National Aeronautics and Space Administration
Goddard Earth Science Data Information and Services Center (GES DISC)

README Document for

MERRA-2 Climate Statistics Products

Last Revised: October 20, 2020

Goddard Earth Sciences Data and Information Services Center (GES DISC)
http://disc.gsfc.nasa.gov
NASA Goddard Space Flight Center
Code 610.2
Greenbelt, MD 20771 USA

Prepared By:

Suhung Shen	**Allison Collow**
Name GES DISC Code 610.2, NASA GSFC	Name Global Modeling and Assimilation Office Code 610.1, NASA GSFC
October 20 2020	
Date	

Reviewed By:

Lena Iredell, Xiaohua Pan	November 10 2020
Reviewer Name GES DISC Code 610.2, NASA GSFC	Date

Goddard Space Flight Center
Greenbelt, Maryland

Revision History

Revision Date	Changes	Author
Oct. 20, 2020	Created document	Suhung Shen, Allison Collow

Table of Contents

1. Introduction

This document provides basic information for using MERRA-2 Climate Statistics Products (MERRA2_ClimStat).

The MERRA2_ClimStat consists of products generated for the focus of climate indices that represent the frequency and intensity of extreme weather events derived from MERRA-2 daily temperature and precipitation data. Also included in this data product is a climatological long term mean and standard deviation representing the interannual variability on a monthly timescale.

1.1 Dataset Description

The variable list of each data collection and the algorithm to generate this dataset can be found in the "File Specification for MERRA-2 Climate Statistics Products" at

https://gmao.gsfc.nasa.gov/pubs/docs/Collow1341.pdf
Please read the file specific documents carefully for selecting and downloading the data.

More MERRA-2 document can be found at NASA GMAO MERRA-2 project page: https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/

1.2 Data Disclaimer

1.2.1 Data citation

Users should use the **data Digital Object Identifiers (DOI)** to cite the data used in research papers by following this example (e.g. M2SMNXEDI_v1):

> Global Modeling and Assimilation Office (GMAO) *(2020) MERRA-2 statM_2d_edi_Nx: 2d, Single-Level, Monthly Extremes Detection Indices V1, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [Data Access Date], 10.5067/QFJ13GEGDI99*

Table 1: DOI of data collection

Product	File_keyString	Description	DOI
M2SMNXEDI	statM_2d_edi_Nx	2d, Single-Level, Monthly Extremes Detection Indices	10.5067/QFJ13GEGDI99
M2SMNXPCT	statM_2d_pct_Nx	2d, Single-Level, Monthly Percentiles	10.5067/JGAV2VDLRY9G
M2TCNXLTM	tavgC_2d_ltm_Nx	2d, Single-Level, Long Term Mean Diagnostics	10.5067/HWSZE7YK7L81
M2TCNPLTM	tavgC_3d_ltm_Np	3d, Long Term Mean 3-Dimensional Meteorological Fields	10.5067/HQR1D9MPSEJN

1.2.2 Contact Information

If you have science questions regarding this dataset:

Email: m2fluid_feedback@gmao.gsfc.nasa.gov

If you need assistance or wish to report a data access problem:

Email: gsfc-dl-help-disc@mail.nasa.gov

Address:
Goddard Earth Sciences Data and Information Services Center NASA Goddard Space Flight Center Code 610.2 Greenbelt, MD 20771 USA

2. Data Organization

This dataset consists four global gridded products, including two monthly climate indices products (i.e., M2SMNXEDI and M2SMNXPCT) and two long-term climatology products (i.e., M2TCNXLTM and M2TCNPLTM). The long-term climatology were calculated with the base period of 30 years from 1981 to 2010.

2.1 File Naming Convention

2.1.1 Monthly Climate Indices Products:

MERRA2.< File_keyString >.<yyyymm>.nc4

Where,

File_keyString: the type of product (see Table 1 in section 1.2.1)
yyyy = 4 digits year number [1980 -].
mm = 2 digits month number [01-12]

Example filename for June 2020 data:

MERRA2.statM_2d_edi_Nx.200006.nc4
MERRA2.statM_2d_pct_Nx.200006.nc4

2.1.2 Long-term Climatology Products

MERRA2.< File_keyString >.<1981mm_2010mm>.nc4

Where,

File_keyString: the type of product (see Table 1 in section 1.2.1)
mm = 2 digits month number [01-12]

Example filename for June climatology data:
MERRA2.tavgC_2d_ltm_Nx.198106_201006.nc4
MERRA2.tavgC_3d_ltm_Np.198106_201006.nc4

2.2 File Format

The files are in NetCDF-4 format. NetCDF is a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data that was developed by UCAR/Unidata (http://doi.org/10.5065/D6H70CW6)

3. Data Structure

3.1 Dimensions

Every MERRA-2 climate statistics collection contains variables that define the dimensions of longitude, latitude, and time. Although time is included, each data file only contains one-time step. Product collections that contain 3-dimensional data will also have a vertical dimension that defines pressure levels (see section 3.3). Dimension variables have an attribute named "units," set to an appropriate string defined by the CF and COARDS conventions that can be used by applications to identify the dimension.

