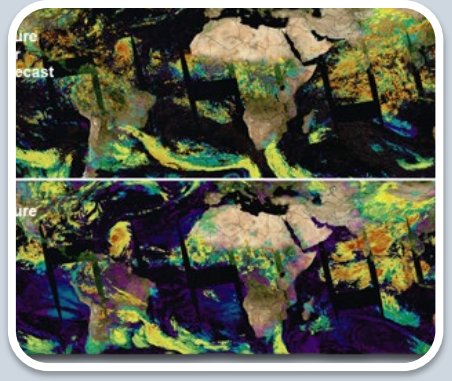
Reanalysis



Seasonal to Sub-Seasonal (S2S) and Decadal Prediction



High Resolution Mesoscale Modeling



Observing System Science

GEOS Forecast and Reanalysis Products

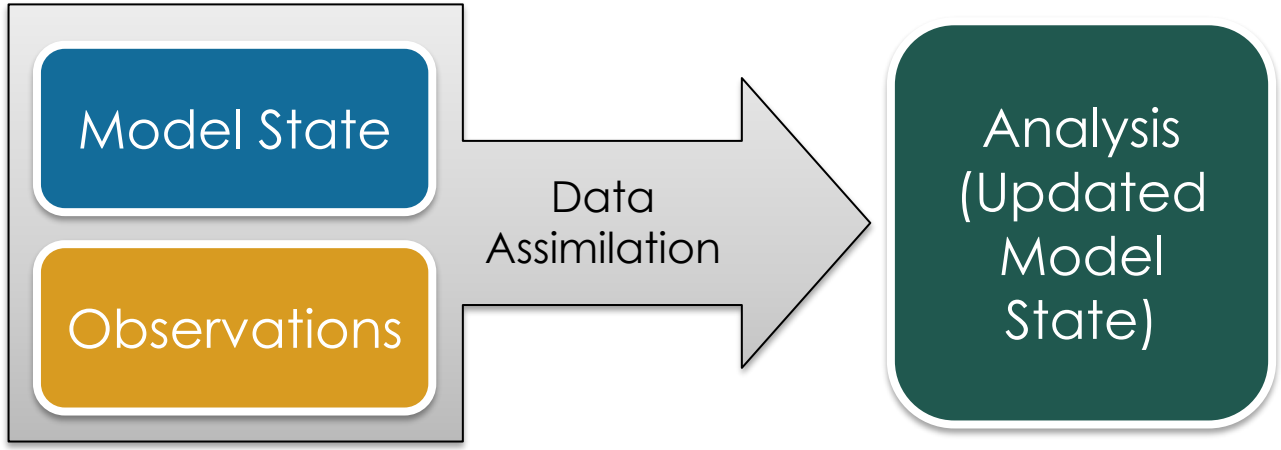
GEOS Forward Processing (GEOS FP) NRT Analysis and Forecast

GEOS-Composition Forecast (GEOS-CF) NRT Forecast

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

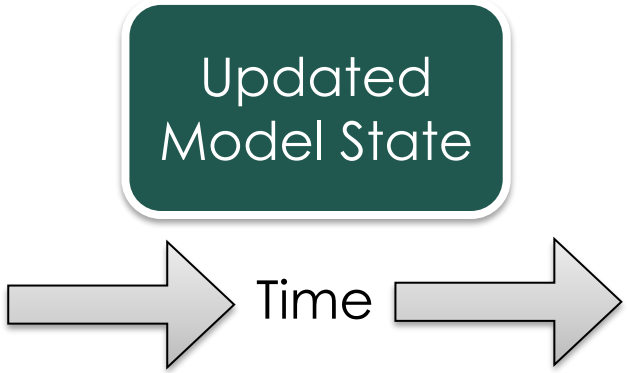
Forecast, Analysis, Reanalysis, Data Assimilation
What are the differences between these?

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.

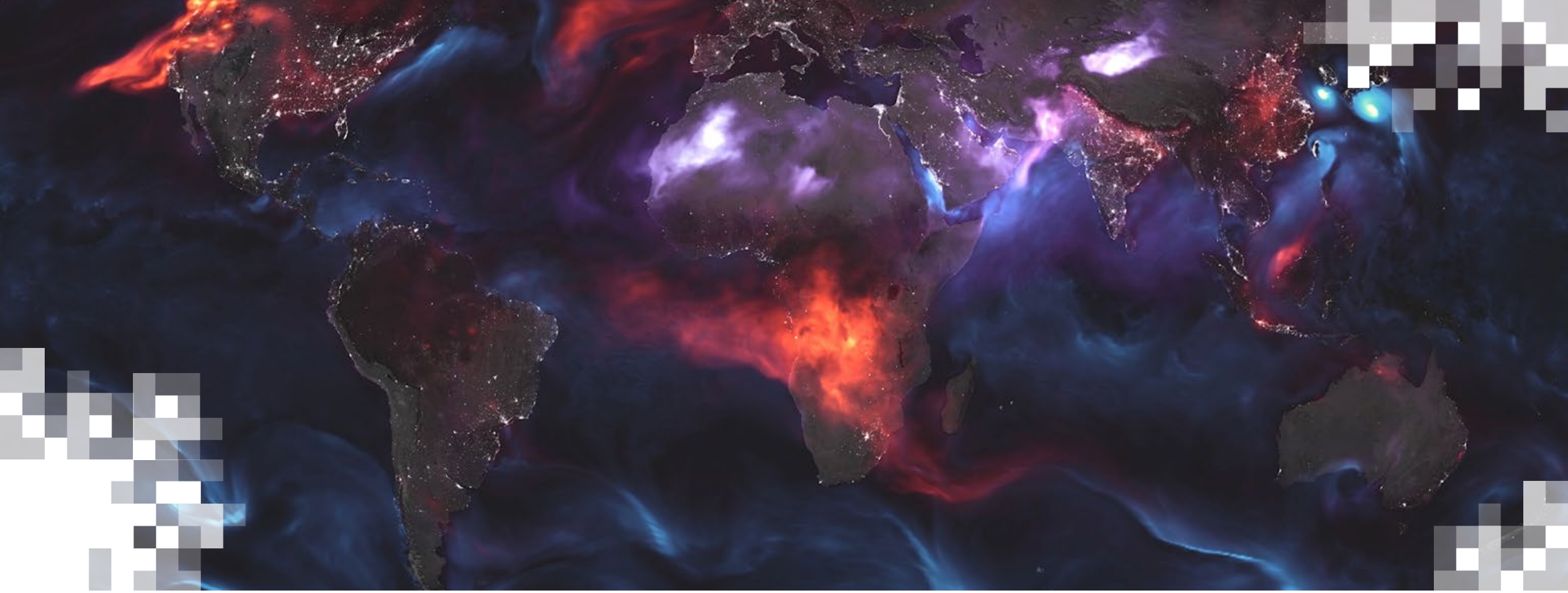
Advantage and Disadvantages of AQ modeling

Strengths:

- Forecasts provide the public with advance notice of the potential for poor air quality.
- Models can be used to test emission reduction strategies.
- Provide information about regions and pollutants that do not have ground observations.

Weaknesses:

- Models are only as good as what you put into them.
  - Emissions
  - Initial Conditions
- Evaluation/Verification
  - Model grid cell vs. a point measurement
  - Model output vs. observation (what is measured)



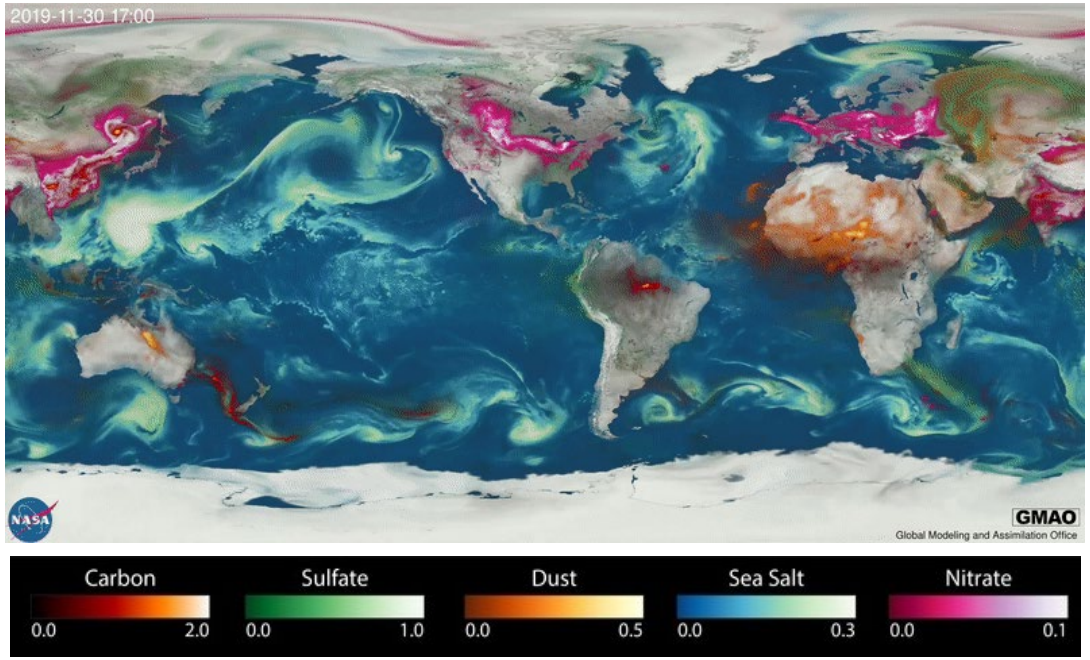
GEOS FP

GEOS FP

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

- State-of-the-science forecast system – model physics or observing system updated every 6-12 months
  - Not suitable for trend analyses
  - GMAO NRT Product Page
- Publicly available
- Includes weather, aerosols, and carbon monoxide (CO) on the same spatial scale
- GOCART aerosol scheme
  - Bulk aerosol scheme
  - Radiatively active
  - Dust, sea salt, black carbon (BC), organic carbon (OC), sulfate, and nitrate

