Global Modeling and Assimilation Office

GMAO Office Note No. 20 (Version 1.1)

File Specification for MERRA-2 Country-Level Surface PM2.5 Monthly Mean Products

Release Date: December 8, 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 Country-Level Surface PM2.5 Monthly Mean Product

Document maintained by Helen Amos (helen.m.amos@nasa.gov)

This document should be cited as

C. Keller, H. Amos, V. Buchard, A. Darmenov, S. Shen, M. Deloria Knoll, F. Villavicencio, M. Bosilovich, 2021: File Specification for MERRA-2 Country-Level Surface PM2.5 Monthly Mean Product. GMAO Office Note No. 20 (Version 1.1), 10 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	December 8, 2021	Baseline
1.1	March 9, 2022	Corrected hyperlink to Randles et al. (2016).
1.1	April 24, 2023	Corrected bug in underlying country mask affecting the Democratic Republic of the Congo and the Republic of the Congo.

TABLE OF CONTENTS

1. Introduction

The Modern Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) is a global reanalysis produced from January 1980 through present, with a frozen version of the NASA's global Earth System model, GEOS (Goddard Earth Observing System) and includes the assimilation of meteorological and aerosol observations (Gelaro et al., 2017).

In the MERRA-2 configuration of the GEOS system, the atmosphere model component (i.e., atmospheric general circulation model/AGCM) is responsible for the meteorology, whereas the Goaddard Chemistry Aerosol Radiation and Transport (GOCART) simulates aerosols in the atmosphere (Chin et al., 2002; Colarco et al., 2010). The GOCART configuration used in MERRA-2 includes five types of aerosols: dust, sea salt, black carbon, organic carbon and sulfate aerosols. Meteorology and aerosols are modeled at a horizontal resolution of approximately 50 km. The atmosphere is partitioned into 72 vertical layers between the surface and about 80 km. The aerosol model considers natural and anthropogenic emissions, losses and chemical production mechanisms. Dust and sea salt emissions are surface wind-speed dependent, whereas emissions of other aerosol types are prescribed from inventories (table 1 in Randles et al., 2017). The model includes convective and large-scale wet removal, dry deposition and sedimentation for dust and sea salt as well as chemical reactions to produce sulfate aerosol from oxidation of sulfur dioxide (SO_2). Mass concentration of aerosols in the lowest layer of MERRA-2 atmosphere is used to quantify speciated and total surface concentration of outdoor ambient aerosol particles. More details can be found in Randles et al. (2016, 2017).

In MERRA-2, assimilation of meteorological parameters such as winds, pressure and temperature fields, and aerosol observations is performed and applied independently to the coupled atmosphere and aerosol components. Aerosol observations are assimilated every three hours, and include aerosol optical depth (AOD at 550 nm) from several ground and spaceborne sensors such as Advanced Very High Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectroradiometer (MODIS), Multi-angle Imaging Spectroradiometer (MISR) over bright surfaces and the Aerosol Robotic Network (AERONET) (Randles et al., 2016). AOD (550 nm) is a vertical integral quantity and as such its assimilation constrains well total aerosol mass in the column. The limited informational content of the assimilated AOD means that aerosol mass speciation and vertical distribution are more weakly constrained by the AOD data (Randles et al., 2017; Buchard et al., 2017).

Further details regarding the MERRA-2 system and MERRA-2 output datasets can be found in Gelaro et al. (2017) and Bosilovich et al. (2016), while a discussion of aerosols and surface PM2.5 in MERRA-2 can be found in Randles et al. (2016, 2017) and Buchard et al. (2017). In addition, readers may want to also familiarize themselves with studies that investigated PM2.5 from the predecessor of MERRA-2 (MERRAero; Buchard et al., 2016; Provencal et al., 2016, 2017).

This document describes the M2_TMAX_PM25 data collection, which provides MERRA-2 country-level monthly-average surface level concentrations of fine particulate matter (PM2.5) for 1980-2020 and is located at https://disc.gsfc.nasa.gov/datacollection/M2_TMAX_PM25_1.html. The M2_TMAX_PM25 data collection is experimental. This document describes the method for

translating gridded MERRA-2 speciated aerosol concentrations to a country-level PM2.5 and applying optional population density weighting (Section 2), the structure of the resulting country-level data (Section 3), file naming conventions (Section 4), the files contained within the M2_TMAX2_PM25 data collection (Section 5), and how to cite the data collection (Section 6).