Table 2: Dimension Variables in NetCDF Files

Name	**Description**	**Type**	***units* attribute**
lon	Longitude	double	degrees_east
Lat	Latitude	double	degrees_north
Lev	pressure or layer index	double	hPa
time	minutes since first time in file	double	minutes

3.2 Horizontal Structure

All fields are provided on the same 5/8 longitude by ½ latitude (0.625 x 0.5) degrees grid as MERRA-2. The GEOS MERRA-2 ***native grid*** is a cubed sphere; however, the output is on a global horizontal grid, consisting of **IMn=576** points in the longitudinal direction and **JMn=361** points in the latitudinal direction. The horizontal native grid origin, associated with variables indexed (*i*=1, *j*=1) represents a grid point located at (180°W, 90°S). Latitude (φ) and longitude (λ) of grid points as a function of their indices (*i, j*) can be determined by:

$$\lambda_i = -180 + (\Delta\lambda)_n (i-1), \quad i = 1, \mathrm{IMn}$$
$$\varphi_j = -90 + (\Delta\varphi)_n (j-1), \quad j = 1, \mathrm{JMn}$$

Where $(\Delta\lambda)_n$=5/8° and $(\Delta\varphi)_n$ = 1/2°.

3.3 Vertical Structure

Gridded products use two different vertical configurations: Horizontal-only (can be vertical averages, single level, or surface values) or pressure-level. Horizontal-only data for a given variable appear as 2-dimensional fields (x, y), while pressure-level data appear as 3-dimensional fields (x, y, z). In all cases the time dimension spans multiple files. Pressure-level data is output on the **LMp=12** pressure levels shown in Table 3.

Table 3: Pressure-level data will be output on the following 12 pressure levels:

Level	***Pressure (hPa)***	***Level***	***Pressure (hPa)***
1	1000	7	400

2	925	8	300
3	850	9	200
4	700	10	100
5	600	11	30
6	500	12	10

4. Data Access Services

All users are required to be registered with the NASA Earthdata Login. Detailed instructions on how to register and receive authorization to access all GES DISC data are provided at: https://disc.gsfc.nasa.gov/data-access .

4.1 HTTPS

The user can download the native archived data files via HTTPS.

Example data:

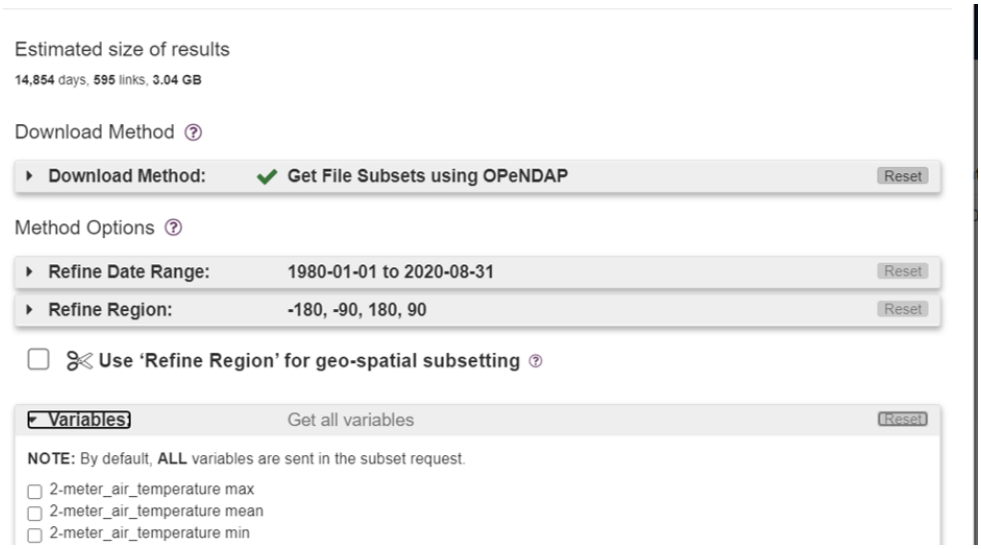
https://goldsmr4.gesdisc.eosdis.nasa.gov/data/MERRA2_CLIM/M2SMNXEDI.1/1980/MERRA2.statM_2d_edi_Nx.198001.nc4

