Users' Guide for Estimated Surface $PM_{2.5}$ from Tropospheric Emissions: Monitoring of Pollution (TEMPO) & Advanced Baseline Imager (ABI) Aerosol Retrievals

Version 2.1, June 2025

Table of Contents

List of Figures

List of Tables

1. Purpose of this Guide

This User's Guide is intended for users of the estimated surface concentrations of particulate matter with a median diameter of 2.5 µm or less ($PM_{2.5}$) Level 4 (L4) files generated from Tropospheric Emissions: Monitoring of Pollution (TEMPO) and Advanced Baseline Imager (ABI) aerosol retrievals. It provides a general introduction to TEMPO and ABI, the estimated surface $PM_{2.5}$ product, and the format and contents of the TEMPO-ABI estimated surface $PM_{2.5}$ L4 data files. This guide serves as an introduction to the more technical TEMPO-ABI $PM_{2.5}$ Algorithm Theoretical Basis Document (ATBD), which is not yet finalized.

2. Points of Contact

For questions or comments regarding this document, please contact Shobha Kondragunta, Hai Zhang or Amy Huff.

3. Document Definitions

Surface concentrations of particulate matter with a median diameter of 2.5 µm or less ($PM_{2.5}$) are estimated using satellite aerosol optical depth (AOD) and aerosol layer height (ALH) through a geographically weighted regression (GWR) algorithm. Table 1 lists additional acronyms and abbreviations used in this document.

4. TEMPO and ABI Overview

TEMPO is the first space-based instrument to monitor air pollutants on an hourly basis across North America during the daytime. It is an ultraviolet (UV, 290-490 nm)–visible (Vis, 540-740 nm) spectrometer. TEMPO flies onboard Intelsat-40e (IS-40e), a commercial telecommunications satellite centered at 91.0° W longitude.

ABI is the primary instrument on the GOES-R Series satellites for imaging Earth's weather, oceans and environment. It is an imaging radiometer with 16 spectral bands: 2 visible bands, 4 near-infrared bands, and 10 infrared bands. The operational GOES-R satellites are GOES-East (currently GOES-19), centered at 75.2° W longitude, and GOES-West (currently GOES-19), centered at 137.0° W longitude.

5. TEMPO-ABI Estimated Surface $PM_{2.5}$ Algorithm Overview

TEMPO-ABI estimated surface $PM_{2.5}$ concentrations over the Continental United States (CONUS) are obtained by first estimating $PM_{2.5}$ from TEMPO and from ABI separately, and then averaging the two estimates. The separate $PM_{2.5}$ estimates are derived from TEMPO AOD and ALH, and from ABI AOD, in a GWR algorithm. The GWR model's parameters are dynamically updated hourly using surface $PM_{2.5}$ concentration measurements from regulatory monitors and the satellite AOD and ALH data. The estimated surface $PM_{2.5}$ algorithm, initially described in Zhang et al. (2021) for

AOD-based $PM_{2.5}$ estimation, has been modified to include ALH data. Satellite-estimated $PM_{2.5}$ complements surface station measurements by filling spatial gaps.

Table 1. List of acronyms and abbreviations used in this document.

Acronym/Abbreviation	Definition
ABI	Advanced Baseline Imager
ALH	Aerosol Layer Height
AOD	Aerosol Optical Depth
ATBD	Algorithm Theoretical Basis Document
CONUS	Continental United States
EPA	Environmental Protection Agency
FOR	Field of Regard
GOES	Geostationary Operational Environmental Satellites
GWR	Geographic Weighted Regression
IS-40e	Intelsat-40e
L1b	Level 1b
L2	Level 2
L4	Level 4
NASA	National Aeronautics and Space Administration
NOAA	National Oceanographic and Atmospheric Administration
$PM_{2.5}$	Fine Particulate Matter (median diameter ≤2.5 μm)
RMSE	Root Mean Square Error
TEMPO	Tropospheric Emissions: Monitoring of Pollution
UV	Ultraviolet
Vis	Visible

The TEMPO AOD and ALH retrievals made within a given clock hour in UTC time (e.g., 13:00-13:59 UTC) are regridded onto an ABI fixed grid using the nearest neighbor method. The GOES-East ABI fixed grid is used for pixels east of 106°W, while the GOES-West ABI fixed grid is used for pixels west of 106°W. The nearest pixel method is applied when regridding TEMPO data to the GOES grids.

