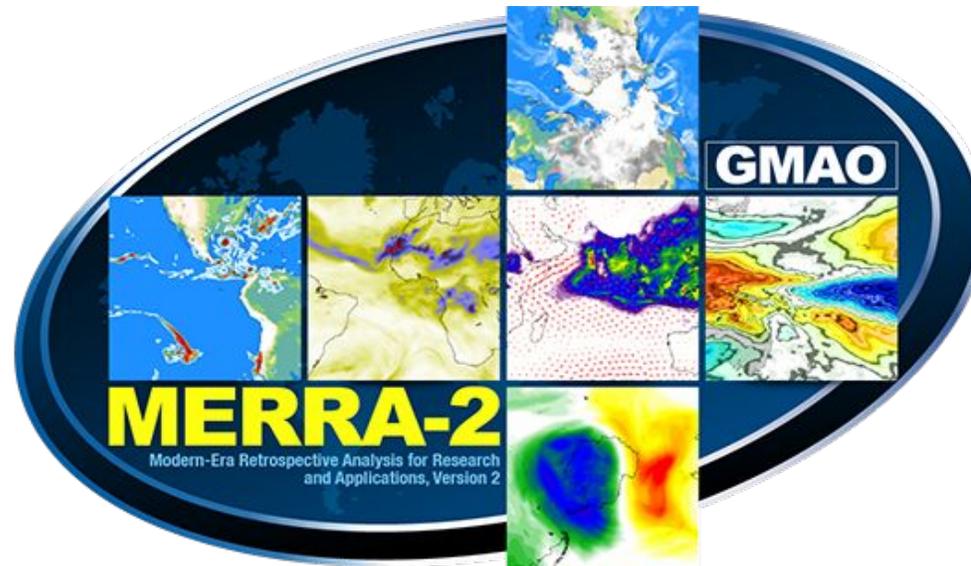


The NASA MERRA-2 Reanalysis Products: Data and Tools Used for Aerosol and Air Quality Studies

Xiaohua Pan, Suhung Shen, Jennifer Wei, Irina Gerasimov, Michael Bosilovich, Barbara Deshong, Lena Iredell, David Meyer

NASA/GSFC/GES DISC

Motivation: Accelerate your science with right data and new technology under Open-source science initiative (NASA).





Introduction: MERRA-2 at GES DISC

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

GMAO Webpage: <https://gmao.gsfc.nasa.gov/reanalysis/>

Data Access Page: <https://disc.gsfc.nasa.gov/datasets?project=MERRA-2>



Key MERRA-2 Collections and Variables Useful for Aerosol and 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} & PM₁₀	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
	M2I3NVASM 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:

$$\text{TOTSPM25} = \text{OCSMASS} + \text{BCSMASS} + \text{SO4SMASS} * 1.375 + \text{DUSMASS25} + \text{SSSMASS25}.$$

@ PM_{1.0} and PM₁₀ the formula can be found at the MERRA-2 project [FAQ](#).

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



Visualize and Analyze MERRA-2 Data with GES DISC Tools

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)
- ❑ [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)





Visualize and Analyze MERRA-2 Data with GES DISC Tools

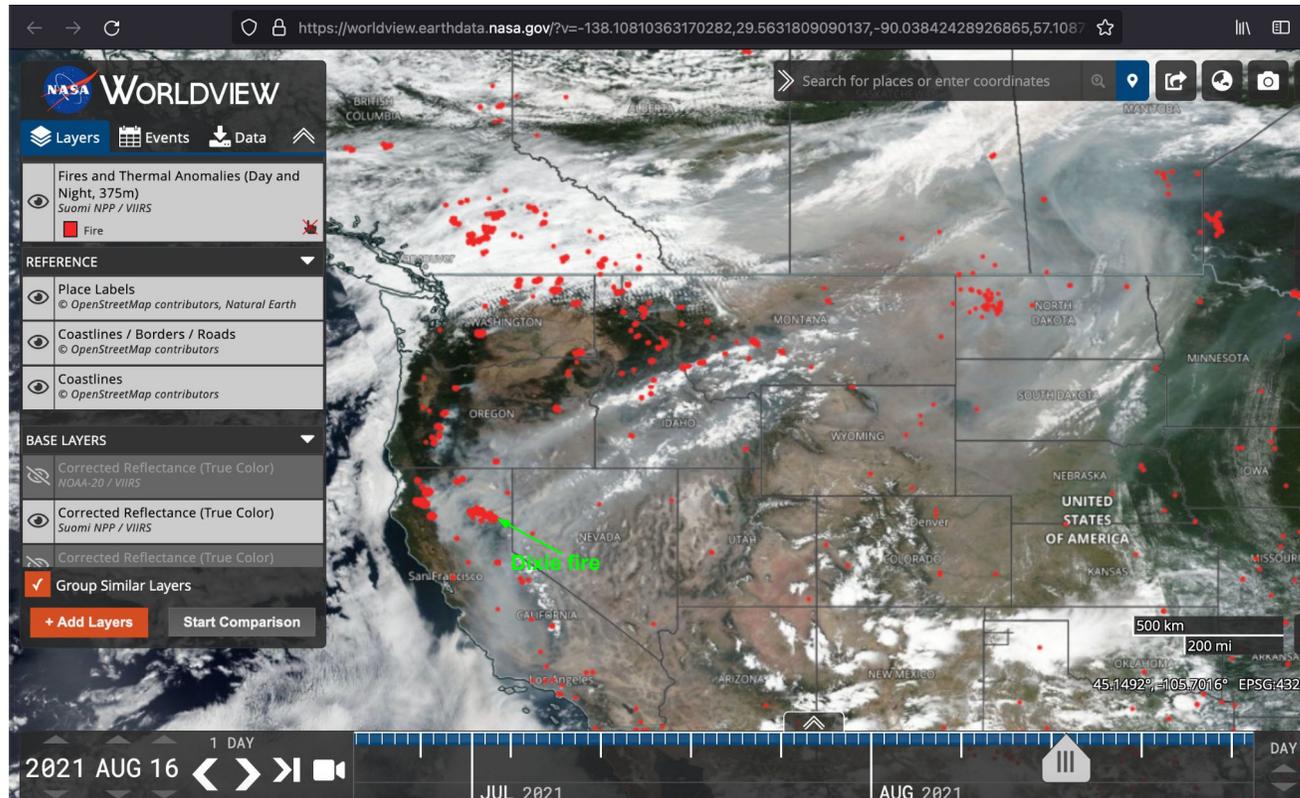
Popular Data Tools and Recipes for Various Use Cases