The following How-to document gives an example to download one or more data files:

How to Download Data Files from HTTPS Service with wget

4.2 Subsetting Service

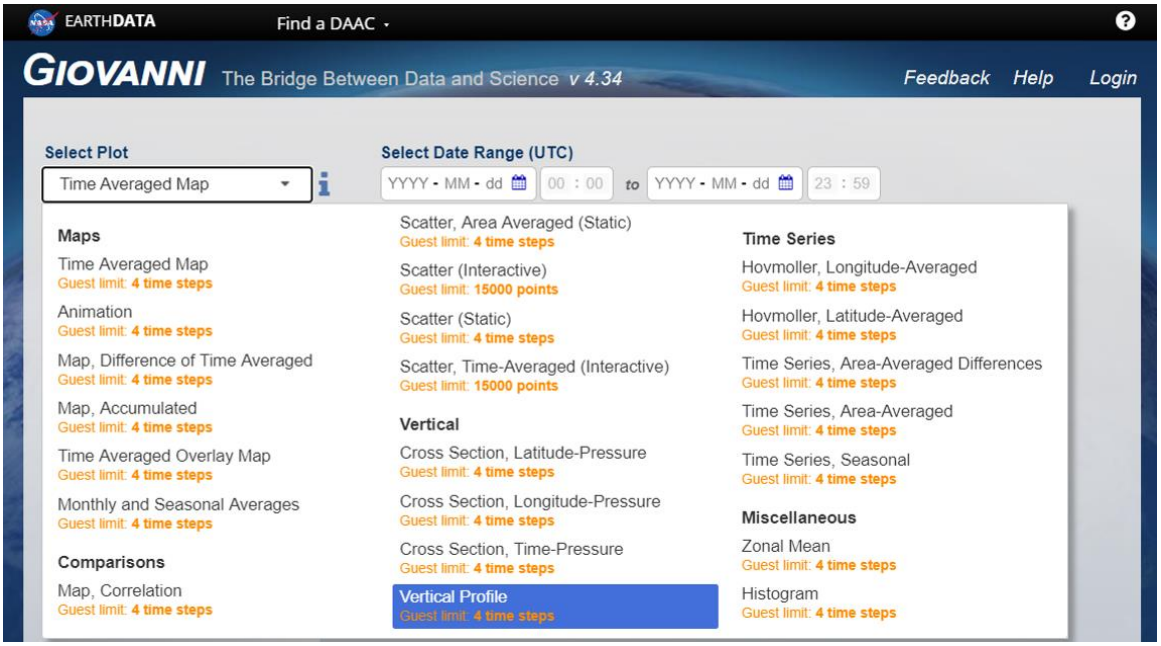
This service enables the variable, spatial, and vertical subsetting of a data file. This makes one to download only the segment of interested data. Example interface:



4.3 Online Visualization:

The monthly climate statistic data are available in Giovanni, which is a GES DISC developed data tool. Giovanni allows a user to visualize data through any browser without downloading data: https://giovanni.gsfc.nasa.gov/giovanni/

The available plot functions are shown in the Figure:



4.4 OPeNDAP:

OPeNDAP enables the remote data accessing and processing without downloading data to local machine via a number of software, such as panoply, IDV, IDL, Matlab, R, python. Please find OPeNDAP Data How-to document.

4.5 Online Help Document

The GESDISC web site (https://disc.gsfc.nasa.gov/) contains many informative articles under the "How-To's", "FAQ", "Data-in-Action", "News", "Glossary", and "Help" . One may search for "MERRA" for MERRA/MERRA-2 related document, for example:

How-to: https://disc.gsfc.nasa.gov/information/howto?keywords=MERRA&page=1

FAQ: https://disc.gsfc.nasa.gov/information/faqs?keywords=MERRA&page=1

Examples:

Earthdata Login for Data Access

How to Download Data Files from HTTPS Service with wget

How to use the Level 3 and 4 Subsetter and Regridder

5. Options and Examples for Reading the Data

5.1 Command Line Utilities

5.1.1 ncdump

The ncdump tool can be used as a simple browser for NetCDF data files, to display the dimension names and sizes; variable names, types, and shapes; attribute names and values; and optionally, the values of data for all variables or selected variables in a netCDF file. The most common use of ncdump is with the –h option, in which only the header information is displayed. Read more https://www.unidata.ucar.edu/software/netcdf/docs/netcdf_utilities_guide.html

ncdump [-c|-h] [-v ...] [[-b|-f] [c|f]] [-l len] [-n name] [-p n[,n]] [-k] [-x] [-s] [-t|-i] [-g ...] [-w] [-Ln] file
Most used options:

