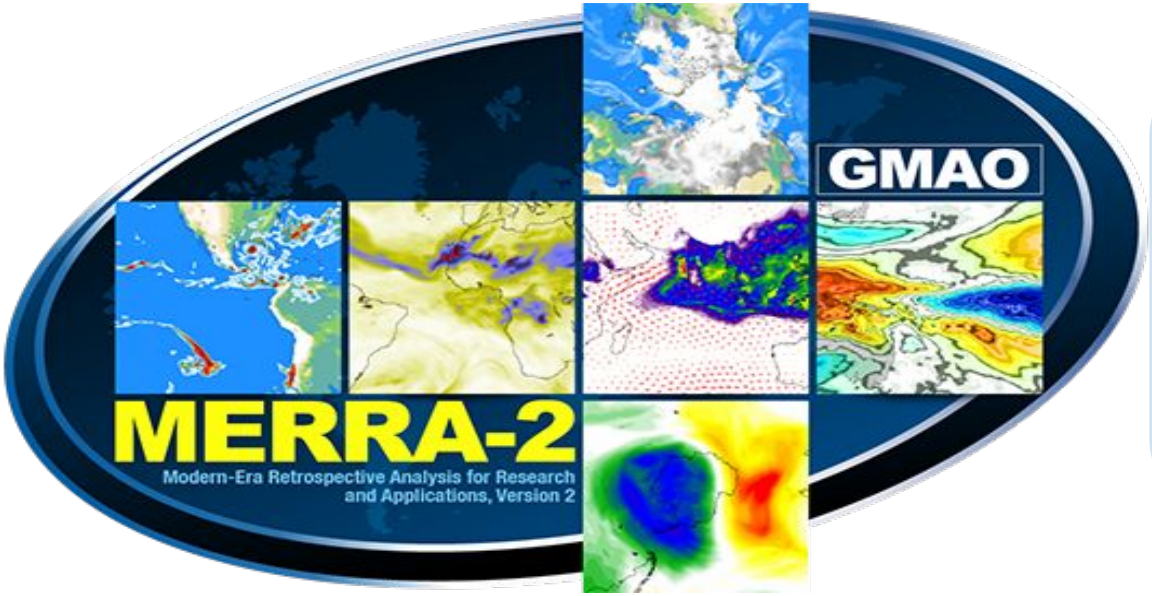
Part 1: Data Assimilation and Reanalysis



Xiaohua Pan

NASA/GSFC
Goddard Earth Sciences Data and Information Services Center
(GES DISC)

ACAM 2023 Training School

Acknowledgement: The NASA ARSET training and Arlindo da Silva at NASA/GMAO

GESDISC

GMAO: https://gmao.gsfc.nasa.gov/reanalysis/
GES DISC: https://disc.gsfc.nasa.gov/information/mission-project?title=MERRA-2

Why data assimilation?

Models

Highly spatial and temporal coverage but difficult in specifying emission, microphysical process and transport, leading to large uncertainty

+

Observations

There are a large number of observations, such as aerosol sensors, but with many blind spots

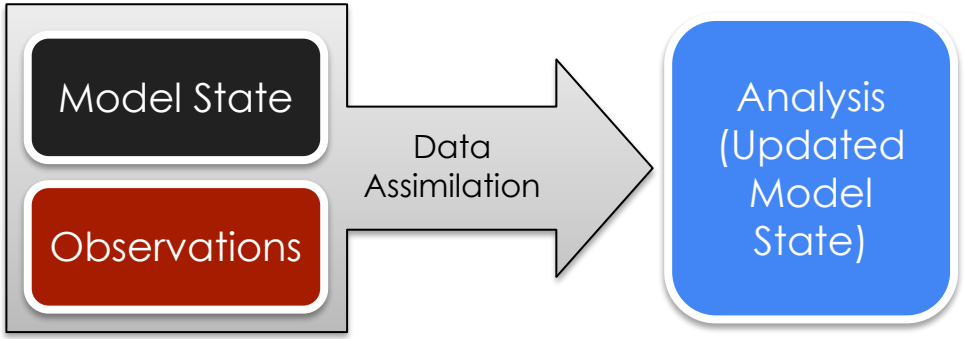
Data assimilation

Integrator of the information from models and observations and conveyor of past observations

GESDISC

Forecast, Analysis, Reanalysis, Data Assimilation How are they related to each other?