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 *

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

AQ-Relevant Collections and Variables

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

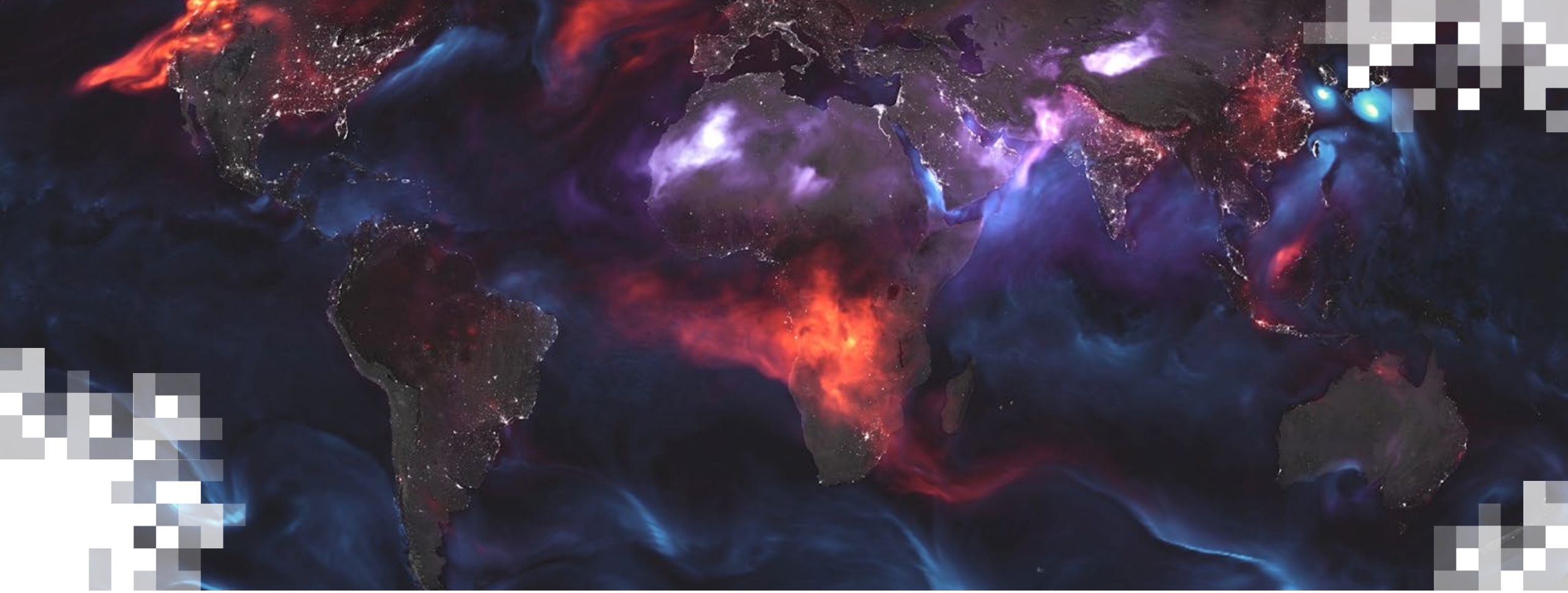
Collection Name	Description
tavg3_2d_aer_Nx	2D time-averaged aerosol diagnostics *Optical properties (Extinction AOT, Scattering AOT, Angstrom parameter)* *Surface concentration (kg/m^3), Column Density (kg/m^2)*
tavg3_2d_adg_Nx	2D time-averaged aerosol diagnostics (extended) *Emissions and removal processes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

To calculate $PM_{2.5}$, use the formula:

$PM_{2.5}$=[DUSMASS25]+[SSSMASS25]+[BCSMASS]+[OCSMASS]+1.375×[SO4SMASS]+1.29×[NISMASS25]

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.

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

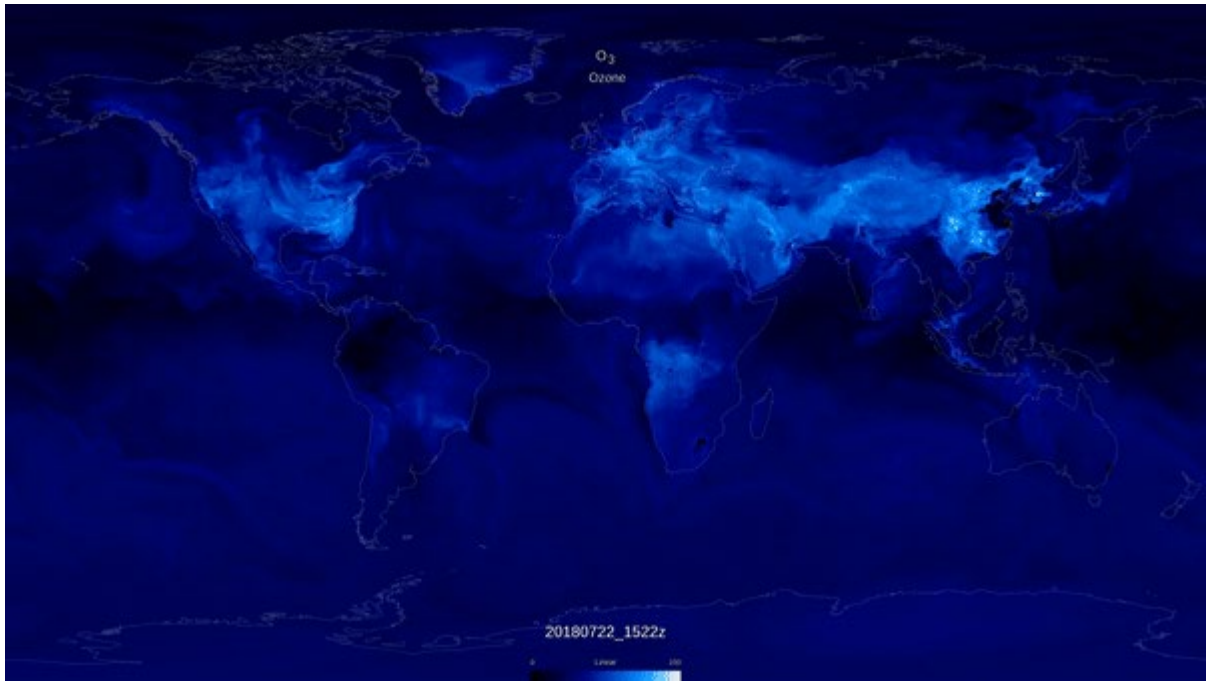


GEOS-CF

NASA Composition Forecasts (GEOS-CF)

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

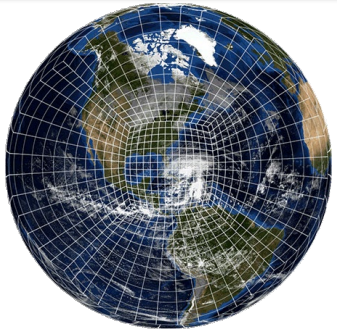
- Trace gas and aerosol forecast using constrained meteorology from GEOS and the GEOS-Chem chemical mechanism.
- Publicly available
- GEOS-Chem is a community-developed global 3-D model of atmospheric chemistry.
  - 250 chemical species
  - 725 chemical reactions
- Questions about GEOS-CF can be sent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



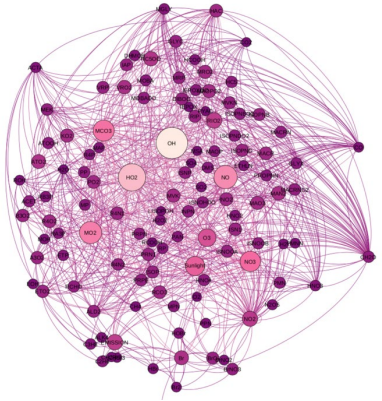
https://svs.gsfc.nasa.gov/4754

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 *

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

AQ-Relevant Collections and Variables

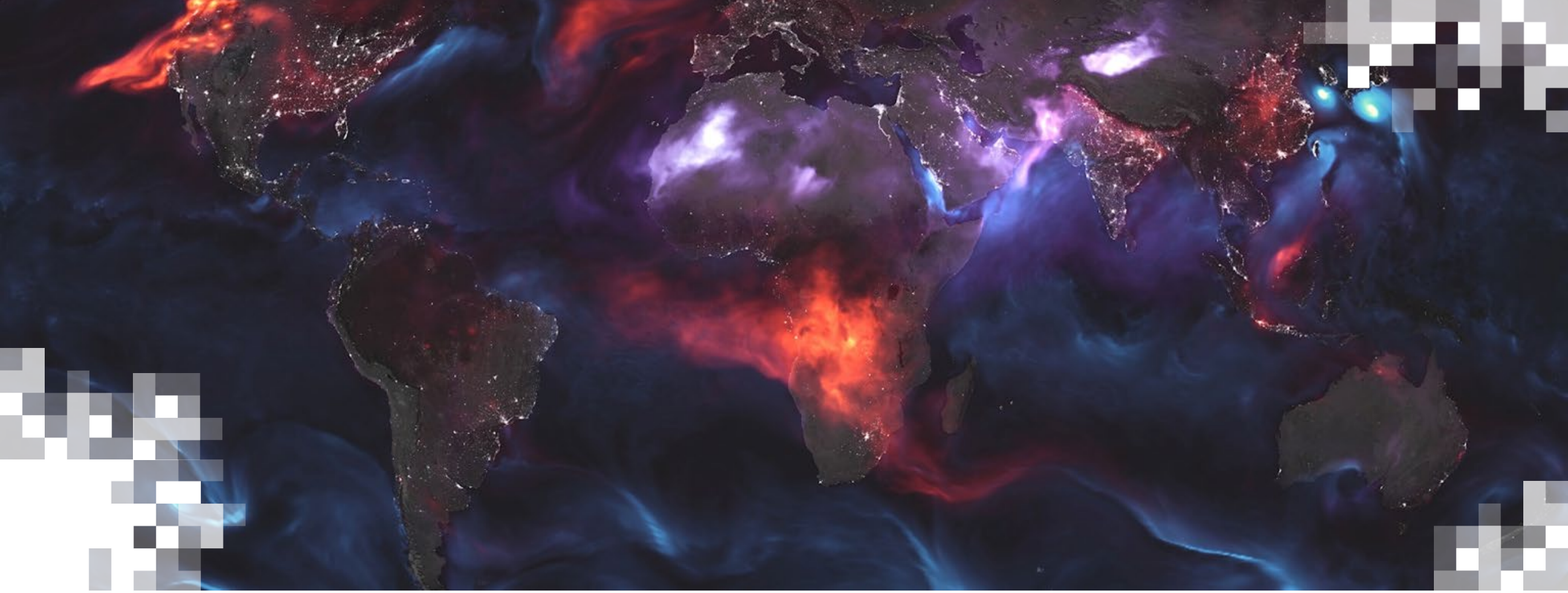
https://gmao.gsfc.nasa.gov/pubs/docs/Knowland1204.pdf *

Collection Name	Description
htf_inst_15mn_g1440x721_x1	High Temporal Frequency Chemistry and Meteorology *Surface CO, NO_2, O_3, SO_2, $PM_{2.5}$ (GCC & GOCART), and meteorology (RH, T, P, etc)*
aqc_tavg_1hr_g1440x721_v1	Air Quality Concentrations *One-hour time-averaged surface CO, NO_2, O_3, $PM_{2.5}$(GCC), SO_2*
chm_tavg_1hr_g1440x721_v1	Chemistry Fields *One-hour time-averaged surface mixing ratios of many chemical species and speciated $PM_{2.5}$ (GCC)*
xgc_tavg_1hr_g1440x721_x1	Extra GEOS-Chem Fields *One-hour time-averaged AOD, column quantities, and removal processes (deposition)*
chm_inst_1hr_g1440x721_p23	Chemistry Fields *3D (23 pressure levels,1000 to 10 hPa) instantaneous CO, NO_2, O_3, $PM_{2.5}$(GCC, speciated), SO_2*

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)

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php



MERRA-2

Reanalysis

What is reanalysis?

- 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

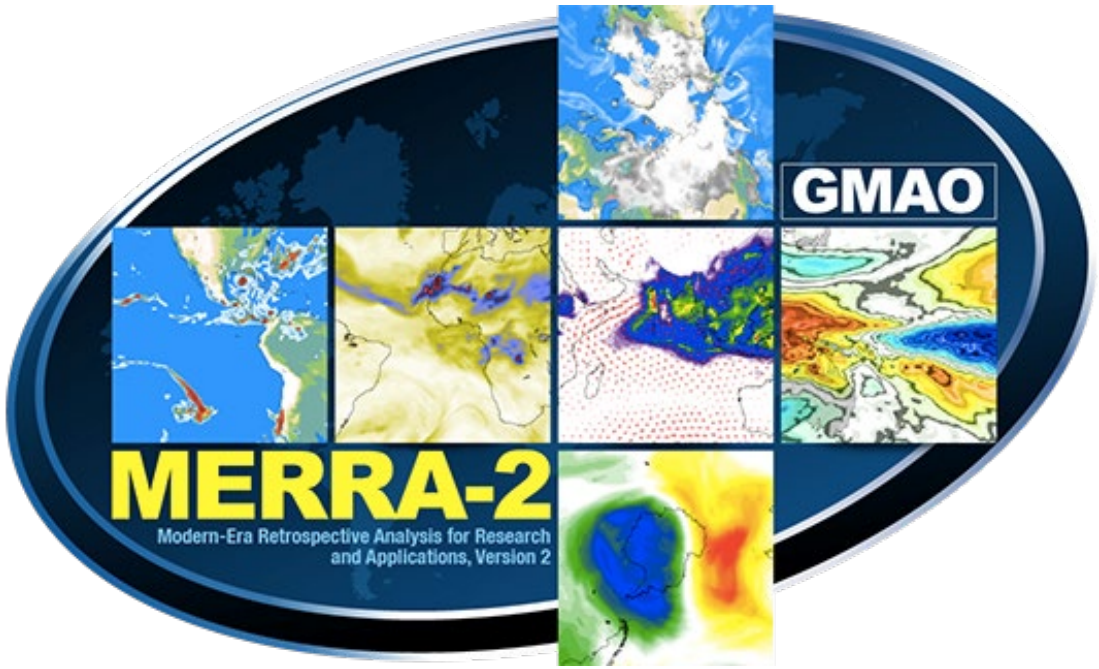
Why do we do it?

