



Global Modeling and Assimilation Office

GMAO Office Note No. 19 (Version 1.1)

File Specification for MERRA-2 Climate Statistics Products

Release Date: 07/07/2021

**Global Modeling and Assimilation Office
Earth Sciences Division
NASA Goddard Space Flight Center
Greenbelt, Maryland 20771**

This page intentionally left blank.

File Specification for MERRA-2 Climate Statistics Products

Document maintained by Allison Collow (GMAO, USRA)

This document should be cited as

Collow, A., M. Bosilovich, A. Dezfuli, and R. Lucchesi, 2021: File Specification for MERRA-2 Climate Statistics Products. GMAO Office Note No. 19 (Version 1.1), 15 pp, available from http://gmao.gsfc.nasa.gov/pubs/office_notes.

Approved by:

Steven Pawson

Date

Head, Global Modeling and Assimilation Office
Code 610.1, NASA GSFC

REVISION HISTORY

Version Number	Revision Date	Extent of Changes
1.0	11/16/2020	Baseline
1.1	07/07/2021	Clarified variable descriptions

TABLE OF CONTENTS

1. INTRODUCTION	1
2. FORMAT AND FILE ORGANIZATION	2
2.1 DIMENSIONS.....	2
2.2 VARIABLES.....	2
2.3 GLOBAL ATTRIBUTES.....	4
3. GRID STRUCTURE	6
3.1 HORIZONTAL STRUCTURE.....	6
3.2 VERTICAL STRUCTURE.....	6
4. FILE NAMING CONVENTIONS	7
4.1 FILE NAMES.....	7
5.2 EARTH SCIENCE DATA TYPES (ESDT) NAME.....	8
5. MERRA-2 CLIMATE STATISTICS DATA COLLECTIONS	9
<i>statM_2d_edi_Nx: Extremes Detection Indices</i>	9
<i>statM_2d_pct_Nx: Percentiles</i>	11
<i>tavgC_2d_ltm_Nx: Long Term Mean</i>	11
<i>tavgC_3d_ltm_Np: Long Term Mean</i>	13
REFERENCES	15
WEB RESOURCES	15

1. Introduction

The Modern Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) contains a wealth of information that can be used for weather and climate studies. By combining the assimilation of observations with a frozen version of the Goddard Earth Observing System (GEOS), a global analysis is produced at an hourly temporal resolution spanning from January 1980 through present (Gelaro et al., 2017). It can be difficult to parse through a multidecadal dataset such as MERRA-2 to evaluate the interannual variability of weather that occurs on a daily timescale, let alone determine the occurrence of an extreme weather event. Furthermore, it was recognized that standard metrics were needed to evaluate climate change among climate models and international research efforts. As a result of these concerns, the Expert Team on Climate Change Detection and Indices (ETCCDI) developed a set of indices that represent the frequency and intensity of extreme weather events using a daily time series of 2-m air temperature (T2m) and precipitation (Alexander et al., 2016). These indices were used as a basis to comprise a list of fields that represent daily extreme temperature and precipitation events, heatwaves, multi-day precipitation, as well monthly percentile statistics from the MERRA-2 dataset. Also included in this data product is a climatological long term mean and standard deviation representing the interannual variability on a monthly timescale.

All percentiles computed based on daily 2-m temperature and precipitation fields were calculated using a running window of +/- 7 days centered on each day of the year for the climatology period of 1981 through 2010. The monthly percentile statistics, climatological long term mean, and standard deviation were also generated for each grid point using the climatology period of 1981 through 2010. Precipitation used to generate the climate statistics is the model generated version that has not been corrected by observations (Reichle et al., 2017). A heatwave is defined as at least three consecutive days in which the daily mean 2-m air temperature exceeds the 90th percentile. Due to the nature of providing data on a monthly basis, indices such as the duration of the longest heatwave event only represent dates within the current month. However, a heatwave event can begin during the preceding month or end in the following month.

This document describes the gridded climate statistics files produced using output from the MERRA-2 reanalysis. Further details regarding MERRA-2 can be found in Gelaro et al. (2017), standard output from MERRA-2 is detailed Bosilovich et al. (2016), while a discussion of extreme indices in MERRA-2 can be found in Collow et al. (2017). An example in the literature for heatwave statistics is Perkins et al. (2013).