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 **forecast** is a model simulation run forward in time to predict a future state, initialized with the best estimate of current conditions using analysis.

A **reanalysis** is the same as analysis but using a single model version and consistent data assimilation techniques to produce a long-term datasets that can be used for longer term analyses

Reanalysis

What is reanalysis?

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

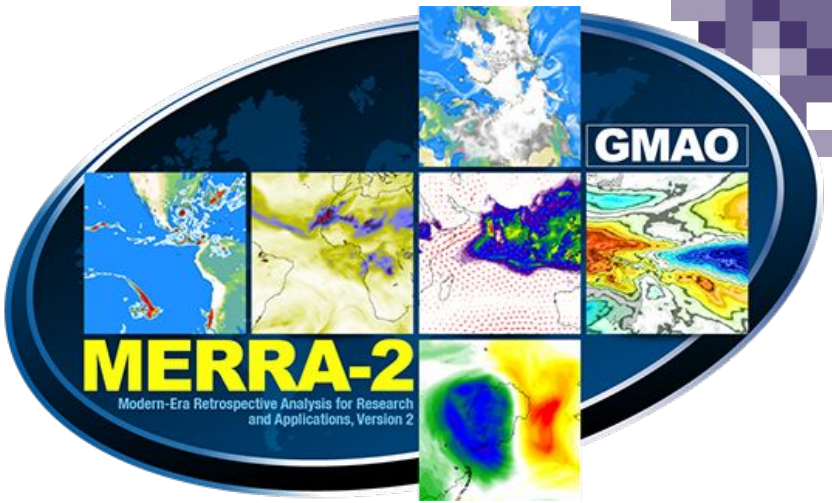
Why do we need it?

- 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, such as MERRA-2

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 reanalyses 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

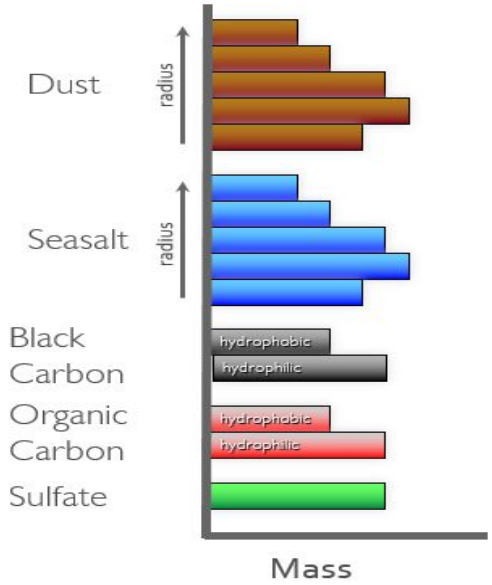


- Three-dimensional variational data analysis (3DVAR) Gridpoint Statistical Interpolation (GSI) meteorological analysis scheme (overviewed in Gelaro et al., 2017)

MERRA-2 Aerosol Model Scheme

Credit: Arlindo da Silva at NASA/GMAO

GOCART in GEOS-5



- Based on the Goddard Chemistry, Aerosol, Radiation, and Transport Model [Chin et al. 2002]
- Sources and sinks for 5 non-interactive species

DUST	wind and topographic source, 5 mass bins
SEASALT	wind driven source, 5 mass bins
BLACK CARBON	anthropogenic and wildfire source, mass hydrophic and hydrophilic
ORGANIC CARBON	anthropogenic, biogenic, and wildfire source, mass hydrophic and hydrophilic
SULFATE	anthropogenic and wildfire source of SO_2, oxidation to SO_4 mass

Nitrate not included

- Convective and large scale wet removal
- Dry deposition (and sedimentation for dust and sea salt)
- Optics based primarily on OPAC

- Aerosols are coupled to the meteorological reanalysis (both radiatively and through emissions/loss processes)

GESDISC

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.
- The MERRA-2 aerosol analysis uses the Goddard Aerosol Assimilation System (GAAS). Every 3 h, this system assimilates quality-controlled AOD at 550 nm into the GEOS-5/GOCART modeling system.
- 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.