- 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° × 0.66° (~50 km) grid
- GOCART aerosol scheme
  - Similar to GEOS FP, but *no nitrate aerosols*



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 (aqc 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	~3 weeks 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 *

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

MERRA-2 Aerosol Observations

- Aerosol assimilation is 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.
- Information to keep in mind when using MERRA-2 aerosol output:
  - 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.

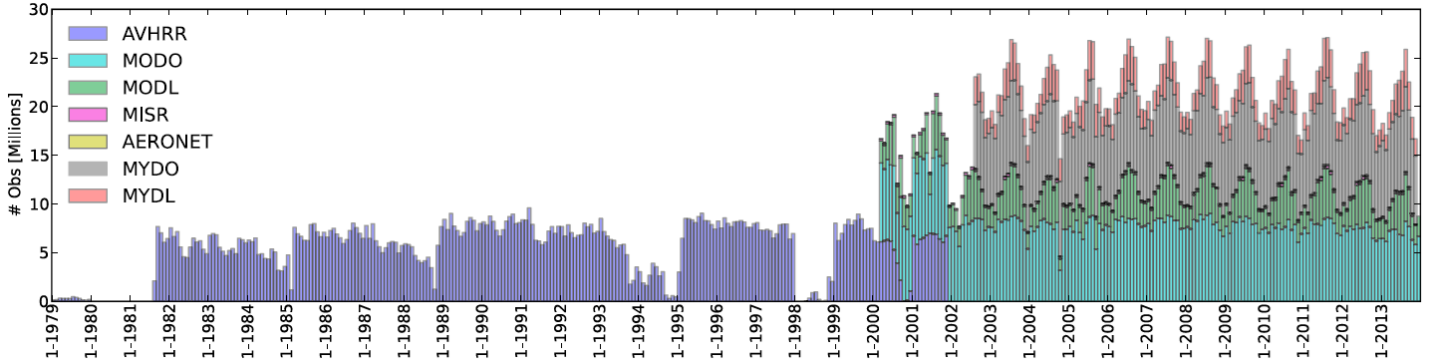


Figure 3 from Randles et al. 2017

AQ-Relevant Collections and Variables

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

Collection Name	Description
tavg1_2d_aer_Nx	Aerosol Diagnostics *Optical properties (Extinction AOT, Scattering AOT, Angstrom parameter)* *Surface concentration (kg/m^3), Column Density (kg/m^2)*
tavg1_2d_adg_Nx	Aerosol Diagnostics (extended) *Emissions and removal processes (deposition, sedimentation, and scavenging)*
tavg1_2d_chm_Nx	2D time-averaged chemistry diagnostics *Surface CO, column CO, emissions, chemical loss, chemical production, total column O_3*
inst3_3d_aer_Nv	3D instantaneous aerosol diagnostics *Mass mixing ratios (kg/kg) of aerosol species in each size bin*
inst3_3d_chm_Nv	3D instantaneous chemistry diagnostics *CO molar mixing ratio (mol/mol), O_3 (not for use in scientific analysis)*

To calculate $PM_{2.5}$, use the formula:

$$PM_{2.5}=[DUSMASS25]+[SSSMASS25]+[BCSMASS]+[OCSMASS]+1.375\times[SO4SMASS]$$

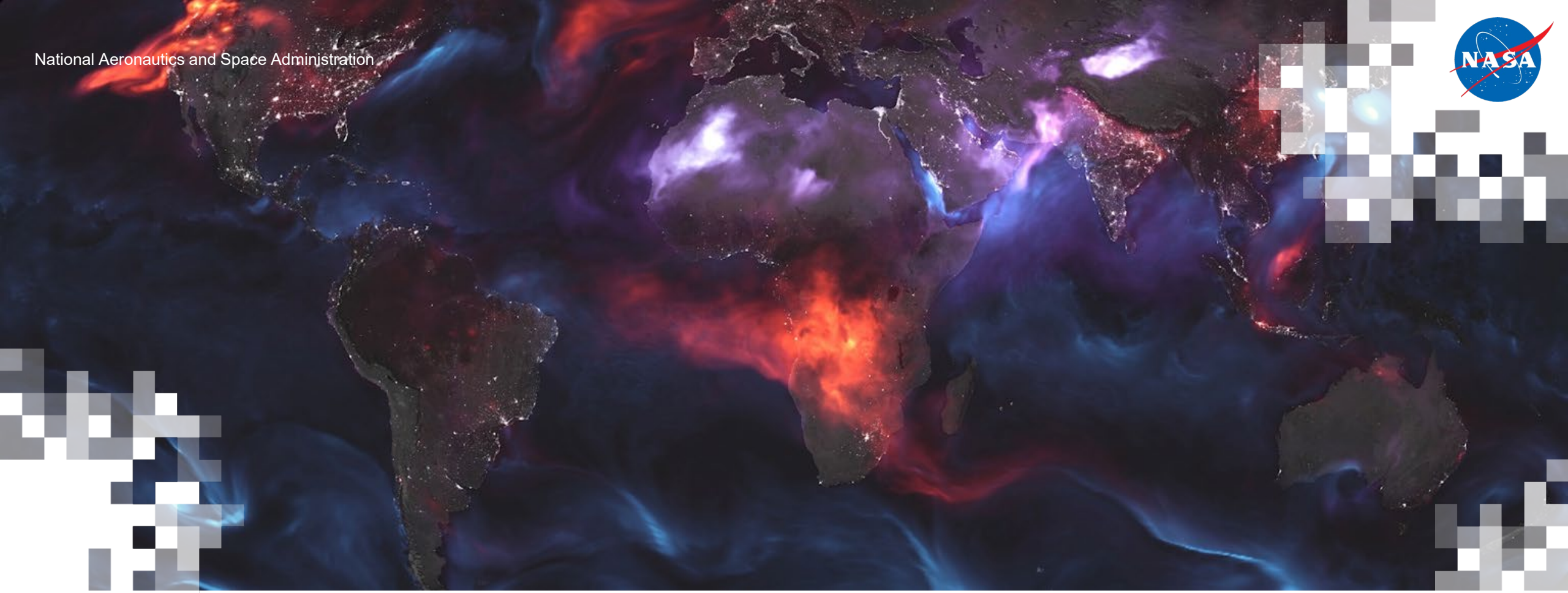
These variables are contained in the *tavg1_2d_aer_Nx* collection

* Find most current File Specification at https://gmao.gsfc.nasa.gov/pubs/office_notes.php

Downloading GEOS FP, GEOS-CF, and MERRA-2 output

- GEOS FP and GEOS-CF output can be downloaded using the GMAO FLUID site
  - GEOS FP: https://fluid.nccs.nasa.gov/weather/
  - GEOS-CF: https://fluid.nccs.nasa.gov/cf/
  - See our earlier training for a demonstration of the download tools
    - Session 2: Introduction and Access to Global Air Quality Forecasting Data and Tools
- MERRA-2 output is available from the GES DISC
  - https://disc.gsfc.nasa.gov/datasets?keywords=merra2&page=1
  - Information and demonstration in the second part of today's session

National Aeronautics and Space Administration



GES DISC Tools for Visualization and Analysis of MERRA-2

Xiaohua Pan, Suhung Shen, Binita KC, and Jennifer Wei (NASA/GSFC/GES DISC)

Learning Objectives

By the end of this session participants will be able to:

- Review and identify different Air Quality (AQ) relevant model outputs available from the Goddard Earth Observing System (GEOS) Earth System Model
- Understand the difference between analysis, reanalysis, and forecasts
- Identify which MERRA-2 data are frequently used in air quality (AQ) studies
- Locate and access MERRA-2 AQ data and additional resources available from the GES DISC website
- Subset, download, and analyze the MERRA-2 AQ data using the web tools and Python API developed by the GES DISC

1 Introduction: GES DISC Data Holding

- The Goddard Earth Sciences Data and Information Services Center (GES DISC), website: https://disc.gsfc.nasa.gov/
- One of 12 NASA Distributed Active Archive Centers (DAACs) since the mid-1990s
- "CoreTrustSeal" certified data repository
- Supports the archive and distribution of data from a range of satellite observations and Earth process models in multiple disciplines
  - Archive Size: ~150 million files covering > 3,000 public and restricted collections (~3.2 PB)
  - Distribution: > 3 billion files (~ 32 PB) (2010-present)
- **Focus Areas in NASA Earth Science Including:**
  - Atmospheric Composition
  - Carbon Cycle/Ecosystems (Atmospheric Carbon)
  - Water/Energy Cycles
  - Climate Variability
  - Weather/Atmosphere Dynamics

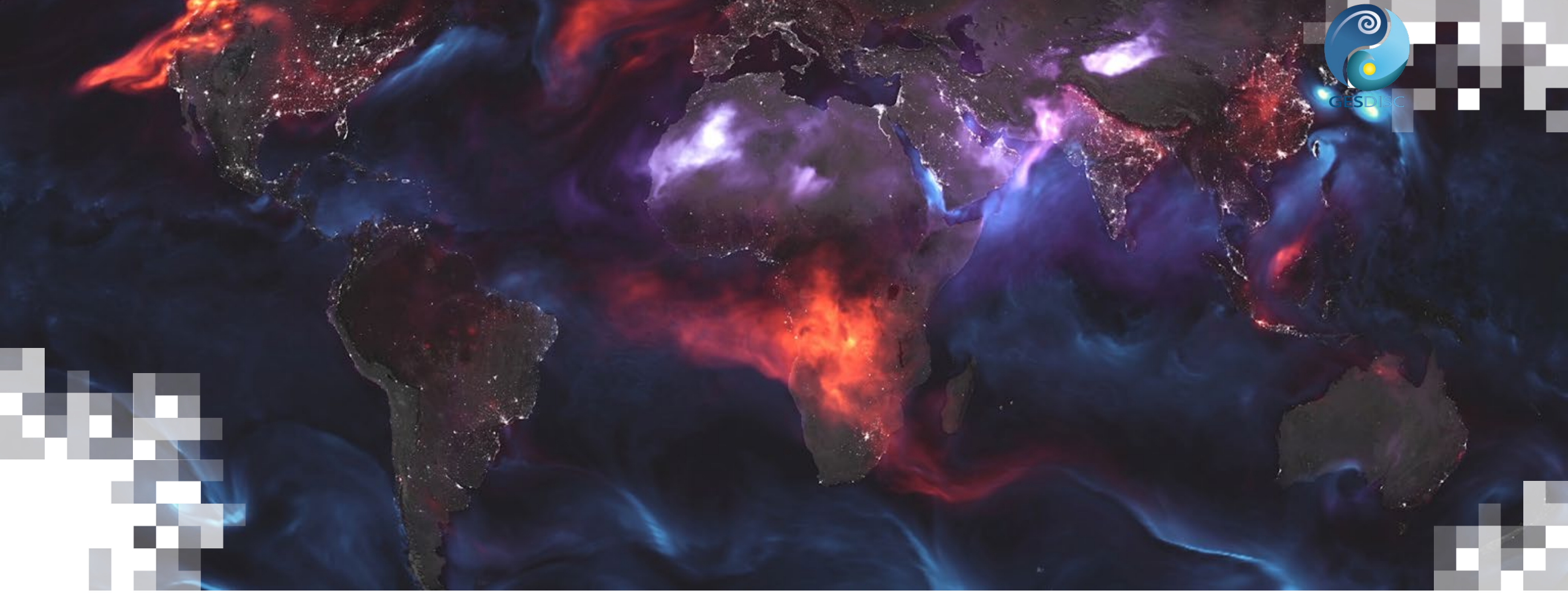
1 Introduction: GES DISC Services and Support