- ❑ [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 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?](#)



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

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



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)

Screenshot from NASA [Worldview](#)

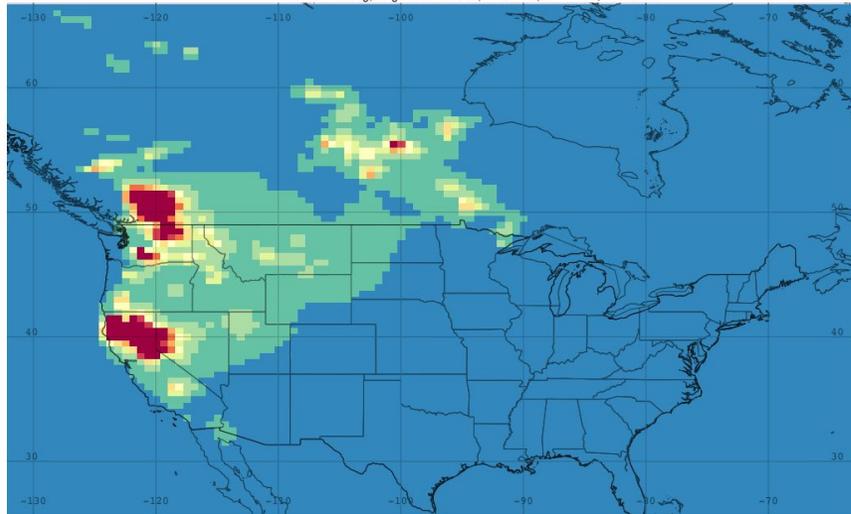




Sample Plots: August Surface PM_{2.5} over California for 2002-2021

1) PM_{2.5} for August 2021

Time Averaged Map of Total Surface Mass Concentration - PM 2.5 monthly 0.5 x 0.625 deg. [MERRA-2 Model M2TMNXAER v5.12.4] kg m-3 over 2021-Aug, Region 132.1875W, 25.7285N, 62.9297W, 66.8613N



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

Interannual Time Series

Average Total Surface Mass Concentration - PM 2.5 monthly 0.5 x 0.625 deg. [MERRA-2 Model M2TMNXAER v5.12.4] kg m-3 for 2012-Aug - 2021-Aug, Shape California



August



GIOVANNI The Bridge Between Data and Science v 4.36 Feedback Help Log

Select Plot Time Series, Recurring A... **Select Seasonal Dates** August 2012 2021 **Select Region (Bounding Box or Shape)** US States : California;

Valid Range: 1980-01-01 to 2021-12-31

Select Variables Number of matching Variables: 0 of 2023 Total Variable(s) included in Plot: 1

Keyword: Search Clear

Variable	Units	Source	Temp.Res.	Spat.Res.	Begin Date	End Date
<input checked="" type="checkbox"/> Total Surface Mass Concentration - PM 2.5 (M2TMNXAER v5.12.4)	kg m-3	MERRA-2 Model	Monthly	0.5 x 0.625 °	1980-01-01	2021-12-31

Observations
 Model (1058)
 Observation (706)

Disciplines
 Aerosols (253)
 Atmospheric Chemistry (297)



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?