2. Methodology

MERRA-2 country-level monthly average surface PM2.5 data are provided in a comma-separated values (CSV) files. These values are calculated from the gridded monthly average MERRA-2 model fields, which are archived under the data collection name M2TMNXAER_5.12.4 (GMAO, 2015a).

2.1 Calculating PM2.5

To convert MERRA-2 aerosol output data to fine particular matter PM2.5, the formulation described in Buchard et al. (2017) is used:

$$PM2.5 = PM2.5^{DU} + PM2.5^{SS} + PM2.5^{OC} + PM2.5^{BC} \\ + (132.14/96.06) \times PM2.5^{SO4} \tag{1}$$

where the right-hand side is the sum of the speciated surface concentration of particles smaller than 2.5 micrometers. Table 1 lists the corresponding aerosol types and variable names (Bosilovich et al., 2016) in the M2TMNXAER files. The factor of 132.14/96.06 is applied to convert sulfate ion (molar mass of 96.06 g mol^{-1}) concentration output by MERRA-2 to ammonium sulfate (132.14 g mol^{-1}), assuming that sulfate is primarily present as neutralized ammonium sulfate.

The PM2.5 data is provided in units of microgram per cubic meter (μg m^{-3}).

Table 1. Aerosol types and names of variables in M2TMNXAER files required to compute PM2.5 (see Equation 1).

aerosol type	Speciated PM2.5	variable name
dust	$PM2.5^{DU}$	DUSMASS25
sea salt	$PM2.5^{SS}$	SSSMASS25
organic carbon	$PM2.5^{OC}$	OCSMASS
black carbon	$PM2.5^{BC}$	BCSMASS
sulfate	$PM2.5^{BC}$	SO4SMASS

2.2 Country-Level Aggregation and Popluation Weighting

The monthly averaged country-level MERRA-2 PM2.5 products are calculated from the following data sources:

1. Gridded MERRA-2 PM2.5 concentration fields.
2. Gridded country masks, to determine the geographic area over which the PM2.5 concentration shall be averaged.
3. Gridded population density, to provide the weights given to the individual cells within an area.

Country-level PM2.5 concentrations are derived from the gridded MERRA-2 PM2.5 concentration (Eq. 1) and a mask file for geo-poltical boundaries (hereafter referred to as the "country mask"). The country mask is created from The Gridded Population of the World, Version 3 (GPWv3) National Identifier Grid data (CIESIN, 2005), as provided by the Emissions of atmospheric Compounds and Compilation of Ancillary Data (ECCAD) database (https://eccad3.sedoo.fr/#GPW3), and includes countries, territories and islands. The mask is used to identify which MERRA-2 model grid cells fall within which geopolitical boundaries. The resulting PM2.5 is the area-weighted average of grid cells within the boundaries of the GPWv3 country, territory and island boundaries and is hereafter referred to as "country-level PM2.5" for simplicity. Due to mask artifacts, PM2.5 values may appear as zero (0) for some small islands.

In addition, population weighted PM2.5 concentrations are provided to users as a separate data file. These values are calculated the same as above, except that the country-wide averaging is performed by weighting all grid cells within a given country based on population density, i.e., giving higher weights to grid cells with a higher population. The GPWv4 data set (CIESIN, 2018) is used as input for the population density map. GPWv4 contains estimates of population density data for years 2000, 2005, 2010, 2015, and 2020, and the closest available year is used for any given year.

2.3 Python Scripts

A python tutorial demonstrating how to aggregate native gridded MERRA-2 output to country-level PM2.5 values and optionally apply population weighting is available here on the GES DISC website.

3. Spatial Structure

3.1 Horizontal Structure

All values are aggregated by country from the original MERRA-2 output available at 5/8° longitude by 1/2° latitude grid.

3.2 Vertical Structure

PM2.5 values are calculated for the surface model level. The thickness of the surface layer is variable with typical values between approximately 61 meters in winter and 64 meters in summer over land (GMAO, 2015b).

4. Naming Conventions

The monthly averaged country-level MERRA-2 PM2.5 products are a single data collection consisting of two data files. Naming conventions for the data collection and associated data files are described in this section.