- Metadata support, documentation, metrics
  - Assign DOIs, generate landing pages
  - Generate metadata records, publish to CMR
  - Provide distribution metrics to EMS and data provider
- Web-based discovery and access to products
- Value-added services on data
  - Giovanni
  - Sub-setting, reformatting and re-gridding
  - Access protocols (e.g., OPeNDAP, OGC)
- Multi-tiered user support
  - GES DISC Help Desk: Contact us at gsfc-dl-help-disc@mail.nasa.gov
  - Forum: forum.earthdata.nasa.gov
  - Social Media Platform (e.g., Twitter @NASAEarthdata, and Youtube NASA Earthdata channel)
- User community engagement
  - User working group (UWG)
  - Workshops/webinars/training, e.g., by collaborating with NASA ARSET and Earthdata
  - Conferences, publications, news releases

1 Introduction: MERRA-2 at GES DISC

- Generated by NASA Global Modeling and Assimilation Office (GMAO)
- Archived and curated at NASA GES DISC
- ***Total of 100 collections*** *(95 standard collections, 4 derived climate statistical collections, and 1 value-added collection to be released very soon - monthly $PM_{2.5}$ at country level)*

- Model Version: 5.12.4
- Format: NetCDF4
- Temporal Range: 1980-01-01 to Present
- Temporal Resolution: Hourly, 3-Hourly, Daily, Monthly, and Monthly Diurnal
- Spatial Coverage: Global
- Spatial Resolution:
- 2D: 0.5° × 0.625°
- 3D: 0.5° × 0.625° × 72 model levels or
- 0.5° × 0.625° × 42 pressure levels
- **Data Latency: ~ 3 weeks**
- **Users: Over 7000 in 2021**



MERRA-2 Data for Air Quality Studies

Key MERRA-2 Collections and Variables Useful for Air Quality Studies

Measurements	Data Collection (data DOI hyperlink)	Parameters (variable long-name in the collection)	Temporal Feature
Aerosol Optical Depth (AOD)	M2T1NXAER_5.12.4 (i.e., tavg1_2d_aer_Nx)	total aerosol extinction aot [550 nm]	1 hourly averaged
	M2TMNXAER_5.12.4 (i.e., tavgM_2d_aer_Nx)	total aerosol extinction aot [550 nm]	monthly mean
& $PM_{2.5}$	M2T1NXAER_5.12.4 (i.e., tavg1_2d_aer_Nx)	surface mass concentration of SO_4, BC, OC, dust, and sea salt	1 hourly averaged
	M2TMNXAER_5.12.4 (i.e., tavgM_2d_aer_Nx)	surface mass concentration of SO_4, BC, OC, dust, and sea salt	monthly mean
@ $PM_{1.0}$ & PM_{10}	M2I3NVAER_5.12.4 * (i.e., inst3_3d_aer_Nv)	mixing ratio of SO_4, BC, OC, dust, and sea salt	3 hourly instantaneous
Carbon Monoxide (CO)	M2T1NXCHM_5.12.4 (i.e., tavg1_2d_chm_Nx)	CO surface concentration	1 hourly averaged
	M2TMNXCHM_5.12.4 (i.e., tavgM_2d_chm_Nx)	CO surface concentration	monthly mean
Meteorological Conditions	M2T1NXSLV_5.12.4 (i.e., tavg1_2d_slv_Nx)	wind, specific humidity, air temperature at 2-meter, 10-meter, 850 hPa	1 hourly averaged
	M2TMNXSLV_5.12.4 (i.e., tavgM_2d_slv_Nx)	wind, specific humidity, air temperature at 2-meter, 10-meter, 850 hPa	monthly mean
	M2T1NXFLX_5.12.4 (i.e., tavg1_2d_slv_Nx)	PBL height, precipitation	1 hourly averaged
	M2TMNXFLX_5.12.4 (i.e., tavgM_2d_slv_Nx)	PBL height, precipitation	monthly mean
	M2T3NVASM_5.12.4 * (i.e., tavg3_3d_asm_Nv	wind, relative humidity, air temperature, pressure, height at model levels	3 hourly averaged

& Surface $PM_{2.5}$ can be derived by variables listed in that collection: TOTSPM25 = OCSMASS+ BCSMASS+ SO4SMASS*1.375+ DUSMASS25 + SSSMASS25.

@ $PM_{1.0}$ and PM_{10} the formula can be found at the MERRA-2 project FAQ .

* In 3D (the rest collections in this table are in 2D).

MERRA-2 Collection Naming Convention Used by ESDIS

- MERRA-2 outputs are organized into file *collections* that contain related variables.
- The 9-character shortname of a collection has the form, e.g., M2TMNXAER

M2TimeFrequencyDimensionsGroup

Time (1 letter)

Time description
- C = constant
- I = instantaneous
- T = time-average
- S = statistics

Frequency (1 letter or number)

Frequency or averaging interval
- 1 = Hourly
- 3 = 3-Hourly
- 6 = 6-Hourly
- M = Monthly mean
- D = Daily statistics
- U = Monthly-Diurnal mean: consists monthly mean of data at each sub-daily time stamp.
- C = Climatology monthly mean (30 years mean from 1981 to 2010)

Dimensions (two letters)

Dimensions of variables
- NX = 2D (single level)
- NP = 3D at 42 Pressure levels
- NV = 3D at at 72 model layer center
- NE = 3D at at 73 model layer edge

Group (3 letters)

Type of variables
- These are also used in the shortname
  Ex. AER = Aerosol fields
- See documentation for full list

More Info: MERRA-2 Quick Guide - Data Collection

MERRA-2 Collection Naming Convention Used by GMAO

- MERRA-2 outputs are organized into file *collections* that contain related variables.
- The 12 or 13-character shortname of a collection has the form, e.g., tavgM_2d_aer_Nx

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)

Dimensions

Dimensions of variables

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

Group

Three letter abbreviation for the type of variables

- These are also used in the shortname
- 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

More Info: MERRA-2 Quick Guide - Data Collection

MERRA-2 File Naming Conventions

- Each MERRA-2 file has the form, e.g., MERRA2_400.tavgM_2d_aer_Nx.202110.nc4

MERRA2_SVv.collection.timestamp.nc4

Stream and Version

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

Collection

Collection Naming Convention used by GMAO

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

Timestamp

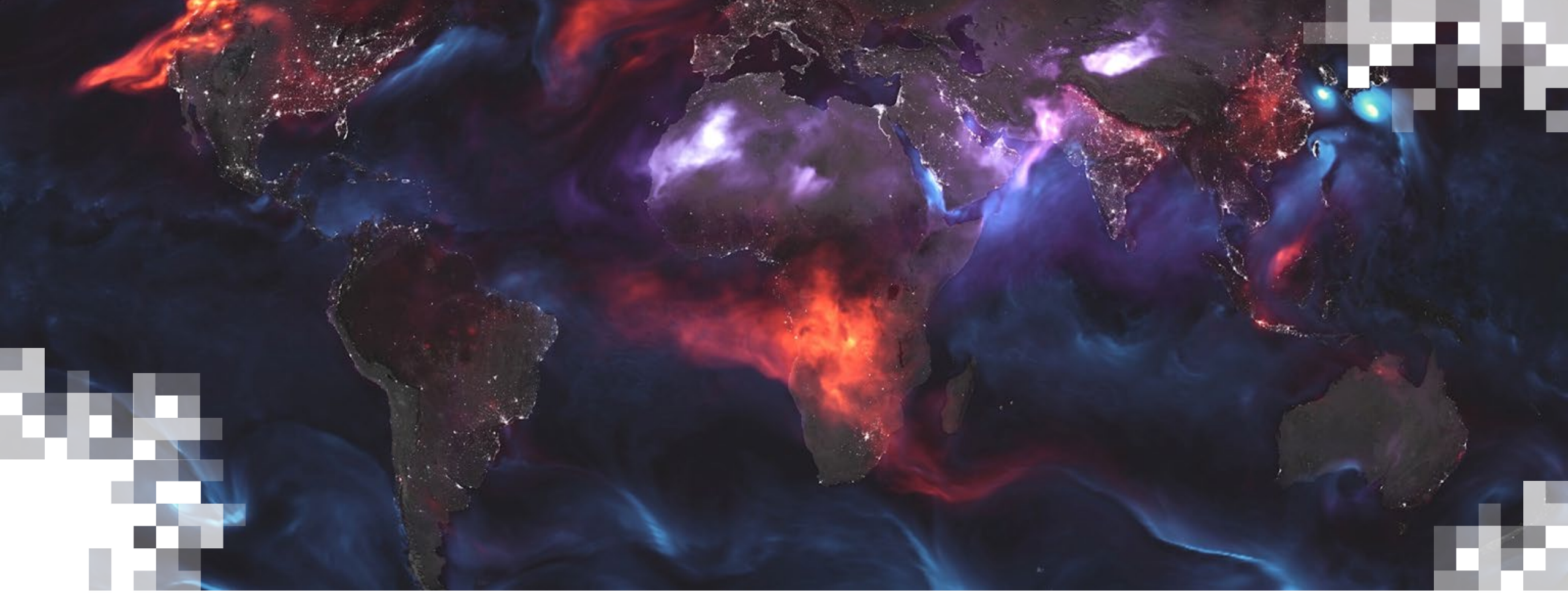
Date and time of data file

- 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

Part 2 from the Series Introduction and Access to Global Air Quality Forecasting Data and Tools