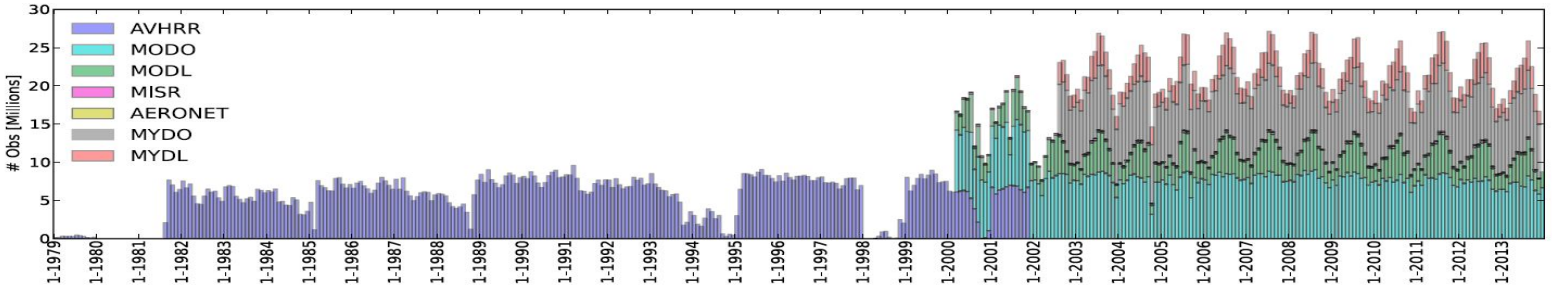


Figure 3 from Randles et al. 2017

MERRA-2 Global Mean AOD Reanalysis

- Aerosol reanalysis for the modern satellite era (1980-onward)
- This is the first multidecadal reanalysis within which meteorological and aerosol observations are jointly assimilated into a global assimilation system

