



# Data, Application, and User Need: the NASA MERRA-2 Reanalysis Products

Xiaohua Pan<sup>1,2</sup>, Suhung Shen<sup>1,3</sup>, Jennifer Wei<sup>1</sup>, Irina Gerasimov<sup>1,2</sup>, Michael Bosilovich<sup>1</sup>, Barbara Deshong<sup>1,2</sup>, Mary Greene<sup>1,4</sup>, Lena Iredell<sup>1,2</sup>, David Meyer<sup>1</sup>

<sup>1</sup>GES DISC, NASA Goddard Space Flight Center, Greenbelt, MD, USA <sup>2</sup>ADNET Systems Inc., Lanham, MD, USA <sup>3</sup>George Mason University, Fairfax, VA, USA

<sup>4</sup>Telophase Corp., Arlington, VA, USA **Contact:** [xiaohua.pan@nasa.gov](mailto:xiaohua.pan@nasa.gov)



NASA/Goddard EARTH SCIENCES DATA and INFORMATION SERVICES CENTER (GES DISC), website: <https://disc.gsfc.nasa.gov/>. Contact us at [gsfc-dl-help-disc@mail.nasa.gov](mailto:gsfc-dl-help-disc@mail.nasa.gov)

## 1. Introduction

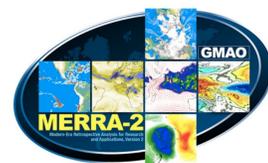
The NASA Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) is reanalysis data spanning from 1980 to the present. It is produced by the NASA Global Modeling and Assimilation Office (GMAO) and distributed by the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). MERRA-2 data has 100 collections, including Earth system variables (mainly from an atmospheric model) such as meteorological fields, radiation fields, and aerosol fields. MERRA-2 data has been one of the most popular datasets from NASA that is widely used in interdisciplinary research and applications, with a consistently increasing number of users. For example, there were more than 7000 unique users accessing the MERRA-2 data at GES DISC in the year 2021.

The GES DISC curates and distributes data and provides various data services to help users accelerate scientific discovery, e.g., by answering users' questions (tickets), providing data service, recipes, and tutorials, and analyzing the data distribution metrics and data usage. In this study, we will also update the status of the cloud migration of the MERRA-2 data to Amazon Web Services (AWS).

## 2. MERRA-2 Data

### General features:

- Generated by NASA GMAO (<https://gmao.gsfc.nasa.gov/reanalysis/>)
- Archived and curated at NASA GES DISC (<https://disc.gsfc.nasa.gov>)
- Model Version: 5.12.4
- Total of 100 collections (95 standard collections, 4 derived climate statistical collections, and 1 value-added collection - monthly PM<sub>2.5</sub> at country level)
- 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 7,000 in 2021



### New products:

- Monthly PM<sub>2.5</sub> at country level (read [data release news](#))
- Four version 2 of monthly statistical collections based on 1991-2020 climatology (read [data release news](#))

### Reprocessed products:

- Data during June - September 2021 (see full list [here](#))

### Products in the AWS Cloud:

- 11 MERRA-2 collections are in AWS US west-2 region as of November 30, 2022. [Earthdata Cloud](#)

### Recommendation to users:

- Please cite the dataset DOI when using MERRA-2 data (see example shown below). We collect and promote your paper in our publication database



Global Modeling and Assimilation Office (GMAO) (2022), MERRA-2 statM\_2d\_ed1\_Nx: 2d, Single-Level, Monthly Extremes Detection Indices based on 1991-2020 V2, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [Data Access Date], 10.5067/O8AX56D060MI

## 3. GES DISC Services and Support

- Value-added services for data
  - Giovanni
  - Subsetting, reformatting, and regridding
  - Access protocols (e.g., OPeNDAP)
  - [Cloud dashboard](#)
- Multi-tiered user support
  - GES DISC Help Desk: Contact us at [gsfc-dl-help-disc@mail.nasa.gov](mailto:gsfc-dl-help-disc@mail.nasa.gov)
  - Forum: [forum.earthdata.nasa.gov](https://forum.earthdata.nasa.gov)
  - Social Media Platform (e.g., Twitter @NASAEarthdata, and Youtube [NASA Earthdata channel](#))



## 4. Data Usage in Applied Science

Articles in Various Applied Science Areas Mentioning MERRA-2 during Jan 2017- Nov 2022 (Source: Google Scholar)

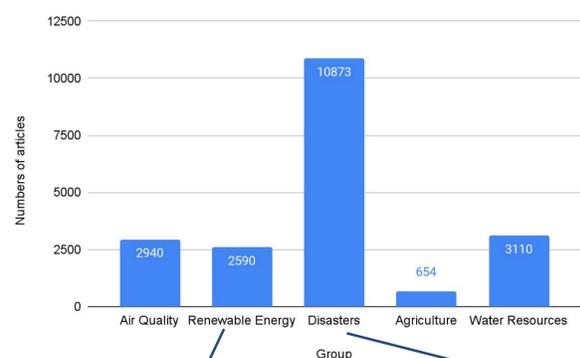
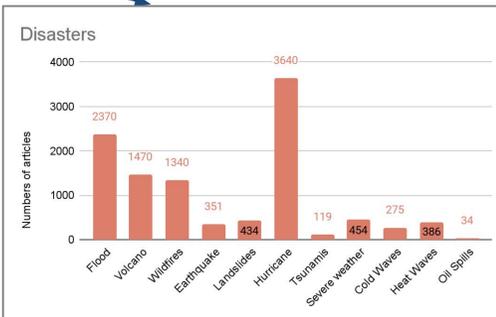
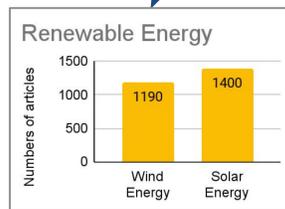


Chart 1 Numbers of articles in various applied science areas mentioning the MERRA-2 data found in Google Scholar search dated from January 2017 to November 2022 (Upper panel); Numbers of articles for sub-sectors in renewable energy (lower left panel); Numbers of articles for sub-sectors in disasters (lower right panel).