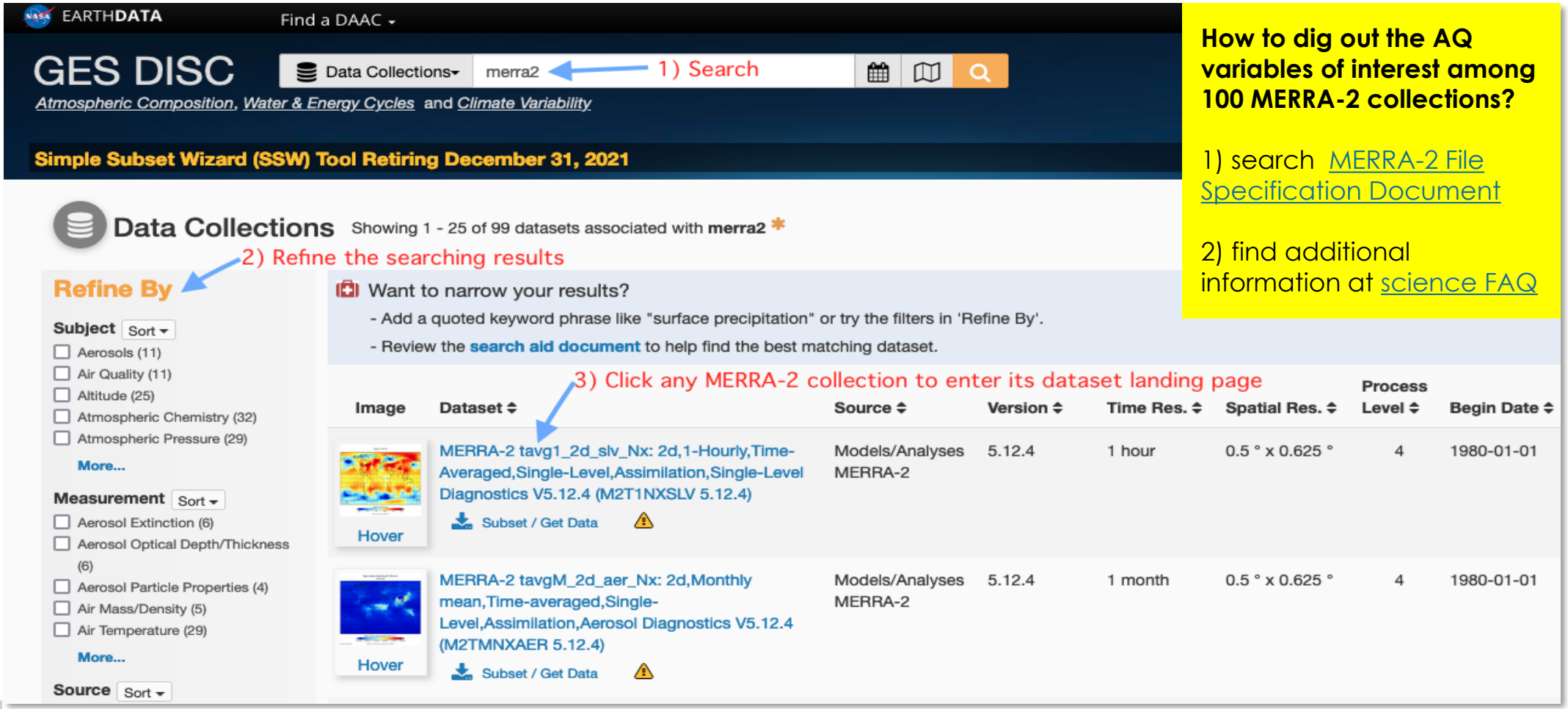
More Info: MERRA-2 Quick Guide - Data filenaming convention



Find and Access MERRA-2 AQ Data and Resources on the GES DISC Website

2. Discover and Download the MERRA-2 AQ Collections

List of MERRA-2 Collections: https://disc.gsfc.nasa.gov/datasets?keywords=merra2&page=1



How to dig out the AQ variables of interest among 100 MERRA-2 collections?

1) search MERRA-2 File Specification Document

2) find additional information at science FAQ

Example of a MERRA-2 Dataset Landing Page: The Collection of M2T1NXAER

https://disc.gsfc.nasa.gov/datasets/M2T1NXAER_5.12.4/summary?keywords=M2T1NXAER_5.12.4

1) Read the documents and resources

GES DISC | Data Collections | M2T1NXAER_5.12.4

Atmospheric Composition, Water & Energy Cycles and Climate Variability

Simple Subset Wizard (SSW) Tool Retiring December 31, 2021

HOW-TO:	How to read and plot the data.
PROJECT HOME PAGE:	The GMAO MERRA-2 documentation and FAQs
PI DOCUMENTATION:	MERRA-2 File Specification Document
READ-ME:	README Document
GENERAL DOCUMENTATION:	MERRA-2 Data Access – Quick Start Guide
IMPORTANT NOTICE:	Records of MERRA-2 Data Reprocessing and Service Changes
GENERAL DOCUMENTATION:	FAQs about MERRA-2 data access

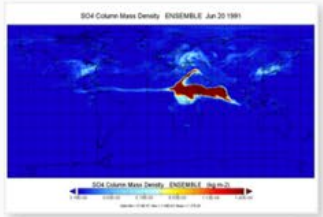
Read them

2) Search the shortname of collection of your interest

Back to search results

The second Modern-Era Retrospective analysis for Research and Applications

MERRA-2 tavg1_2d_aer_Nx: 2d,1-Hourly,Time-averaged,Single-Level,Assimilation,Aero V5.12.4 (M2T1NXAER)



View Full-size Image

M2T1NXAER (or tavg1_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 FAQs under the Documentation tab of this page. 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 Mod ...more

3) Access or download data

Data Access

- Online Archive
- Earthdata Search
- Giovanni
- Web Services
- Subset / Get Data

How to dig out the AQ variables of interest among 100 MERRA-2 collections?

1) search MERRA-2 File Specification Document

2) find additional information at science FAQ

Product Summary | Data Citation | Documentation | References

4) Cite dataset algorithm in your ppt and paper

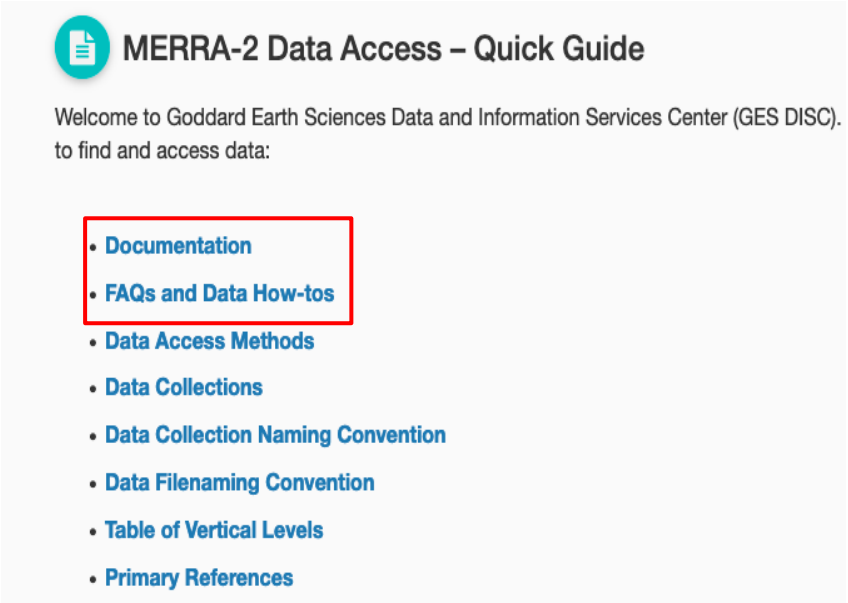
Also cite the dataset DOI in your ppt and paper!

To cite the data in publications:

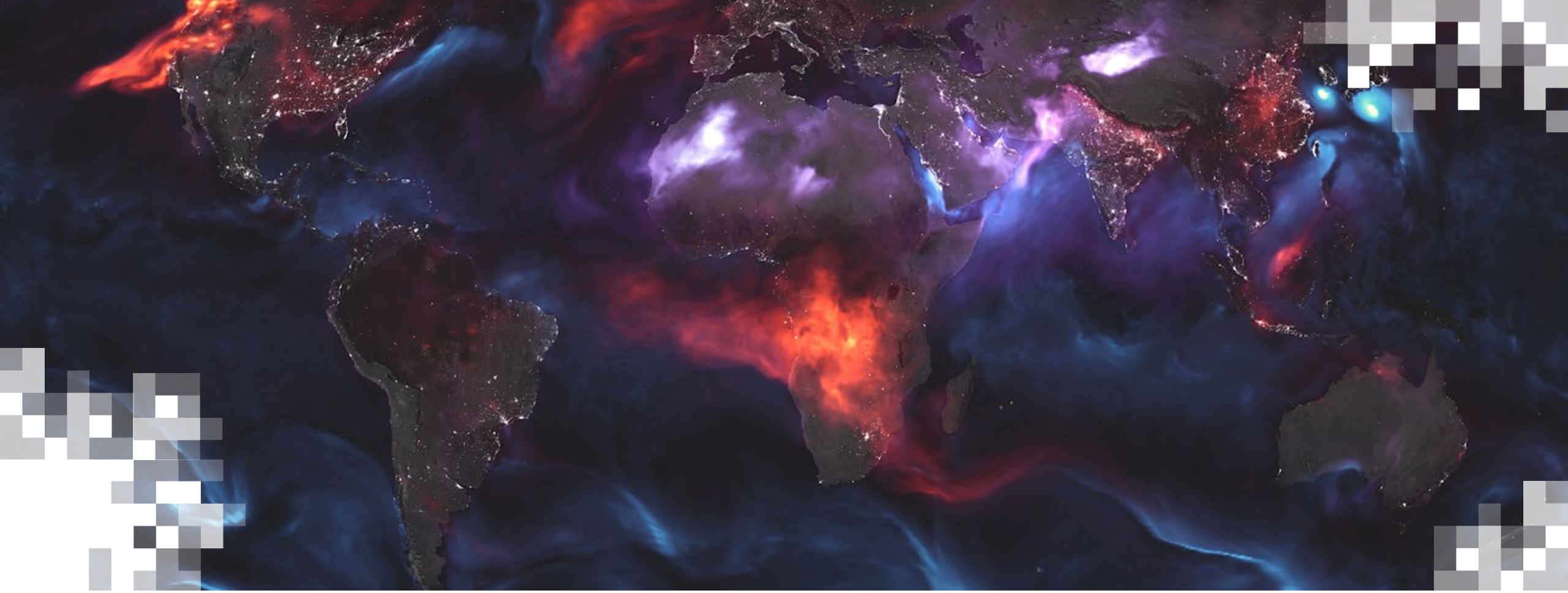
Global Modeling and Assimilation Office (GMAO) (2015), MERRA-2 tavg1_2d_aer_Nx: 2d,1-Hourly,Time-averaged,Single-Level,Assimilation,Aerosol Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [*Data Access Date*], 10.5067/KLICLTZ8EM9D

3. Find Assistance for Using the MERRA-2 Data

- Read the GES DISC Documentation: MERRA-2 Data Access – Quick Guide (nasa.gov)



- Contact Us - GES DISC Help Desk
  - When you need any assistance or to report a data access problem
  - When you want to sign up for the MERRA-2 listserv to receive announcements on the latest data information, tools and services that become available, data announcements from GMAO and more …
- Contact GMAO - merra-questions@lists.nasa.gov
  - When you have questions on science content such as the definition of a variable



Subset, Download, and Analyze MERRA-2 AQ Data Using the GES DISC Web Tool and Python API Developed by the GES DISC

4. Visualize and Analyze MERRA-2 Data with GES DISC Tools

4.1 Web tools

Above is the dataset landing page of the collection M2T1NXAER:
https://disc.gsfc.nasa.gov/datasets/M2T1NXAER_5.12.4/summary?keywords=M2T1NXAER_5.12.4

- ❑ Giovanni - Interactive visualization webtool (demonstrated later)
- ❑ **Web Services**
  - ❑ **GDS:** remotely access data with GrADS or other scripts (e.g., Python)
  - ❑ **OPENDAP:** subset and download data or remotely access the data via data tools or scripts (e.g., Panoply, IDV, Python, etc.).
  - ❑ **THREDDS:** subset, aggregate, and download data, or remotely access time series with data tools or scripts (e.g., Panoply, ArcGIS, Python, etc.).
- ❑ **Subset / Get Data** → subset, regrid, and download data, and compute daily statistics (mean, minimum, maximum) on-the-fly (demonstrated later)

4. Visualize and Analyze MERRA-2 Data with GES DISC Tools (cont.)

4.2 Python Jupyter Notebooks (access and visualize MERRA-2 data)

- How to Use the Web Services API for Subsetting MERRA-2 Data (subset and download data with a Python API) - demonstrated later from Google Colab