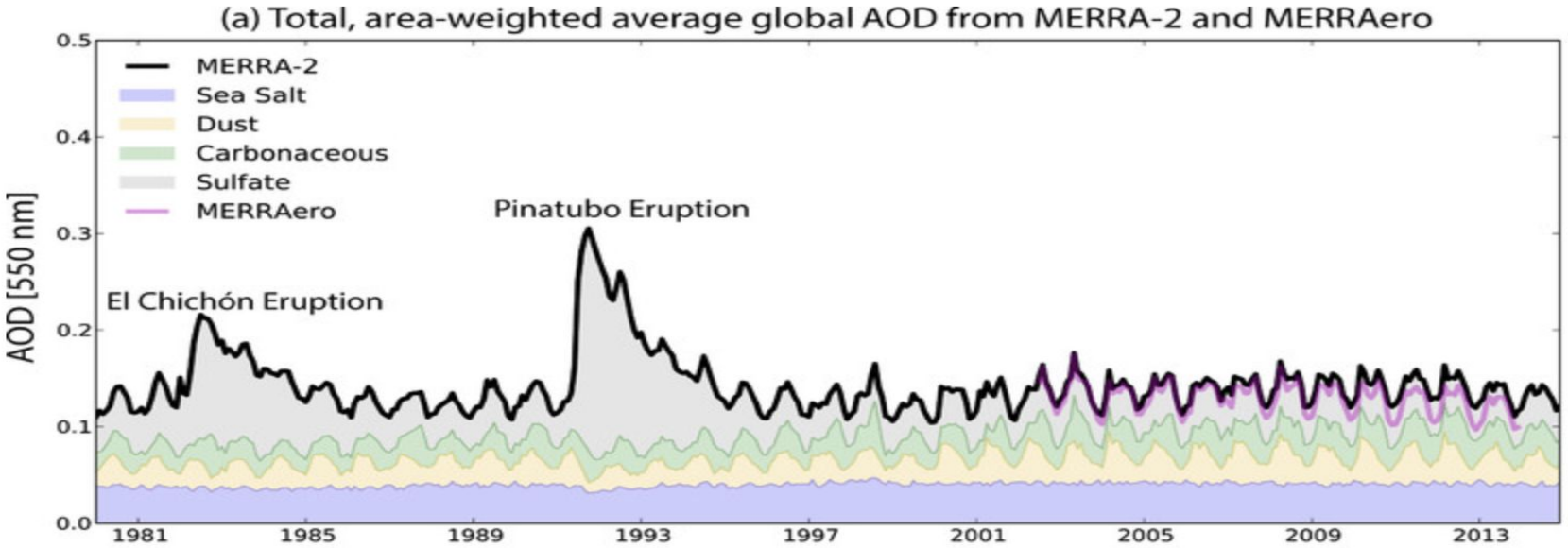


Figure 5a from Randles et al. 2017

Overview of MERRA- 2

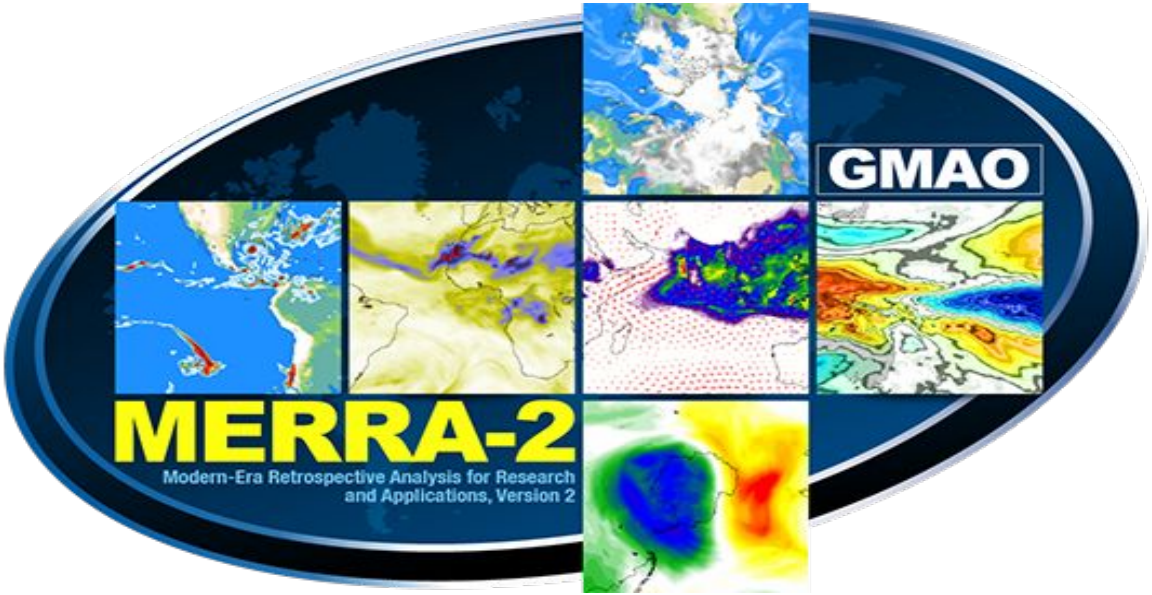
- Generated by NASA Global Modeling and Assimilation Office (GMAO)
- Archived and distributed at NASA GES DISC
- ***Total of 99 collections*** *(95 standard collections, 4 derived climate statistical collections) available in both the on-prem archive and the AWS S3 buckets (list).*
- *One value-added collection monthly $PM_{2.5}$ at country level available in on-prem archive only (link)*

- 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 project page at the GES DISC

AWS: Amazon Web Services. **S3:** Simple Storage Service.