-h : Header information only, no data

-c: Coordinate variable data and header information
-v var1[,...] : Data for variable(s) <var1>,... only

Example,

ncdump -h MERRA2.statM_2d_edi_Nx.202008.nc4

It will return:

```
netcdf MERRA2.statM_2d_edi_Nx.202008 {
dimensions:
      lon = 576 ;
      lat = 361 ;
      time = 1 ;
variables:
      float TN10p(lat, lon) ;
            TN10p:long_name = "percentage of time when daily min 2-m temperature < 10th percentile" ;
            TN10p:_FillValue = -9999. ;
            TN10p:units = "%" ;

.....

// global attributes:
            :ShortName = "M2SMNXEDI" ;
            :LongName = "MERRA-2 statM_2d_edi_Nx: 2d, Single-Level, Monthly Extreme Detection Indices" ;
 ....
```

5.1.2 HDFView

NtCDF4 files are legal HDF5 files with additional bookkeeping information managed by the NetCDF4 library. It is therefore possible to inspect and copy data out of the NetCDF4 files by using the HDF5 utilities and libraries maintained by the HDF Group (https://www.hdfgroup.org/products/hdf5_tools/index.html) or by using the HDF5 interface in your favorite programming language. However, the two libraries should not be considered fully interchangeable.

HDFView is a Java based graphical user interface created by the HDF Group which can be used to browse HDF files. The utility allows users to view all objects in an HDF file hierarchy which is represented as a tree structure. Additional information about HDFView can be found at https://support.hdfgroup.org/products/java/hdfview/

5.2 Tools/Programming

The data files can be read and queried using the NetCDF4 library and tools maintained by Unidata (http://www.unidata.ucar.edu/software/netcdf/). Support for reading NetCDF is offered in many data tools and programming languages, including **Panoply, Python, R, Matlab, IDL, IDV, ArcGIS, GrDAS, NCO,**

CDO, C/C++ and Fortran.

5.2.1 Python

The following code snippet shows how to read the variable lat, lon, and TN10p from an example file. Some basic information about the size of the variable arrays are also shown. More examples can be found at GES DISC Python How-to .

Example file: MERRA2.statM_2d_edi_Nx.202008.nc4

```
import netCDF4
from netCDF4 import Dataset

nc_fid = netCDF4.Dataset('MERRA2.statM_2d_edi_Nx.202008.nc4' ,mode='r',format='NETCDF4')

#read in the variables
lat = nc_fid.variables['lat'][:]
lon = nc_fid.variables['lon'][:]
data = nc_fid.variables['TN10p'][:]

# print out the minimum, maximum, and dimensions for the three variables
print("-- lat  Min/Max values", lat[:].min(), lat[:].max())
print("lat.shape:", lat.shape)
print("-- lon Min/Max values:", lon[:].min(), lon[:].max())
print("lon.shape:", lon.shape)
print("-- data Min/Max values:", data[:].min(), data[:].max())
print("data.shape:", data.shape)
```

5.2.2 GrADS

The Grid Analysis and Display System (GrADS) is an interactive desktop tool for easy access, manipulation, and visualization of earth science data. The documentation and software for GrADS can be found at http://cola.gmu.edu/cola.php.

Example to plot a lon-lat map from a file: MERRA2.statM_2d_pct_Nx.202008.nc4

$ grads

ga-> sdfopen MERRA2.statM_2d_pct_Nx.202008.nc4

GrADS returns:

```
Scanning self-describing file:  MERRA2.statM_2d_pct_Nx.202008.nc4
SDF file MERRA2.statM_2d_pct_Nx.202008.nc4 is open as file 1
```

```
LON set to 0 360
LAT set to -90 90
LEV set to 0 0
Time values set: 2020:8:1:0 2020:8:1:0
E set to 1 1
```