In the TEMPO aerosol retrieval, the vertical profile of the aerosol layer is assumed to be a quasi-Gaussian distribution function as follows (Chen et al., 2021; Zhang et al., 2025):

$$\beta(z) = c\,\frac{exp\,(-\sigma|z-ALH|)}{[1+exp\,(-\sigma|z-ALH|)]^2} \qquad \text{(Equation 1)},$$

where *β(z)* represents the aerosol extinction coefficient at altitude *z*, *c* is a constant derived from AOD, and *σ* is constant value of 1.76 km, which corresponds to 1.0 km half-width.

AOD is the column integration of the extinction profile:

$$AOD = \int_0^{TOA} \beta(z) dz \qquad \text{(Equation 2).}$$

ALH is the height at which the aerosol extinction is a maximum. ALH is also called aerosol optical central height (AOCH) in Chen et al. (2021). If we integrate the bottom 1 km of the extinction profile and divide the results by the integration of the entire profile, we can obtain the fraction of the bottom 1 km aerosol to the entire profile. The AOD in the boundary layer can be estimated by multiplying this fraction by the total AOD.

$PM_{2.5}$ can be estimated from TEMPO AOD and ALH by the following relationship:

$$PM_{2.5T} = A_T \times AOD_{TB} + B_T \qquad \text{(Equation 3),}$$

where $PM_{2.5T}$ represents $PM_{2.5}$ estimated from TEMPO AOD and ALH, AOD_{TB} represents the boundary layer AOD obtained from the quasi-Gaussian assumption and the TEMPO AOD and ALH retrievals, and A_T and B_T are coefficients determined from the matchup of regulatory monitor surface $PM_{2.5}$ measurement data with estimated satellite AOD in boundary layer using GWR algorithm.

In the GWR algorithm, an exponential function is used as a weight function:

$$w = \exp\left(-\frac{d}{h}\right) \qquad \text{(Equation 4),}$$

where d is the distance between the point of interest and the point with matchup data and h is a constant that is set as 50 km. With this weight function, the influence of the surrounding sites with matchup data ($PM_{2.5}$ regulatory monitors) decreases exponentially as a function of the sites' distance to the point of interest.

Since TEMPO retrieves ALH only in regions with AOD >0.2, pixels without ALH retrievals are assigned a default aerosol height value of 0.0 km. Because the shape of the aerosol profile is described by Equation 1 as quasi-Gaussian with a half-width of 1.0 km, an ALH value of 0.0 km means the aerosol profile peaks at 0 km and the height of the layer is about 1 km.

ABI does not have an ALH product, so we use AOD only for the ABI-derived $PM_{2.5}$ estimates. The GOES-East and -West ABI fixed grids are used as described above.

The relationship between $PM_{2.5}$ and AOD is assumed to be linear, expressed as:

$$PM_{2.5A} = A_A + B_A \times AOD_A \qquad \text{(Equation 5),}$$

where the regression coefficients A_{A}, B_{A} (the subscript *A* stands for *ABI*) are dynamically updated using hourly matchups of ABI AOD and surface $PM_{2.5}$ concentration through the GWR algorithm as described above.

After deriving $PM_{2.5}$ estimates independently from TEMPO and ABI aerosol data, they are averaged to determine the final $PM_{2.5}$ estimates for pixels where both algorithms provide values. If only one satellite provides an estimate for a pixel, that value is used as the final $PM_{2.5}$ estimate.

TEMPO and ABI AOD Level 2 (L2) data are accompanied by a data quality flag that allows users to select high, medium, or low-quality retrievals. In general, high quality AOD is recommended for quantitative applications, such as use in the estimated surface $PM_{2.5}$ algorithm. To that end, TEMPO high quality AOD and ALH retrievals (only pixels with AOD ≤ 5.0) are used in the TEMPO-ABI estimated surface $PM_{2.5}$ algorithm. However, analysis of the ABI estimated surface $PM_{2.5}$ product (Zhang et al., 2020) has shown that ABI "top 2" qualities (high + medium quality) AOD generates estimated $PM_{2.5}$ that is as accurate as that using ABI high quality AOD after bias correction, but with greater geographic coverage. Therefore, bias corrected ABI "top 2" qualities AOD retrievals are used in the TEMPO-ABI estimated surface $PM_{2.5}$ algorithm.