Part 2: Discover and Access MERRA-2 Air Quality Data



Xiaohua Pan

NASA/GSFC
Goddard Earth Sciences Data and Information Services Center
(GES DISC)

ACAM 2023 Training School

GMAO: https://gmao.gsfc.nasa.gov/reanalysis/
GES DISC: https://disc.gsfc.nasa.gov/information/mission-project?title=MERRA-2

1. 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

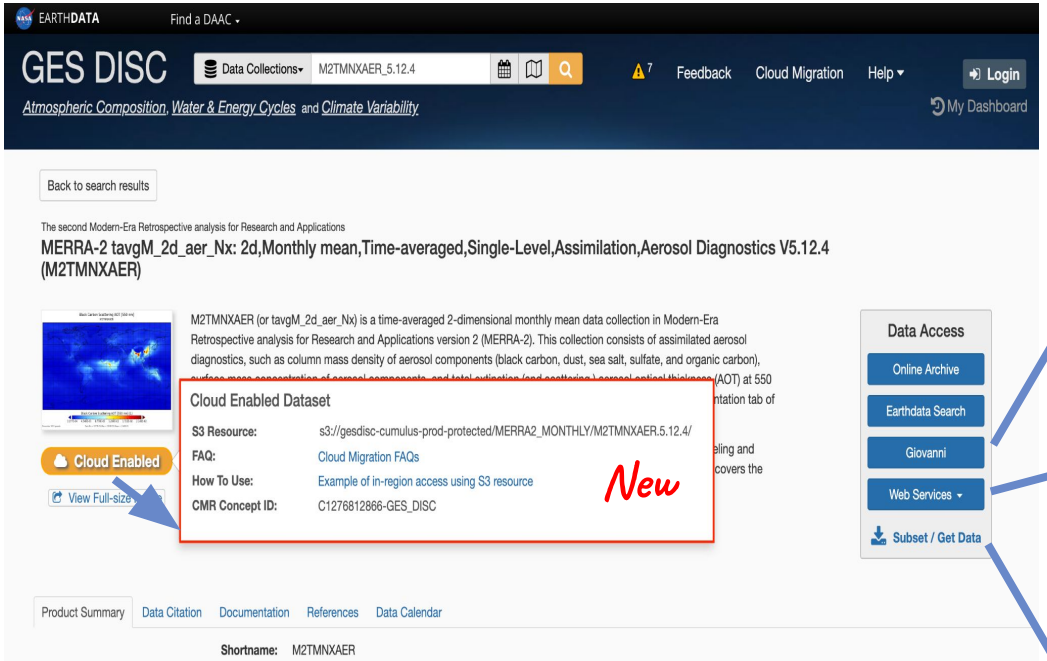
& Surface $PM_{2.5}$ can be derived by variables listed in that collection: TOTSPM25 = OCSMASS+ BCSMASS+ SO4SMASS*1.375+ DUSMASS25 + SSSMASS25

@ $PM_{1.0}$ 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).

2. Search the MERRA-2 collection and variables

3. Access MERRA-2 Data



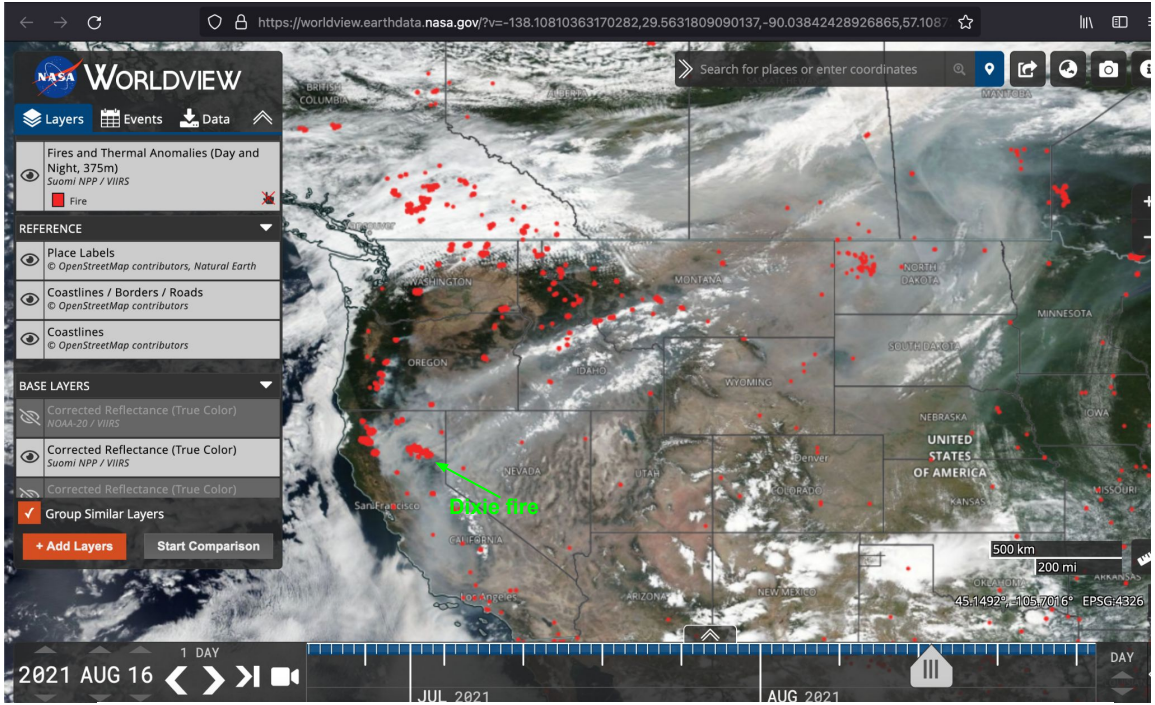
The Dataset Landing Page of the Collection M2TMNXAER