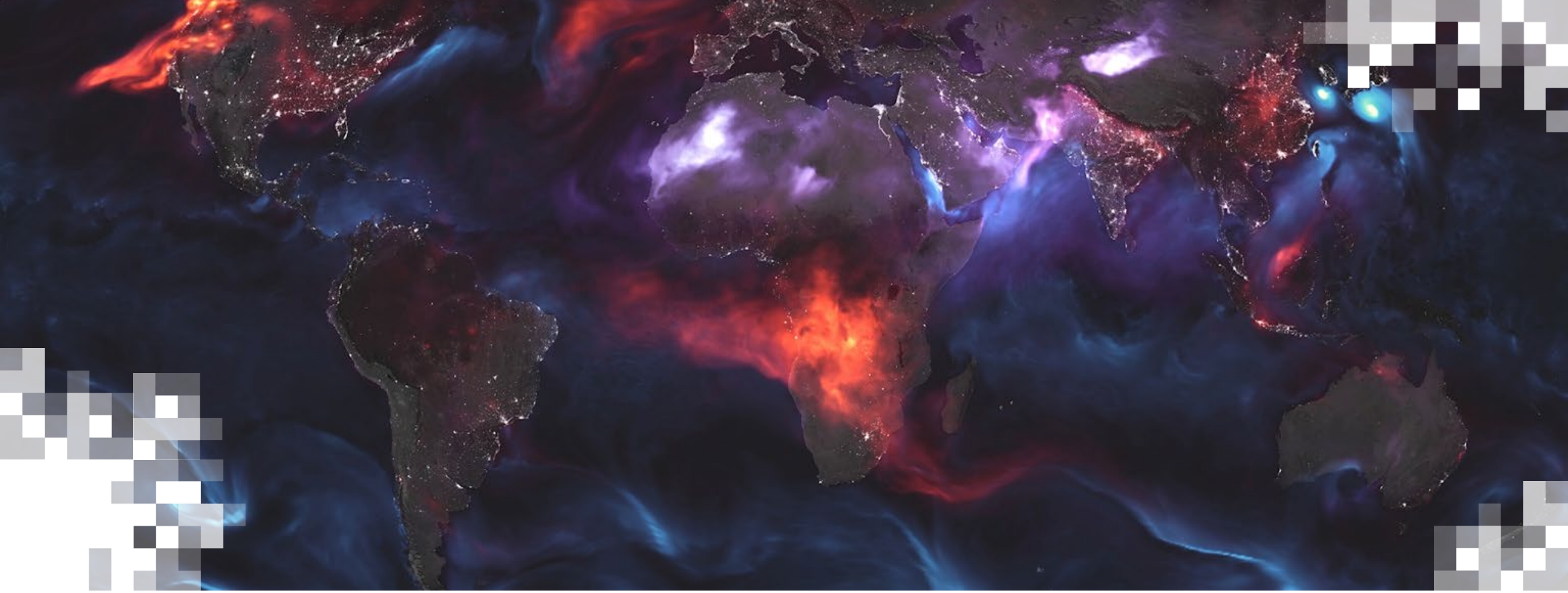
- How to remotely access *MERRA*-2 with Python3 and calculate monthly average surface PM2.5 for world countries (remotely access MERRA-2 monthly PM2.5 related variables without downloading them, and then do data analysis and visualization)

- How to Access *MERRA*-2 Data using OPeNDAP with Python3 and Calculate Daily/Weekly/Monthly Statistics from Hourly Data (remotely access MERRA-2 hourly aerosol optical depth data without downloading them, and then do data analysis and visualization)

4. Visualize and Analyze MERRA-2 Data with GES DISC Tools (cont.)

4.3 Miscellaneous Tools for Various Use Cases

- How can I subset and download a large amount of the MERRA-2 data?
- How do I obtain a time series for MERRA-2 data?
- How to convert data format from netCDF4 to ASCII?
  - How to use the Level 3 and 4 Subsetter and Regridder
  - How to Obtain a Time Series at a Single Point using TDS
  - How can I obtain the data in ASCII format?/How do I use OPeNDAP?
- How to convert data format from netCDF4 to GeoTiff?
  - Giovanni has this image output option



Demonstration: How to Subset, Download, and Analyze MERRA-2 AQ Data using the GES DISC Web Tool and Python API developed by the GES DISC

5. Demonstrate 3 Tools

a. Demo: How to use Giovanni to visualize data

(A web tool to visualize and analyze the data, no need to download data and no need to know any programming language)

b. Demo: How to use the Level 3 and 4 Subsetter and Regridder

(A web tool to subset and download data, then later a programming language is needed to visualize or analyze data)

c. Demo: How to Use the Web Services API for Subsetting *MERRA*-2 Data

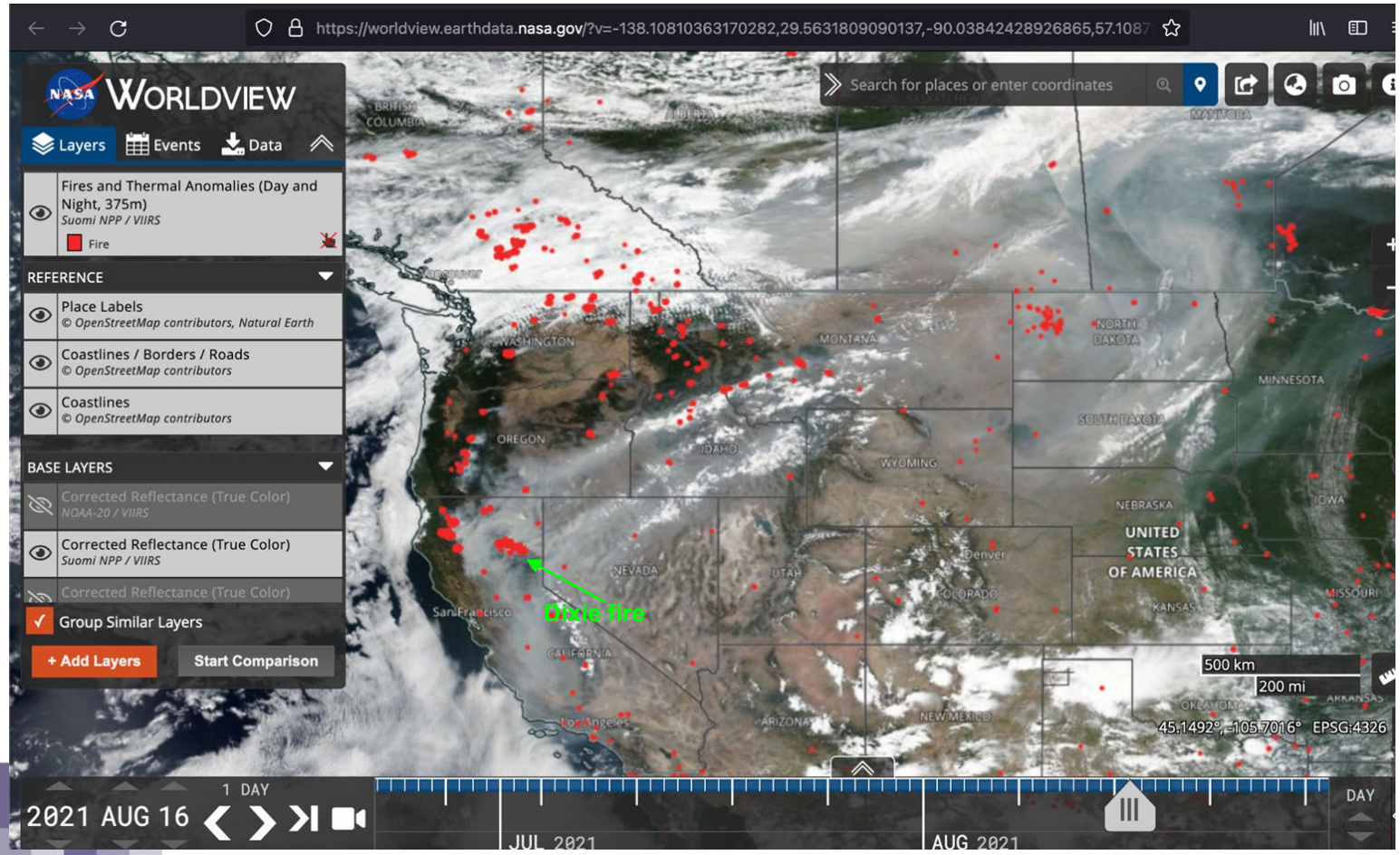
(A Python Jupyter notebook to automate the process of subsetting and downloading data, need to know Python)

Note: The difficulty of 3 demos (a→c) increases with the requirement of programming skill (zero→ high)

5. Demonstrate 3 Tools (cont.)

Case Study: Examine the impact of 2021 summer California wildfire events on air quality

Suomi NPP/VIIRS RGB and Hotspots Map on 08/16/2021



Screenshot from NASA Worldview

Science Questions:

- ❑ How severe were the impacts of 2021 summer California wildfires on air quality?
- ❑ Was it the most severe wildfire season in recent 10 years (2012-2021)?

More Info: What the numbers tells us about a catastrophic year of wildfires (the Guardian)

2021 California wildfires (Wikipedia)

Who am I?
Student Amy in Environmental Science without programming experience

What am I trying to do?
- How severe were the impacts of 2021 summer California wildfires on air quality?
- Was it the most severe wildfire season in recent 10 years (2012-2021)?

What do I need?
- Data: Monthly PM2.5 over California for the period of 2012 to 2021
- A web-based visualization tool

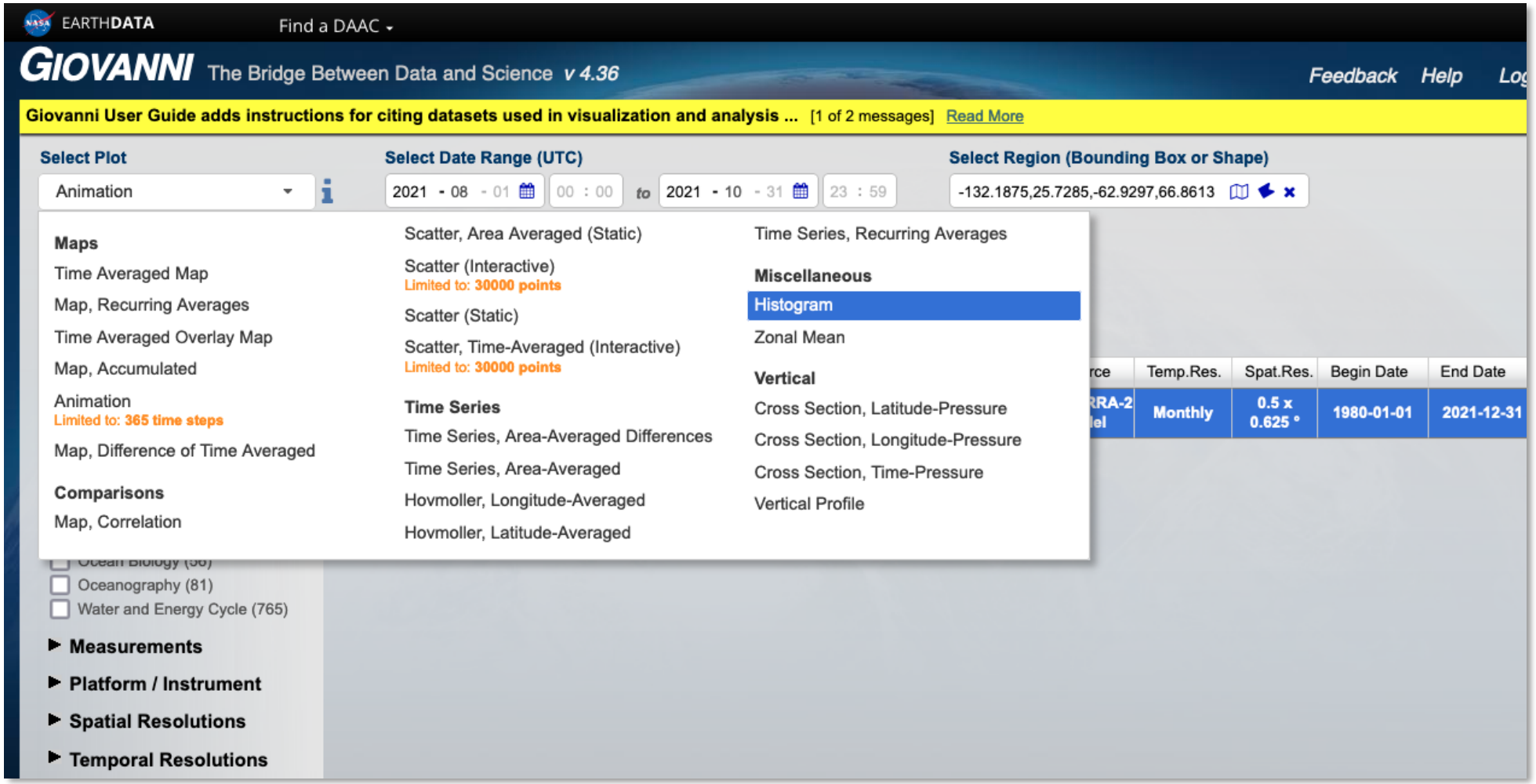
What would the "easy" button look like?
- Just click the buttons from the webpage to select variables, region, time, and the plotting type, and then the figures are displayed.