The MERRA-2 climate statistics data products are available online through the Goddard Earth Sciences (GES) Data and Information Services Center (DISC) (<http://disc.sci.gsfc.nasa.gov/mdisc/>). All data collections are provided on a monthly time scale at the same horizontal grid as MERRA-2. This grid has 576 points in the longitudinal direction and 361 points in the latitudinal direction, corresponding to a resolution of 0.625° x 0.5°.

2. Format and File Organization

MERRA-2 climate statistics data files are provided in netCDF-4 format. Since netCDF-4 files are actually HDF-5 files that are structured in a special way, netCDF-4 files can also be read by HDF-5 tools. The data files adhere to the netCDF “classic” data model, which will allow source code used to read older netCDF formats to still work when compiled with the netCDF-4 and HDF-5 libraries. The data products will adhere to the older COARDS metadata conventions and many of the CF metadata conventions, although the files are not fully CF-compliant. The conventions for identifying dimension information are followed, which should allow MERRA-2 files to be used by many tools that are CF-compliant.

2.1 Dimensions

Every MERRA-2 climate statistics collection will contain 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 3.2). 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 2.1-1. Dimension Variables Contained in GMAO NetCDF Files

Name	Description	Type	<i>units</i> attribute
lon	Longitude	double	degrees_east
lat	Latitude	double	degrees_north
lev	pressure or layer index	double	hPa or layer
time	minutes since first time in file	double	minutes

2.2 Variables

The MERRA-2 climate statistics data product is available as netCDF-4 files. This allows applications written to read netCDF files to easily read variables without having to modify code. Variable names are listed in Section 6 along with the number and sizes of dimensions. One can quickly list the variables in the file by using common utilities such as *ncdump*, which is distributed with the netCDF-4 library. With the ‘-h’ flag, this utility will display all information about the file and its contents, including metadata associated with each variable. The variable name is analogous to the *short name* in the MERRA-2 climate statistics files. A short description of the variable is provided in the *long_name* and *standard_name* metadata parameters.

Each variable has several useful metadata attributes. Many of these attributes are required by the [CF](#) and [COARDS](#) conventions, while others are specific for GMAO products. The following table

lists required attributes. Other attributes may be included for internal GMAO use and can be ignored.

Table 2.2-1 Metadata attributes associated with each variable.

Name	Type	Description
_FillValue	32-bit float	Floating-point value used to identify missing data. Required by CF.
long_name	String	A brief description of the variable contents taken from the <i>Description</i> column of the tables in Appendix D.
units	Char String	The units of the variable. Must be a string that can be recognized by UNIDATA's Uunits package.

2.3 Global Attributes

In addition to dataset variables and dimension scales, global metadata is also stored in GMAO netCDF-4 files. Some metadata are required by the CF/COARDS conventions, some are present to meet EOSDIS requirements, and others as a convenience to users of GMAO products. A summary of global attributes present in all MERRA-2 files is shown in Table 2.3-1. All global metadata parameters are of type character

Table 2.3-1 Global metadata attributes associated with each SDS.

Name	Description
Institution	“NASA Global Modeling and Assimilation Office”
Conventions	CF1.7
Format	“NetCDF-4”
SpatialCoverage	global
VersionID	The version of the data product
Temporal Range	The beginning and ending dates of the data used to produce the granule
identifier_product_doi_authority	“http://dx.doi.org”
ShortName	Product short name used by GESDISC
RangeBeginningDate	Date corresponding to the first timestep in this file.
RangeBeginningTime	Time corresponding to the first timestep in this file.
RangeEndingDate	Date corresponding to the last timestep in this file.
RangeEndingTime	Time corresponding to the last timestep in this file.
ProcessingLevel	“Level 4”
GranuleID	Filename for this product
ProductionDateTime	Production date & time of this granule.
LongName	Description of product type.
Title	Description of product type.
MapProjection	Latitude-Longitude

Name	Description
SouthernmostLatitude	"-90.0"
NorthernmostLatitude	"90.0"
WesternmostLatitude	"-180.0"
EasternmostLatitude	"179.375"
LatitudeResolution	"0.5"
LongitudeResolution	"0.625"
identifier_product_doi	Unique Digital Object Identifier
Source	CVS tag: GEOSadas-5_12_4
Contact	"http://gmao.gsfc.nasa.gov"

3. Grid Structure

3.1 Horizontal Structure

All fields are provided on the same $5/8^\circ$ longitude by $1/2^\circ$ latitude 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 ($180^\circ\text{W}, 90^\circ\text{S}$). Latitude () and longitude () of grid points as a function of their indices (i, j) can be determined by:

$$\lambda_i = -180 + (\Delta\lambda)_n(i - 1), \quad i = 1, \text{IMn}$$

$$\varphi_j = -90 + (\Delta\varphi)_n(j - 1), \quad j = 1, \text{JMn}$$

Where $(\Delta\lambda)_n = 5/8^\circ$ and $(\Delta\varphi)_n = 1/2^\circ$. For example, ($i = 289, j = 181$) corresponds to a grid point at ().