- 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

4. Demonstrate data tools

4.1 Demo: How to use Worldview to visualize severe air quality events

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



Screenshot from NASA Worldview

Science Questions:

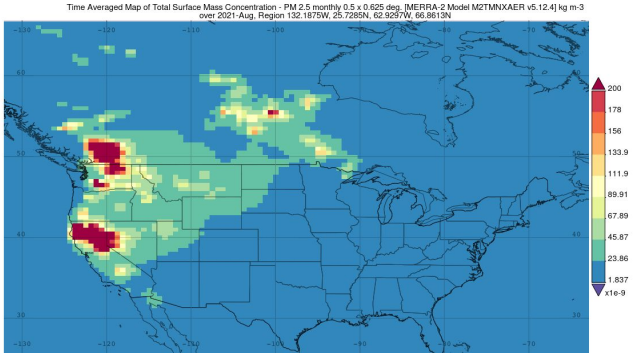
- ❑ How severe were the impacts of 2021 summer California wildfires on air quality?
- ❑ Was it the most severe wildfire season in recent 10 years (2012-2021)?

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

2021 California wildfires (Wikipedia)

4.2 Demo: How to use Giovanni to visualize MERRA-2 PM2.5

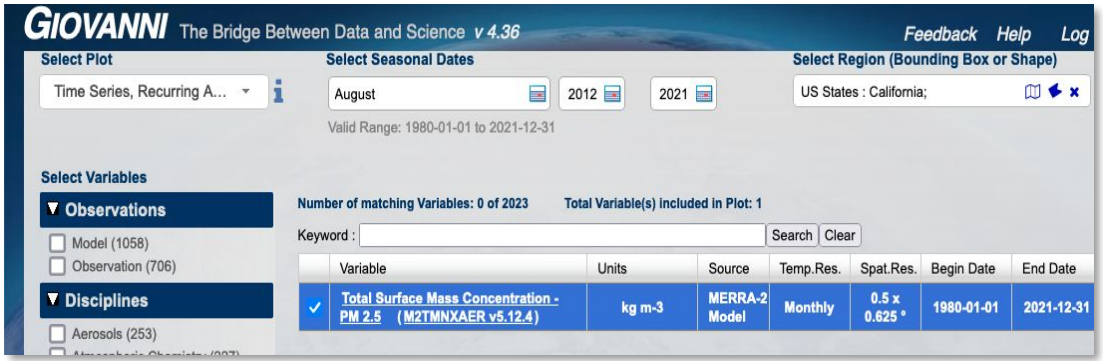
1) $PM_{2.5}$ for August 2021



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

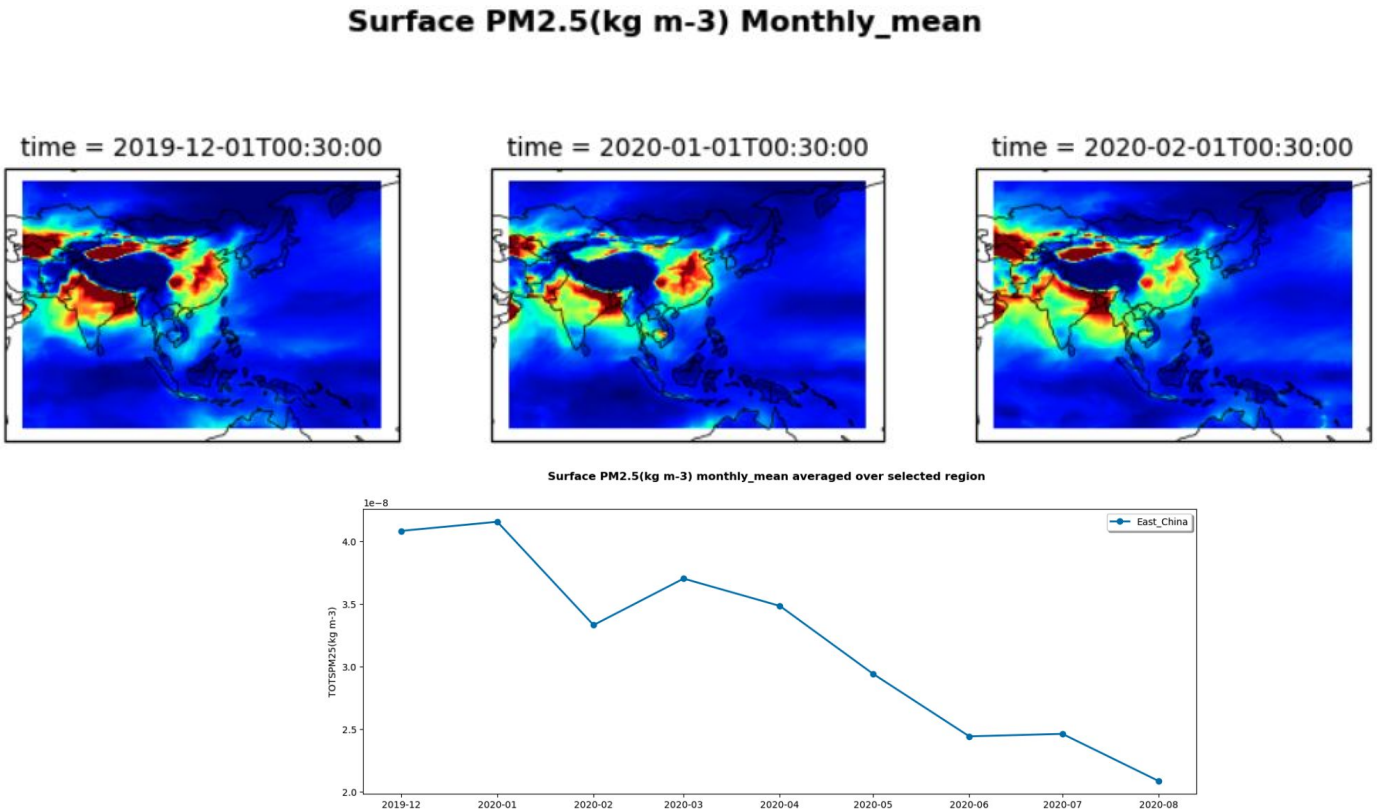
2) Interannual variation of August mean $PM_{2.5}$ over CA for 2012-2021

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



4. Demonstrate data tools

4.2 Demo: How to use Python notebook to visualize MERRA-2 $PM_{2.5}$



5. Find Assistance for Using the MERRA-2 Data

- Read the MERRA-2 project page at the GES DISC
- 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

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

- 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
  - Gelaro et al.,. 2017. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Journal of Climate. Vol. 30, No. 14, pp. 5419-5454. DOI: 10.1175/JCLI-D-16-0758.1 ISSN: 0894-8755. https://journals.ametsoc.org/view/journals/clim/30/14/jcli-d-16-0758.1.xml
  - 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

Additional slides

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

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)

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

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

More info, MERRA-2 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