→ **Giovanni is recommended**

How do I solve the problem now?
1) Create an Earthdata account and Link GES DISC with your account to access data from Giovanni
2) Find monthly PM2.5 data from the MERRA-2 in Giovanni

5. Demonstrate 3 Tools (cont.)

a. Demo: How to use Giovanni to visualize data



5. Demonstrate 3 Tools (cont.)

a. Demo: How to use Giovanni to visualize data

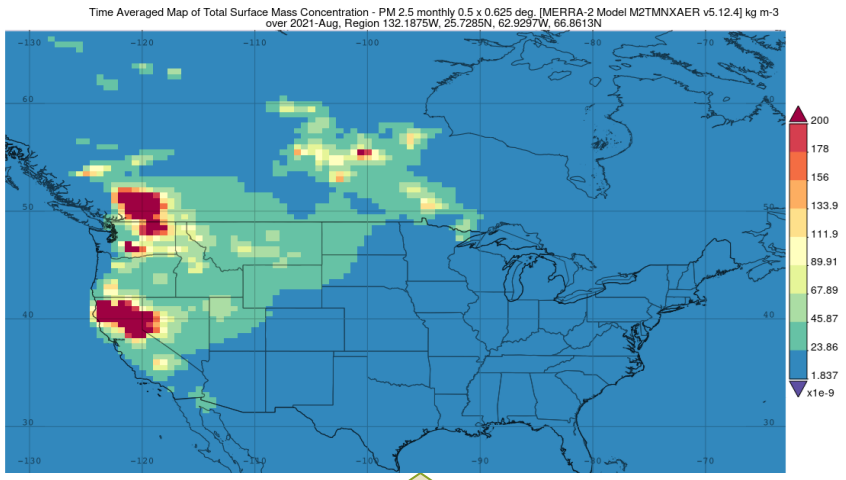
Demo Objectives:

- Show the Giovanni user interface (https://giovanni.gsfc.nasa.gov)
  - Earthdata login, Feedback
  - Plot Function, Date Range, Area of Interest, Parameter(s)
  - Search with keywords ("MERRA-2, total pm2.5, monthly") and refine the returns

- Demonstrate how to make **Map** plots
  - **Plot:** Time-averaged map
  - **Variable:** Total surface mass concentration PM2.5 (M2TMNXAER v5.12.4)
  - **Region:** North America (133°W-62°W, 25°N-67°N)
  - **Time:** August 2021
- Demonstrate how to make **Time-Series** plots
  - **Plot:** Recurring-Time-Averaged time-series
  - **Variable:** Total surface mass concentration PM2.5 (M2TMNXAER v5.12.4)
  - **Region:** California
  - **Time:** August 2012 - August 2021

Sample Plots: August Surface $PM_{2.5}$ over California for 2002-2021

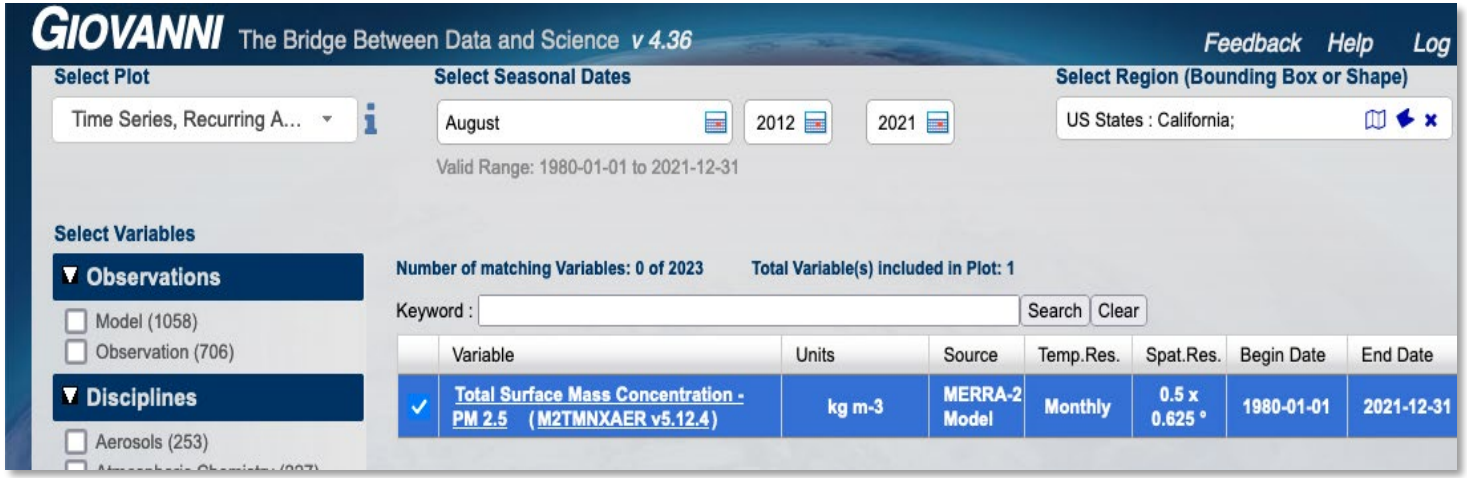
1) $PM_{2.5}$ for August 2021



Dataset: MERRA-2 → monthly $PM_{2.5}$ in collection M2TMNXAER
Visualization Tool: Giovanni
Select Plot: Map, Time averaged Map Average
Select Region: Rectangle

2) Interannual variation of August mean $PM_{2.5}$ over CA for 2012-2021

Dataset: The same
Visualization Tool: Giovanni
Select Plot: Time Series, Recurring Average
Select Region: Shapefile for CA



5. Demonstrate 3 Tools (cont.)

a. Demo: How to use Giovanni to Visualize Data

5. Demonstrate 3 Tools (cont.)

a. Demo: How to use Giovanni to Visualize Data

Giovanni How-Tos and FAQs

- How to view GPM IMERG data using **Giovanni**
- How to Obtain Data for Conducting Hurricane Case Study
- How to Create Hovmöller Plots with **Giovanni** and Panoply and Quantified ASCII Text Hovmöller Output with **Giovanni**, Panoply, and Excel
- How to calculate greenhouse gas growth rates (Download the companion **Jupyter Notebook)**
- How to Convert **Giovanni's** Mapped Image Data Format from NetCDF to ASCII Text Using Excel
- How to Create Difference Maps with **Giovanni** and Panoply: Creating Quantified ASCII Text Difference Map Output with **Giovanni**, Panoply, and Excel
- How to Create Maps with **Giovanni** and Panoply: Creating Quantified ASCII Text Map Output with **Giovanni**, Panoply, and Excel
- How can I download the **Giovanni** animation that I created?
- How can I obtain TMPA data in ASCII?
- How can I put a **Giovanni** output image into Google Earth?
- How do I acknowledge the use of **Giovanni** in a research paper? If I use **Giovanni** should I have someone on the GES DISC staff as a co-author?
- Are GPM data available in Giovanni?
- How do I subscribe to the giovanni-news email mailing list for news about Giovanni development and related scientific and educational topics delivered in newsletters, approximately quarterly ?
- Is there a way to download several months of data at one time from Giovanni?
- I waited several hours to generate a Giovanni visualization, and then it appeared that my request had failed. Why did that happen?
- What are the Recommended Guidelines for Use of Giovanni Data in Publications?
- What is Displacement Height, DISPH?
- Why are Java applications blocked by your security settings with the latest Java?
- Why are there gaps in a daily Level 3 data file?
- Why are there so many missing data points at 1000mb in MERRA-2 over my region of interest?
- Why did I get a failure message when I attempted to plot a three-dimensional parameter?

Tutorial (Video): A Clearer View of the Haze – Using NASA GES DISC Data Tools to Examine the June 2020 Sahara Dust Event over Barbados

Who am I?
Faculty Bob in atmospheric department who masters a programming language

What am I trying to do?

- How severe were the impacts of 2021 summer California wildfires on air quality?
- Was it the most severe wildfire season in recent 10 years (2012-2021)?

What do I need?

- Data: Monthly PM2.5 over California for the period of 2012 to 2021
- Visualization: Do my own visualization or analysis → adjust color bar and compare with other data not available in Giovanni.

What would the "easy" button look like?

- Subset data (e.g., variables, regions, and time range) with buttons

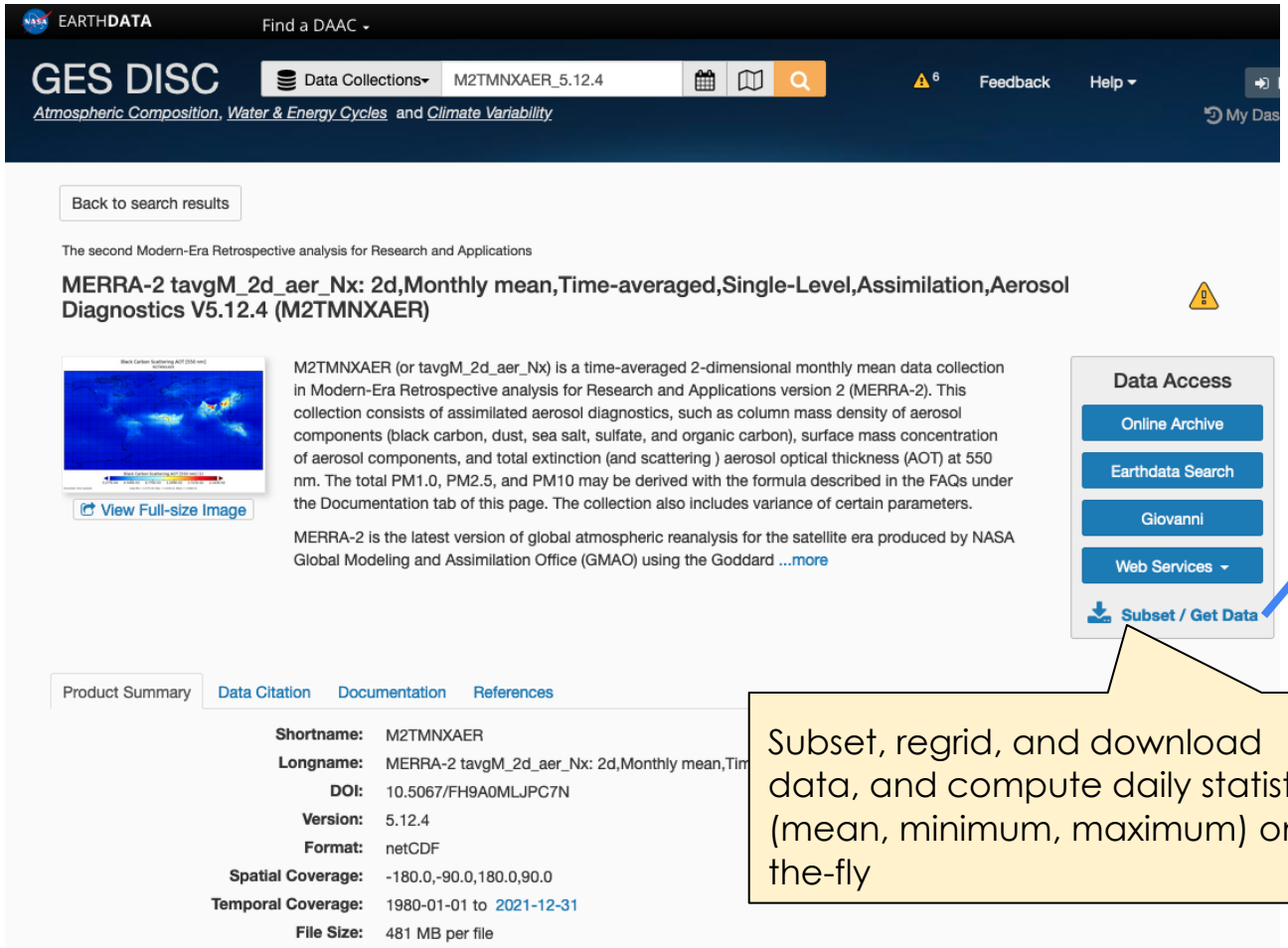
→ **The web-based tool "Level 3 and 4 Subsetter and Regridder" is recommended**