MERRA-2 data have been migrated to AWS Cloud

- ❑ Benefits: Analyzing all NASA data in the AWS (Amazon Web Service) cloud without downloading data
- ❑ NASA EARTHDATA Cloud: [MERRA-2](#)
- ❑ 11 MERRA-2 collections are in AWS west-2 region as of Sep 20, 2022
- ❑ GES DISC URL for S3 (Simple Storage Service) access is:
e.g., `s3://gesdisc-cumulus-prod-protected/MERRA2/M2T1NXSLV.5.12.4/`
- ❑ Cloud dashboard (Migration Status, FAQ, How-To, and more): [Link](#)

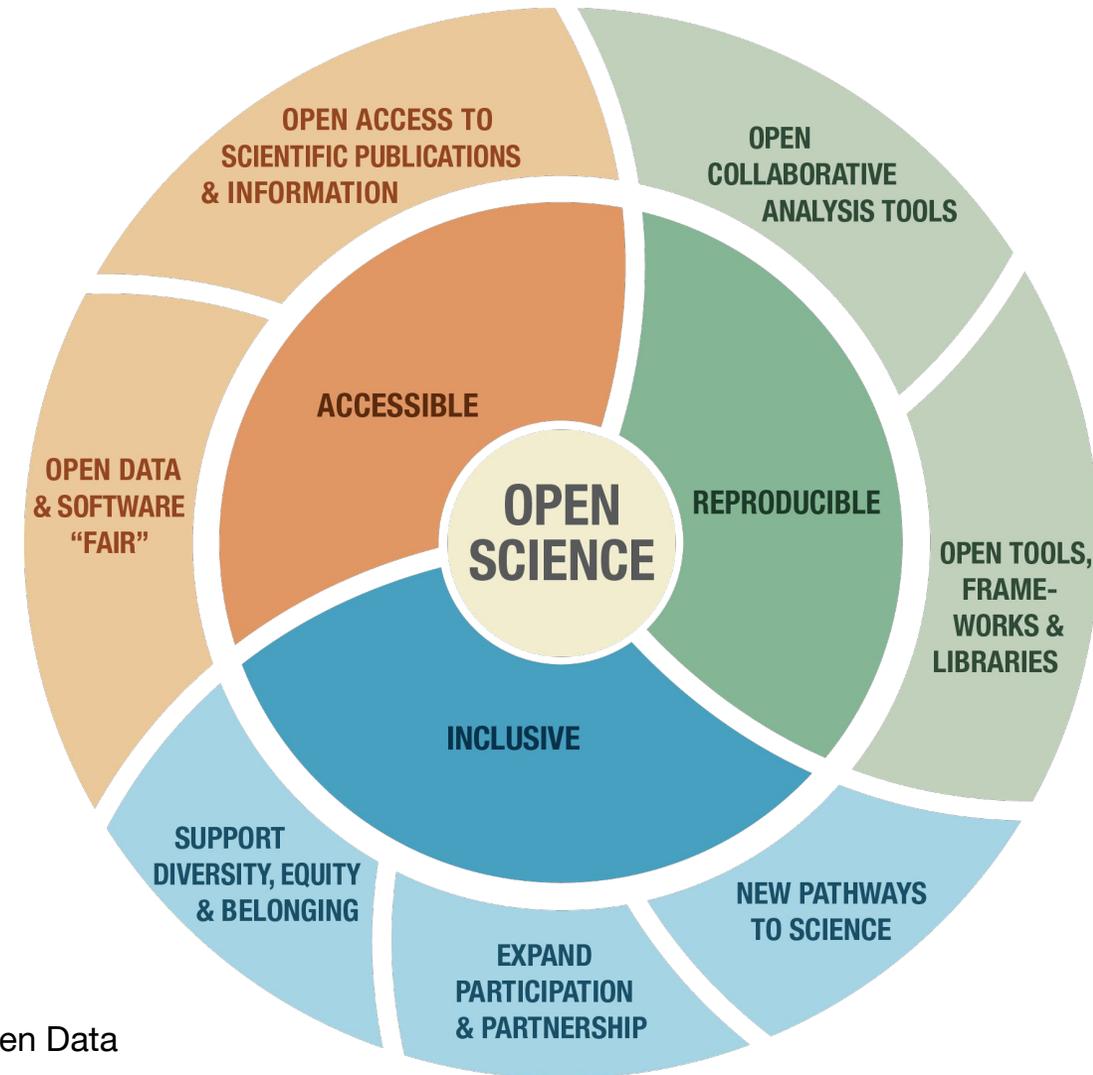


Image Credit:

<https://www.earthdata.nasa.gov/learn/webinars-and-tutorials/how-cloud-earth-scientists>

What is Open Science?

A **collaborative culture** enabled by **technology** that empowers the **open sharing of data, information, and knowledge** within the **scientific community and the wider public** to accelerate scientific research and understanding.



Ramachandran, R., Bugbee, K., & Murphy, K. J. Moving from Open Data to Open Science. Earth and Space Science, Wiley Publication
<https://doi.org/10.1029/2020EA001562>



A NASA OPEN-SOURCE SCIENCE INITIATIVE: **TOPS: TRANSFORM TO OPEN SCIENCE**

12

POC: Dr. Chelle Gentemann, TOPS Program Officer,
CSDO



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., M2**T**M**N**X**A**E**R**

M2**Time****Frequency****Dimensions****Group**

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)

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 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