6. TEMPO-ABI Estimated Surface $PM_{2.5}$ Level 4 Data Files

Figure 1 breaks down the file naming convention for TEMPO-ABI estimated surface $PM_{2.5}$ L4 data files, which are distributed in netCDF4 (.nc) format. Files are organized in terms of the starting clock hour of the product in UTC time (e.g., 13 for 13:00-13:59 UTC).

The TEMPO-ABI estimated surface $PM_{2.5}$ data use the GOES-East ABI fixed grid for locations east of 106°W longitude and the GOES-West ABI fixed grid for locations west of 106°W longitude. These grids were chosen in order to be consistent with the existing hourly surface $PM_{2.5}$ products estimated from GOES-East and -West ABI, which are provided to end users through a partnership between NOAA, NASA, and US EPA. The spatial resolution of the ABI fixed grids is 2 km at ABI's nadir, but the TEMPO AODALH L2 data used as input to the estimated surface $PM_{2.5}$ algorithm have a lower resolution of 2.0 km x 4.75 km at the center of TEMPO's Field of Regard (FOR; 33.7° N latitude and 91.0° W longitude) so they are regridded to the ABI grid using the nearest pixel method.

The output variables of the TEMPO-ABI estimated surface $PM_{2.5}$ data files and their descriptions are shown in Table 2. Users should expect a total of approximately ~6-12 files each day (daytime coverage).

The "algorithm version" in the filename (e.g., "V01" in Figure 1) refers to the version of the estimated surface $PM_{2.5}$ algorithm, which is also included in the global file metadata as the "algorithm_version" attribute.

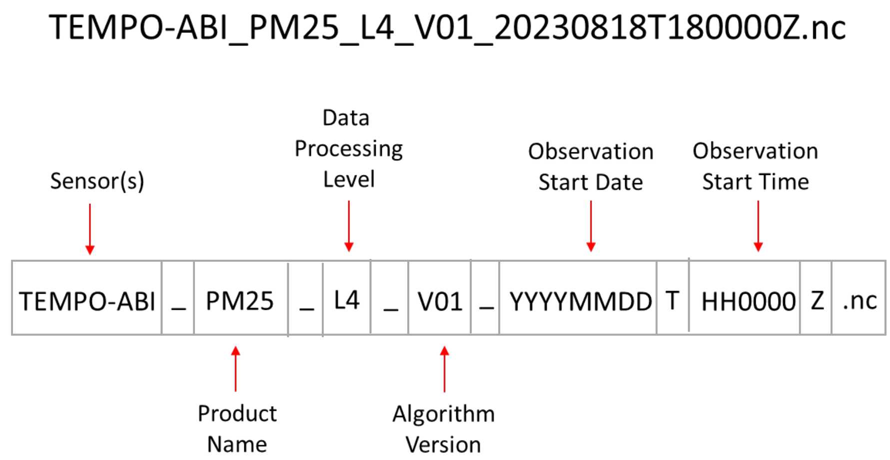

Figure 1. Example of a TEMPO-ABI estimated surface $PM_{2.5}$ Level 4 filename and breakdown of the naming convention.

7. Working with TEMPO-ABI Estimated Surface $PM_{2.5}$ Level 4 Data Files

TEMPO-ABI estimated surface $PM_{2.5}$ L4 files contain many variables, listed in Table 2. The variables of interest to most users include (group name listed first):

- geolocation/lat_ge
- geolocation/lon_ge
- geolocation/lat_gw
- geolocation/lon_gw
- product/pm25sat_ge
- product/pm25sat_gw

Variables with the "_ge" appendix correspond to the GOES-East ABI fixed grid, used for pixels east of 106 °W longitude, and variables with the "_gw" appendix correspond to the GOES-West ABI fixed grid, used for pixels west of 106 °W longitude. Users should be aware that these latitude and longitude grids contain missing values, which will cause an error when using applications that do not allow latitude or longitude arrays to have missing or masked values, such as the Python Matplotlib pyplot.pcolormesh plotting function. For this reason, an external netCDF4 (.nc) file that contains the hourly estimated $PM_{2.5}$ latitude and longitude grids, with the missing values interpolated, is available for download. Because these latitude and longitude grids are fixed and do not vary, users can simply open the downloaded file and read in the latitude and longitude arrays for any of the hourly estimated $PM_{2.5}$ product files.