How do I solve the problem now?

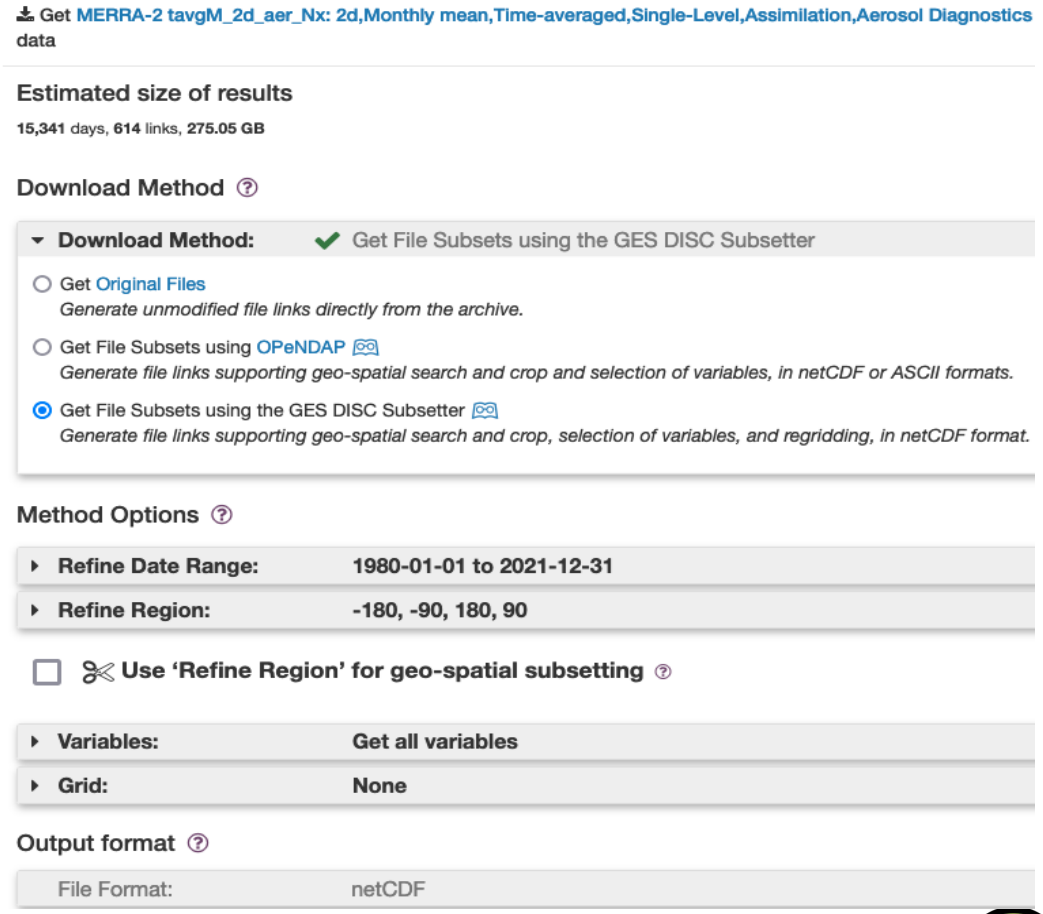
1) Create an Earthdata account and Link GES DISC with your account to access data from GES DISC website
2) Identify the MERRA-2 collection which has the monthly PM2.5 data in GES DISC website by searching the MERRA-2 File Specification Document first. In this case, it is identified as the collection M2TMNXAER.
3) Know how to calculate PM2.5 in MERRA-2 by reading science FAQ (#4)
4) Subset data via the link of **Subset / Get Data**
5) Download the subsetted files with wget or curl (How-to)
6) Visualize the downloaded data with my own script

5. Demonstrate 3 Tools (cont.)

b. Demo: How to Use the Level 3 and 4 Subsetter and Regridder



Subset, regrid, and download data, and compute daily statistics (mean, minimum, maximum) on-the-fly



The above is the dataset landing page of the collection M2TMNXAER:
https://disc.gsfc.nasa.gov/datasets/M2TMNXAER_5.12.4/summary?keywords=M2TMNXAER_5.12.4

5. Demonstrate 3 Tools (cont.)

b. Demo: How to Use the Level 3 and 4 Subsetter and Regridder

Demo Objectives:

- ❑ Identify the MERRA-2 collection which has the monthly PM2.5 data in GES DISC website by
  - ❑ Go to the GES DISC Documentation: MERRA-2 Data Access – Quick Guide (nasa.gov)
  - ❑ Searching "PM 2.5" in the MERRA-2 File Specification Document to identify the collection shortname. In this case, it is identified as the collection M2TMNXAER.
  - ❑ Knowing how to calculate PM2.5 in MERRA-2 by reading science FAQ (#4)
  - ❑ Entering the dataset landing page of M2TMNXAER
- ❑ Demonstrate how to use the web tool "Level 3 and 4 subsetter" to subset data
  - ❑ **Collection:** monthly MERRA-2 collection M2TMNXAER
  - ❑ **Variable:** OCSMASS, BCSMASS, SO4SMASS, DUSMASS25 SSSMASS25
    Note that TOTSPM25 = OCSMASS+ BCSMASS+ SO4SMASS*1.375+ DUSMASS25 + SSSMASS25.
  - ❑ **Region:** North America (133°W-62°W, 25°N-67°N)
  - ❑ **Time Range:** 2012 to 2021 (only 2021-01 to 2021-12 is demonstrated to save time)
- ❑ Demonstrate how to download "the links list" (generated after the above selections) to a Mac laptop with **wget** or **curl** commands (click "Instructions for downloading")
  - ❑ wget --content-disposition --load-cookies ~/.urs_cookies --save-cookies ~/.urs_cookies --keep-session-cookies --content-disposition -i links_list.txt
  - ❑ cat links_list.txt | tr -d '\r' | xargs -n 1 **curl** -LJO -n -c ~/.urs_cookies -b ~/.urs_cookies
- ❑ NOT demonstrate how to analyze and visualize the downloaded data

5. Demonstrate 3 Tools (cont.)

b. Demo: How to Use the Level 3 and 4 Subsetter and Regridder

Who am I?

Scientist Carson in geophysics lab who is good at Python

What am I trying to do?

- How severe were the impacts of 2021 summer California wildfires on air quality?
- Was it the most severe wildfire season in recent 10 years (2012-2021)?

What do I need?

- Data: Monthly PM2.5 over California for the period of 2012 to 2021
- Visualization: Do my own visualization or analysis → adjust color bar and compare with other data not available in Giovanni.

What would the "easy" button look like?

- No need of clicking any buttons
- Automate the subsetting (e.g., variables, regions, and time range) and downloading process with an API

→ **The Python API tool is recommended**

How do I solve the problem now?

1) Create an Earthdata account and Link GES DISC with your account to access data from GES DISC website
2) Identify the MERRA-2 collection which has the monthly PM2.5 data in GES DISC website by searching the MERRA-2 File Specification Document first. In this case, it is identified as the collection M2TMNXAER.
3) Know how to calculate PM2.5 in MERRA-2 by reading science FAQ (#4)
4) Use the Python API to subset and download data from your local machine or from any cloud platform such as Google Colab
5) Visualize the downloaded data with my own script

5. Demonstrate 3 Tools (cont.)

c. Demo: How to Use the Web Services API for Subsetting *MERRA*-2 Data

Demo Objectives:

- ❑ Identify the MERRA-2 collection which has the monthly PM2.5 data in GES DISC website by
  - ❑ Go to the GES DISC Documentation: MERRA-2 Data Access – Quick Guide (nasa.gov)
  - ❑ Searching "PM 2.5" in the MERRA-2 File Specification Document to identify the collection shortname. In this case, it is identified as the collection M2TMNXAER.
  - ❑ Knowing how to calculate PM2.5 in MERRA-2 by reading science FAQ (#4)

  Note: the procedure above is the same as that in the Demo 2 thus skipped in the Demo 3.

- ❑ Demonstrate how to run the Python API (a Jupyter notebook) from Google Colab for this training
- ❑ Demonstrate where (in the step 4-5) to modify this Python API tool to subset and download your data
  - ❑ **Collection:** monthly MERRA-2 collection M2TMNXAER
  - ❑ **Variable:** OCSMASS, BCSMASS, SO4SMASS, DUSMASS25 SSSMASS25
    Note that TOTSPM25 = OCSMASS+ BCSMASS+ SO4SMASS*1.375+ DUSMASS25 + SSSMASS25.
  - ❑ **Region:** North America (133°W-62°W, 25°N-67°N)
  - ❑ **Time Range:** 2012 to 2021(only 2021-06 to 2021-08 is demonstrated to save time)

- ❑ Demonstrate how to move the downloaded data from Google Colab to the local disk
- ❑ NOT demonstrate how to analyze and visualize the downloaded data

5. Demonstrate 3 Tools(cont.)

c. Demo: How to Use the Web Services API for Subsetting *MERRA*-2 Data

Acronyms

API	Application Programming Interface
AQ	Air Quality
CMR	Common Metadata Repository
DAAC	NASA Distributed Active Archive Center
EMS	ESDIS Metrics System
EOSDIS	NASA's Earth Observing System Data and Information System
ESDIS	NASA's Earth Science Data and Information System Project
GEOS FP	Goddard Earth Observing System Model Forward Processing
GEOS-CF	Goddard Earth Observing System Model – Composition Forecast
GES DISC	Goddard Earth Sciences Data and Information Services Center
GMAO	NASA Global Modeling and Assimilation Office
MERRA-2	Modern-Era Retrospective analysis for Research and Applications, Version 2
OGC	Open Geospatial Consortium
OPeNDAP	Open-source Project for a Network Data Access Protocol

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
  - 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


APPLIED REMOTE SENSING TRAINING PROGRAM
TRAIN - EMPOWER - ADVANCE

Thank You!