3.2 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 3.2-1.

Table 3.2-1: Pressure-level data is output on the following 12 pressure levels:

<i>Level</i>	<i>Pressure (hPa)</i>	<i>Level</i>	<i>Pressure (hPa)</i>
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. File Naming Conventions

Each GEOS-5 product file will have a complete file name identified in the EOSDIS metadata as "LocalGranuleID". EOSDIS also requires eight-character abbreviated naming indices for each Earth Science Data Type (ESDT). In MERRA-2 each file collection has a unique ESDT index. The ESDT index convention is described in section 4.2.

4.1 File Names

The standard full name for the MERRA-2 Climate Statistics products will consist of three dot-delimited nodes:

runid.collection.timestamp

The node fields, which vary from file to file, are defined as follows:

collection:

All MERRA-2 Climate Statistics data are organized into file *collections* that contain fields with common characteristics. These collections are used to make the data more accessible for specific purposes. Collection names are of the form *freq_dims_group_HV*, where the four attributes are:

freq: statistics (**stat***F*) or time-average (**tavg***F*), where *F* indicates the frequency or averaging interval and can be any of the following:

M = Monthly Value

C = Climatological Value

dims: **2d** for collections with only 2-dimensional fields or **3d** for collections with a mix of 2- and 3-dimensional fields.

group: A three-letter mnemonic for the type of fields in the collection. It is a lowercase version of the group designation used in the ESDT name, as [listed in the next section](#).

HV: Horizontal and Vertical grid.

H can be:

N: Native (5/8 x 1/2) horizontal resolution

V can be:

x: horizontal-only data (surface, single level, etc.); *dims* must be **2D**

p: pressure-level data (see Section 3.2 for levels); *dims* must be **3D**

timestamp:

This node defines the date and time associated with the data in the file. It has the *yyyymm* for monthly statistics files and *yyyymm_yyyyymm* for long term mean files to denote the climatology period used.

yyyy - year string (e.g. , "2002")

mm - month string (e.g., "09" for September)

EXAMPLE:

MERRA2.statM_2d_edi_Nx.200209.nc4

This is an example of a MERRA-2 Climate Statistics filename. The data are monthly statistics ("statM"), two-dimensional ("2d"), extremes detection indices products ("edi"), at native horizontal resolution ("Nx"). The file contains a monthly value for September 2002 and is in "nc4" format.

5.2 Earth Science Data Types (ESDT) Name

To accommodate EOSDIS toolkit requirements, all files are associated with a nine-character ESDT. The ESDT is a short handle for users to access sets of files. Like with MERRA-2, the ESDT will be used to identify the *Mainstream collections* and consists of a compressed version of the collection name of the form:

M2TFHVGGG

where

M2: MERRA-2

T: Time Description

T = Time-averaged

C = Time-independent

F: Frequency

M = Monthly

0 = Not Applicable

H: Horizontal Resolution

N = Native

V: Vertical Location

X = Two-dimensional

P = Pressure

GGG: Group

EDI = extremes detection indices

LTM = long term mean

PCT = percentiles

5. MERRA-2 Climate Statistics Data Collections

This section lists the variables in each data collection.

statM_2d_edi_Nx: Extremes Detection Indices

Frequency: *Monthly*

Spatial Grid: *2D, single-level, full horizontal resolution*

Dimensions: *longitude=576, latitude=361*

Granule Size: *~26 MB*

<i>Name</i>	<i>Dim</i>	<i>Description</i>	<i>Units</i>
CDD	yx	consecutive dry days (maximum number of consecutive days when precipitation < 1 mm)	days
CSDI	yx	cold spell duration index (count when at least 6 consecutive days of min 2-m temperature < 10th percentile)	days
CWD	yx	consecutive wet days (maximum number of consecutive days when precipitation >= 1 mm)	days
drydays	yx	count of days with < 1 mm of precipitation	1
DTR	yx	diurnal 2-m temperature range	K
HWA	yx	Heatwave Amplitude (daily mean 2-m temperature on hottest day satisfying the heatwave criteria of at least three consecutive days above the 90 th percentile)	K
HWD	yx	Heatwave Duration (length of the longest number of consecutive days satisfying the heatwave criteria of at least three consecutive days above the 90th percentile)	days
HWF	yx	Heatwave Frequency (count of days satisfying the heatwave criteria of at least three consecutive days above the 90 th percentile)	1
HWM	yx	Heatwave Magnitude (average 2-m temperature	K