### Summary:

Google Scholar advanced search was used to find MERRA-2 relevant articles published between 2017 and 2022. With the keywords as "MERRA-2" OR "MERRA2", we found 12,800 articles. To further gain insight into which applied sciences MERRA-2 data are associated with, additional keywords were added to search queries, such as "flood", "hurricane", etc. The above charts show the statistics for each applied science area. Note that the results may contain peer-reviewed and non-peer-reviewed articles along with meeting abstracts, as well as duplicates which likely occurred in overlapping areas, e.g., "Disasters" and "Air Quality".

### Applied science areas:

- 10,873 results are related to "Disasters"
  - 3,640 are associated with "Hurricane" application
  - 2,370 are related to "Flood"
  - 1,470 are associated with "Volcano"
- 2,940 results are related to "Air Quality"
- 3,100 results are related to "Water Resources" (or "drought")
- 2590 results are related to "Renewable Energy"
- 654 results are related to "Agriculture" (or "food security")

## 5. User Needs for Data Services in FAIR Principle

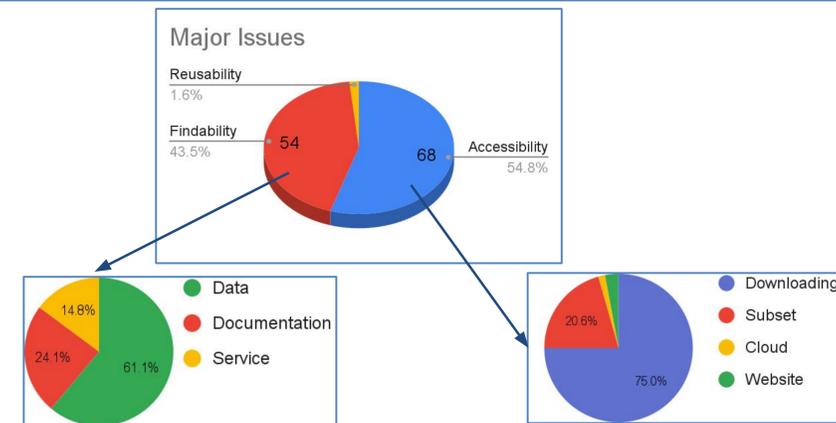


Chart2. The percentage and number of various issues reported in the 124 MERRA-2 tickets during the year 2021

### Summary:

GES DISC dealt with a total of 124 MERRA-2 user assistance tickets in the year 2021:

- 54.8% (68) are related to Accessibility
  - Downloading (51)
  - Subset (14)
  - Cloud (1)
  - Website (2)
- 43.5% (54) are related to Findability
  - Data (32)
  - Documentation (14)
  - Service (8)
- 1.6% (2) are related to Reusability

### Definitions:

- Accessibility:** Users know where the desired data or service is, but failed to access.
- Findability:** Users did not know where the desired data, documentation (including data recipes), or services are
- Reusability:** Users requested to redistribute a subset of MERRA-2 data for a certain application, such as renewable energy.

MERRA-2 is one of the most distributed datasets at GES DISC. There were more than 7,000 unique users accessing MERRA-2 data at GES DISC in the year 2021. 54.8% of 124 user assistance tickets in 2021 are associated with access issues (downloading issues are dominant due to delayed distribution, OPeNDAP issues such as full cache, or reprocessing). 43.5% of tickets are related to finding data, documentation, or services.

## 6. Improvement of Data Service

### Documentation:

- Added [MERRA-2 Data Access – Quick Start Guide](#)
- Added [Records of MERRA-2 Data Reprocessing and Service Changes](#)



### Data How-to documents:

- How to remotely access MERRA-2 with Python3 and calculate monthly average surface PM<sub>2.5</sub> for world countries ([link](#))
- How to Use the Web Services API for Subsetting MERRA-2 Data - revised ([link](#))
- How to Access MERRA-2 Data using OPeNDAP with Python3 and Calculate Daily/Weekly/Monthly Statistics from Hourly Data ([link](#))
- How to Compile Your Own Variable List from Multiple GES DISC's Products using My Dashboard ([link](#))
- How to Directly Access MERRA-2 Data from an S3 Bucket with Python from a Cloud Environment ([link](#))

### GES DISC Webpage:

- GES DISC subsetter: Make "Get Data" button inactive until one variable is selected!

### Data Tutorial:

- ARSET training, "[Tools for Analyzing NASA Air Quality Model Output](#)", showing how to search for MERRA-2 variables and documents on the GES DISC website, and demonstrating how to use data tools for the MERRA-2 data such as Giovanni and the Level 3 and Level 4 subsetter (Part 1 of the training series)