4.1 Data Collection Name

To conform with NASA's Earth Observing System Data and Information System (EOSDIS) requirements, files are associated with a nine-character Earth Science Data Types (ESDT) name. The ESDT "Shortname" for this data collection is M2_TMAX_PM25 and takes the form:

M2_TFHV_PM25

where
  M2: Shorthand for MERRA-2

  T: Time description
        T = Time-averaged

  F: Frequency
        M = Monthly

  H: Horizontal resolution
        A = Country level (i.e., administrative level 0)

  V: Vertical resolution
        X = horizontal-only (surface, single level)

  PM25: Shorthand for fine particulate matter PM2.5

4.2 File Names

The standard full name for the MERRA-2 PM2.5 products consists of five dot-delimited nodes:

runid.collection.mode.version.daterange.csv

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

> ***runid***
>   All files will begin with the runid = "MERRA2"

> ***collection***

The collection name is of the form *freq_dims_group_HV*, where the four attributes are:

freq: **tavgM** for monthly average

dims: **2d** for two-dimensional

group: **pm25** for fine particulate matter (PM2.5) concentrations

HV: Horizontal and Vertical resolution
  H is **admin0** for country-level

  V is **x** for horizontal-only data (surface, single level)

mode
  All PM2.5 data are aggregated at the country level.
  pw = population weighting applied (optional)

version
  The version number, beginning with "v01"

daterange
  The date range = *startdate_enddate* with dates formatted as yyyy-mm-dd (year-month-day)

csv
  Files are formatted as comma-separated values and given the suffix ".csv"

EXAMPLE #1: Country-level PM2.5

MERRA2.tavgM_2d_pm25_admin0x.v01.19800101_20201231.csv

This is an example of MERRA-2 ("MERRA2") monthly average ("tavgM") two-dimensional ("2d") PM2.5 concentrations ("pm25") at the country-level ("admin0") for horizontal-only data ("x"), version 1 ("v01"). The file contains surface-level monthly averages for world countries from 01 January 1980 ("19800101") through 31 December 2020 ("20201231"). The file is in comma-separated values ("csv") format.

EXAMPLE #2: Population weighted country-level PM2.5

MERRA2.tavgM_2d_pm25_admin0x_pw.v01.19800101_20201231.csv

This is an example of MERRA-2 ("MERRA2") monthly average ("tavgM") two-dimensional ("2d") PM2.5 concentrations ("pm25") at the country-level ("admin0") for horizontal-only data ("x"), version 1 ("v01"). Population weighting has been applied (see Section 2). The file contains monthly averages for world countries from 01 January 1980 ("19800101") through 31 December

2020 ("20201231"). The file is in comma-separated values ("csv") format.

5. Data Files

The M2_TMAX_PM25 data collection consists of two data files:
    MERRA2.tavgM_2d_pm25_admin0x.v01.19800101_20201231.csv
    MERRA2.tavgM_2d_pm25_admin0x_pw.v01.19800101_20201231.csv

Plus the ancillary file:
    mask_worldCountries_MERRA2.nc
    mask_worldPopulation_MERRA2.nc

5.1 File Organization

A single csv PM2.5 data file contains the MERRA-2 monthly average country-level surface PM2.5 concentrations for 229 world countries and territories. The difference between the two csv data files is whether population weighting has been applied (see Section 2). For each month starting from January 1980, the county-average PM2.5 values are stored in a single column with the first column indicating the date (yyyy-mm) and the next 229 columns containing the PM2.5 concentrations for countries and territories. The column names are given in the first row (header) of the file.

5.2 PM2.5 Data Files

MERRA2.tavgM_2d_pm25_admin0x.v01.*daterange*.csv

    Frequency: *Monthly*
    Spatial information: *2D, surface, single-level, country-level*
    Dimensions: *rows*=number of months in *daterange* + 1*, columns*= 230
    File Size: *>650 KB*

MERRA2.tavgM_2d_pm25_admin0x_pw.v01.*daterange*.csv

    Frequency: *Monthly*
    Spatial information: *2D, single-level, country-level, population weighted*
    Dimensions: *rows*=number of months in *daterange* + 1*, columns*= 230
    File Size: *>650 KB*

5.3 Ancillary Files

mask_worldCountries_MERRA2.nc

    Frequency: *Once*
    Spatial information: *2D, single-level, gridded (MERRA-2 model grid), world countries and territories*
    Dimensions: *lon=576, lat=361*
    File Size: *>190 MB*

mask_worldPopulation_MERRA2.nc

    Frequency: *Once*
    Spatial information: *2D, single-level, gridded (MERRA-2 model grid)*
    Dimensions: *lon=576, lat=361*
    File Size: *>190 MB*

6. Data Citation

Users should cite the data as:

Global Modeling and Assimilation Office (GMAO) (2021) MERRA-2 tavgM_2d_pm25_admin0x: 2d, Single-Level, Country-Level, Monthly Average PM2.5, V1, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [Data Access Date], 10.5067/FY616726UXSR

Acknowledgements

References

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.