		anomaly on days satisfying the heatwave criteria of at least three consecutive days above the 90 th percentile)	
HWN	yx	Heatwave Number (count of events satisfying the heatwave criteria of at least three consecutive days above the 90 th percentile)	1
LCS	yx	length of longest cold spell of at least 6 consecutive days below the 10 th percentile	days
LWS	yx	length of longest warm spell of at least 6 consecutive days above the 90 th percentile	days
PRCPTOT	yx	total precipitation on days ≥ 1 mm	mm day ⁻¹
R10mm	yx	count of days with ≥ 10 mm of precipitation	1
R20mm	yx	count of days with ≥ 20 mm of precipitation	1
R90d	yx	count of days with precipitation $> 90^{\text{th}}$ percentile	1
R90p	yx	total precipitation from days $> 90^{\text{th}}$ percentile	mm day ⁻¹
R95d	yx	count of days with precipitation $> 95^{\text{th}}$ percentile	1
R95p	yx	total precipitation from days $> 95^{\text{th}}$ percentile	mm day ⁻¹
R99d	yx	count of days with precipitation $> 99^{\text{th}}$ percentile	1
R99p	yx	total precipitation from days $> 99^{\text{th}}$ percentile	mm day ⁻¹
RX1Day	yx	maximum one-day precipitation amount	mm day ⁻¹
RX5Day	yx	highest precipitation amount for a five-day interval	mm day ⁻¹
RX5Daycount	yx	count of heavy precipitation periods ≥ 50 mm within a five-day interval	1
SDII	yx	ratio of total precipitation to the number of wet days	mm day ⁻¹
T2M	yx	2-m temperature	K
TN10p	yx	percentage of time when daily min 2-m temperature $< 10^{\text{th}}$ percentile	%
TN90p	yx	percentage of time when daily min 2-m temperature $> 90^{\text{th}}$ percentile	%
TX10p	yx	percentage of time when daily max 2-m temperature $< 10^{\text{th}}$ percentile	%
TX90p	yx	percentage of time when daily max 2-m temperature $> 90^{\text{th}}$ percentile	%

wetdays	yx	count of days with ≥ 1 mm of precipitation	1
WSDI	yx	warm spell duration index (count when at least 6 consecutive days of max 2-m temperature $> 90^{\text{th}}$ percentile)	days

statM_2d_pct_Nx: Percentiles

Frequency: *Monthly*

Spatial Grid: *2D, single-level, full horizontal resolution*

Dimensions: *longitude=576, latitude=361*

Granule Size: *~7 MB*

<i>Name</i>	<i>Dim</i>	<i>Description</i>	<i>Units</i>
PRECTOT	yx	Percentile for total precipitation	Percentile
T2MMAX	yx	Percentile for maximum 2-m air temperature	Percentile
T2MMEAN	yx	Percentile for mean 2-m air temperature	Percentile
T2MMIN	yx	Percentile for minimum 2-m air temperature	Percentile

tavgC_2d_ltm_Nx: Long Term Mean

Frequency: *Monthly*

Spatial Grid: *2D, single-level, full horizontal resolution*

Dimensions: *longitude=576, latitude=361*

Granule Size: *~55 MB*

<i>Name</i>	<i>Dim</i>	<i>Description</i>	<i>Units</i>
EMP	yx	evaporation minus precipitation	$\text{kg m}^{-2} \text{s}^{-1}$
EVAP	yx	evaporation from turbulence	$\text{kg m}^{-2} \text{s}^{-1}$
EVLAND	yx	land evaporation	$\text{kg m}^{-2} \text{s}^{-1}$