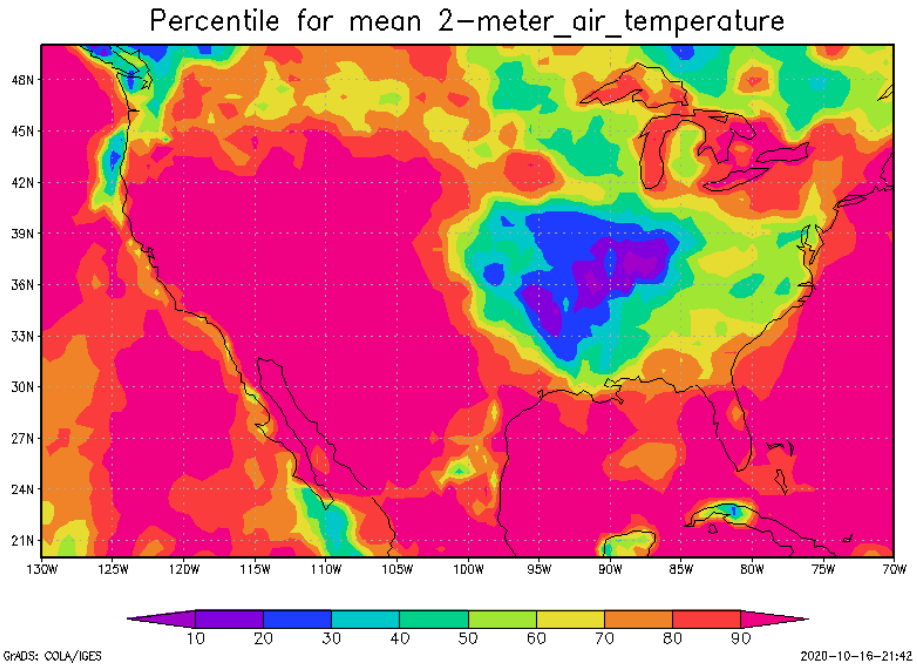
ga-> q ctlinfo

GrADS returns data information:

```
dset MERRA2.statM_2d_pct_Nx.202008.nc4
title MERRA-2 statM_2d_pct_Nx: 2d, Single-Level, Monthly Percentiles
undef -9999
dtype netcdf
xdef 576 linear -180 0.625
ydef 361 linear -90 0.5
zdef 1 linear 0 1
tdef 1 linear 00:30Z01AUG2020 1mn
vars 4
T2MMEAN=>t2mmean 0 y,x Percentile for mean 2-meter_air_temperature
T2MMAX=>t2mmax 0 y,x Percentile for maximum 2-meter_air_temperature
T2MMIN=>t2mmin 0 y,x Percentile for minimum 2-meter_air_temperature
PRECTOT=>prectot 0 y,x Percentile for total_precipitation
endvars
```

ga-> set lon -130 -70
ga-> set lat 20 50
ga-> set gxout shaded
ga-> d t2mmean
ga-> run cbarn
ga-> draw title Percentile for mean 2-meter_air_temperature
ga-> printim sample.statM_2d_pct_Nx.png white

GrADS generates map over US for t2mmean (Percentile for mean 2-meter_air_temperature) as below, and saved in a file named : sample.statM_2d_pct_Nx.png



5.2.3 Panoply

Panoply (https://www.giss.nasa.gov/tools/panoply/)is a cross-platform application that plots geo-referenced and other arrays from NetCDF, HDF, GRIB, and other data formats. The How-to page from the GES DISC provides example on view data with Panoply.

5.2.3 More How-to Examples

More reading data examples by using ArcGIS, R, Python etc. may be found under the "How-To's", https://disc.gsfc.nasa.gov/information/howto

For example:

How to read and plot NetCDF MERRA-2 data in Python

How to Import Gridded Data in NetCDF Format into ArcGIS

How to Display a Shapefile-based Data Subset with GrADS

How to Read Data in NetCDF Format with R

6. More Information

NASA GMAO MERRA-2 project : https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/

7. Acknowledgments

This MERRA-2 Climate Statistics product is produced by the NASA Global Modeling and Assimilation Office (GMAO). The distribution of the dataset is funded by NASA's Science Mission Directorate.

8. References

Bosilovich, M. G., R. Lucchesi, and M. Suarez (2016), MERRA-2: File Specification, GMAO Office Note No. 9 (Version 1.1), https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf

Collow, A. B. M., S. P. Mahanama, M. G. Bosilovich, R. D. Koster, and Siegfried D. Schubert (2017), An Evaluation of Teleconnections Over the United States in an Ensemble of AMIP Simulations with the MERRA-2 Configuration of the GEOS Atmospheric Model, NASA/TM-2017-104606, https://gmao.gsfc.nasa.gov/pubs/docs/Collow963.pdf

Gelaro, R., W. McCarty, M.J. Suárez, R. Todling, A. Molod, L. Takacs, C.A. Randles, A. Darmenov, M.G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A.M. da Silva, W. Gu, G. Kim, R. Koster, R. Lucchesi, D. Merkova, J.E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S.D. Schubert, M. Sienkiewicz, and B. Zhao (2017), The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J. Climate, doi:10.1175/JCLI-D-16-0758.1