Buchard, V., Da Silva, A. M., Randles, C. A., Colarco, P., Ferrare, R., Hair, J., et al. (2016). Evaluation of the surface PM2.5 in Version 1 of the NASA MERRA Aerosol Reanalysis over the United States. *Atmospheric Environment*, 125, 100–111. http://doi.org/10.1016/j.atmosenv.2015.11.004

Buchard, V., C. A. Randles, A. da Silva, A. S. Darmenov, P. R. Colarco, R. C. Govindaraju, R. A. Ferrare, J. W. Hair, A. Beyersdorf, L. D. Ziemba, and H. Yu, 2017. The MERRA-2 Aerosol Reanalysis, 1980-onward, Part II: Evaluation and Case Studies. J. Clim, 30, 6851-6872. DOI: 10.1175/JCLI-D-16-0613.1

Center for International Earth Science Information Network - CIESIN – Columbia University. 2005: Gridded Population of the World, Version 3 (GPWv3): National Boundaries Socioeconomic Data and Applications Center (SEDAC).

Center for International Earth Science Information Network - CIESIN - Columbia University. 2018. Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. NASA Socioeconomic Data and Applications Center, Accessed: 4 November 2018. (SEDAC). https://doi.org/10.7927/H49C6VHW.

Chin, M., and Coauthors, 2002: Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and sun photometer measurements. Journal of the Atmospheric Sciences, 59, 461–483, 10.1175/1520-0469(2002)059

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. Journal of Geophysical Research, 115 (D14207), 10.1029/2009JD012820, http://dx.doi.org/10.1029/2009JD012820.

Gelaro, R. et al., 2017, The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J. Clim., doi: 10.1175/JCLI-D-16-0758.1, 2017

Global Modeling and Assimilation Office (GMAO) (2015a), MERRA-2 tavgM_2d_aer_Nx: 2d,Monthly mean,Time-averaged,Single-Level,Assimilation,Aerosol Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 19 November 2021 , 10.5067/FH9A0MLJPC7N

Global Modeling and Assimilation Office (GMAO) (2015b), MERRA-2 instM_2d_lfo_Nx: 2d,Monthly mean,Instantaneous,Single-Level,Assimilation,Land Surface Forcings V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 10 August 2021, 10.5067/11F99Y6TXN99

Provencal, S., V. Buchard, A. M. da Silva, R. Leduc, and N. Barrette. Evaluation of PM2.5 surface concentration simulated by Version 1 of the NASAs MERRA Aerosol Reanalysis over Israel and Taiwan. Aerosol Air Qual. Res, doi:10.4209/aaqr.2016.04.0145, 2016.

Provencal, S., V. Buchard, A. da Silva, R. Leduc, and N. Barrette. Evaluation of PM surface concentration simulated by Version 1 of the NASA MERRA Aerosol Reanalysis over Europe. Atmos. Pollut. Res, doi:10.1016/j.apr.2016.10.009, 2016.

Randles, C. A., A. da Silva, V. Buchard, P. R. Colarco, A. S. Darmenov, R. C. Govindaraju, A. Smirnov, R. A. Ferrare, J. W. Hair, Y.Shinozuka, and C. Flynn. The MERRA-2 Aerosol Reanalysis, 1980-onward, Part I: System Description and Data Assimilation Evaluation. Journal of Climate, doi:10.1175/JCLI-D-16-0609.1, 2017.

Randles, C.A., A. M. da Silva, V. Buchard, A. Darmenov, P. R. Colarco, V. Aquila, H. Bian, E. P. Nowottnick, X. Pan, A. Smirnov, H. Yu, and R. Govindaraju. The MERRA-2 Aerosol Assimilation, Technical Report Series on Global Modeling and Data Assimilation, Volume 45, NASA/TM–2016-104606/Vol. 45, 2016. [PDF]

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/