GWETROOT	yx	root zone soil wetness	1
GWETOP	yx	surface soil wetness	1
H200	yx	height at 200 hPa	m
H500	yx	height at 500 hPa	m
H850	yx	height at 850 hPa	m
LWTUP	yx	upwelling longwave flux at top of the atmosphere	W m^{-2}
PRECTOT	yx	total precipitation	$\text{kg m}^{-2} \text{s}^{-1}$
PRECTOTCORR	yx	bias corrected total precipitation	$\text{kg m}^{-2} \text{s}^{-1}$
PS	yx	surface pressure	Pa
SLP	yx	sea level pressure	Pa
SWGDN	yx	surface incoming shortwave flux	W m^{-2}
T2MMAX	yx	maximum 2-m air temperature	K
T2MMEAN	yx	2-m air temperature	K
T2MMIN	yx	minimum 2-m air temperature	K
TQV	yx	total precipitable water vapor	kg m^{-2}
TS	yx	surface skin temperature	K
U10M	yx	10-meter eastward wind	m s^{-1}
U200	yx	eastward wind at 200 hPa	m s^{-1}
U2M	yx	2-meter eastward wind	m s^{-1}
U500	yx	eastward wind at 500 hPa	m s^{-1}
U50M	yx	eastward wind at 50 meters	m s^{-1}
U850	yx	eastward wind at 850 hPa	m s^{-1}
UFLXQV	yx	eastward flux of atmospheric water vapor	$\text{kg m}^{-1} \text{s}^{-1}$
V10M	yx	10-meter northward wind	m s^{-1}
V200	yx	northward wind at 200 hPa	m s^{-1}
V2M	yx	2-meter northward wind	m s^{-1}
V500	yx	northward wind at 500 hPa	m s^{-1}
V50M	yx	northward wind at 50 meters	m s^{-1}

V850	yx	northward wind at 850 hPa	m s ⁻¹
VFLXQV	yx	northward flux of atmospheric water vapor	kg m ⁻¹ s ⁻¹

tavgC_3d_ltm_Np: Long Term Mean

Frequency: *Monthly*

Spatial Grid: *3D, single-level, full horizontal resolution*

Dimensions: *longitude=576, latitude=361, pressure=12*

Granule Size: *~153 MB*

<i>Name</i>	<i>Dim</i>	<i>Description</i>	<i>Units</i>
H	zyx	layer heights	m
OMEGA	zyx	vertical pressure velocity	Pa s ⁻¹
QV	zyx	specific humidity	kg kg ⁻¹
RH	zyx	relative humidity	1
T	zyx	air temperature	K
U	zyx	eastward wind	m s ⁻¹
V	zyx	northward wind	m s ⁻¹
Var_V	zyx	variance of northward wind	m s ⁻¹ m s ⁻¹

6. Data Object Identifiers

Digital Object Identifiers are attached to each MERRA-2 climate statistics collection. Users should cite the data used in research papers following these DOI's.

Example Citation:

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

Note that complete citations for each file collection are provided at the GES-DISC download site.

Table 6.1 DOIs for MERRA-2 Climate Statistics Products

Descriptive Short Name	Short Name	DOI
statM_2d_edi_Nx	M2SMNXEDI	10.5067/QFJ13GEGDI99
statM_2d_pct_Nx	M2SMNXPCT	10.5067/JGAV2VDLRY9G
tavgC_2d_ltm_Nx	M2TCNXLTM	10.5067/HWSZE7YK7L81
tavgC_3d_ltm_Np	M2TCNPLTM	10.5067/HQR1D9MPSEJN

References

Alexander L. V., 2016: Global observed long-term changes in temperature and precipitation extremes: A review of progress and limitations in IPCC assessments and beyond. *Weather and Climate Extremes*, 11, 4-6, doi:10.1016/j.wace.2015.10.007.

Bosilovich, M. G., R. Lucchesi, and M. Suarez, 2016: MERRA-2: File Specification. GMAO Office Note No. 9 (Version 1.1), 73 pp, available from http://gmao.gsfc.nasa.gov/pubs/office_notes.

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, Vol. 47, 68 pp, <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*, 30, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.

Perkins, S.E. and L.V. Alexander, 2013: On the Measurement of Heat Waves. *J. Climate*, 26, 4500–4517, doi: 10.1175/JCLI-D-12-00383.1.

Reichle, R.H., Q. Liu, R.D. Koster, C.S. Draper, S.P. Mahanama, and G.S. Partyka, 2017: Land Surface Precipitation in MERRA-2. *J. Climate*, 30, 1643–1664, <https://doi.org/10.1175/JCLI-D-16-0570.1>.

Web Resources

GMAO web site: <http://gmao.gsfc.nasa.gov/>

GMAO Operations page: <http://gmao.gsfc.nasa.gov/products/>

CF Standard Description: <http://cf-pcmdi.llnl.gov/>

FLUID Visualizations: <https://fluid.nccs.nasa.gov/reanalysis/>