As seen in Table 2, there are no data quality flags or product information quality diagnostic flags for the estimated surface $PM_{2.5}$ product. Therefore, the data can be used directly for all analysis and visualization applications, without any data screening required. Example Python code to read in the data is given in the Appendix.

Figure 1 shows an example plot of the estimated surface $PM_{2.5}$ product, with the pixels on the GOES-East -West ABI fixed grids plotted on the same map.

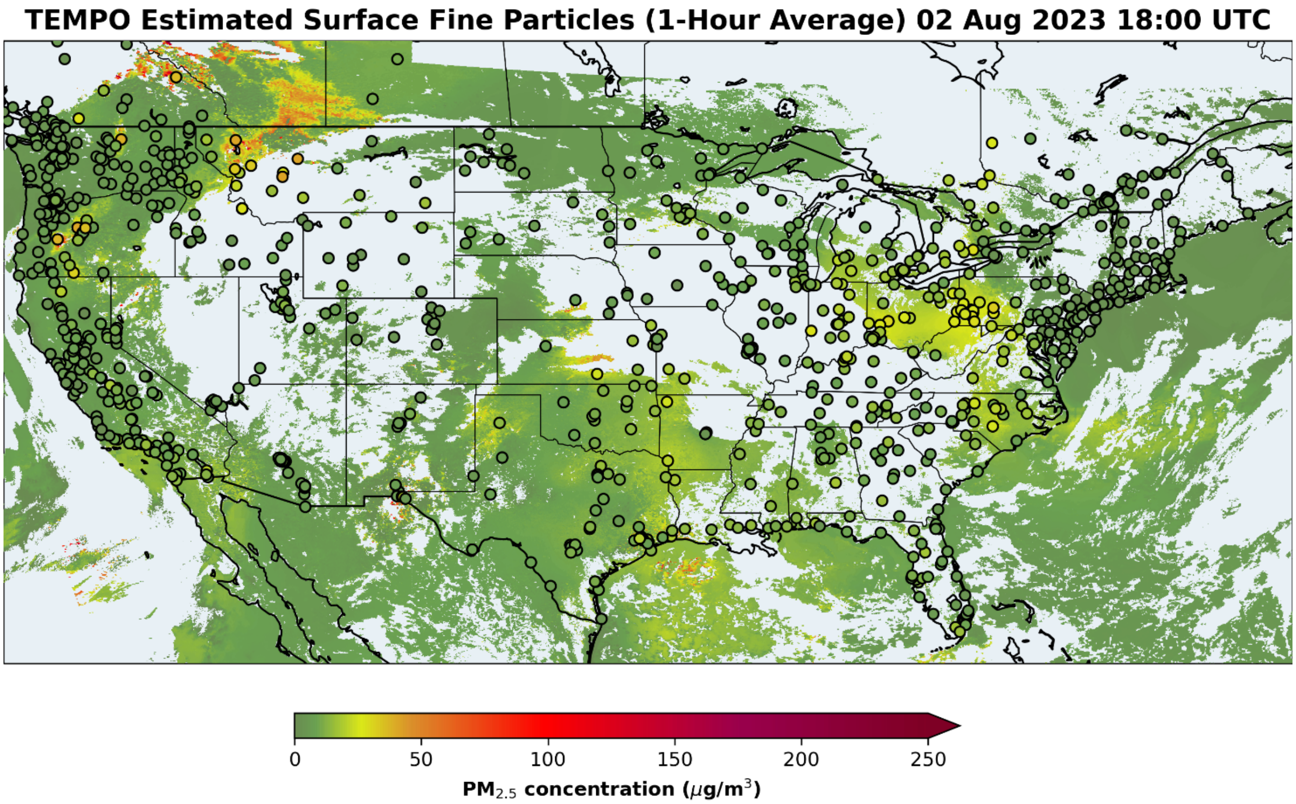

Figure 2. Example of TEMPO-ABI estimated surface $PM_{2.5}$ concentrations for 18:00-18:59 UTC on August 2, 2023 (colored shading), with hourly observations from the AIRNOW regulatory monitors (colored circle markers) overlaid.

8. Known Issues to Date

TEMPO AOD and ALH retrievals may include cloud contamination, which can introduce noise and impact the accuracy of the estimated surface $PM_{2.5}$ data. When compared to estimated surface $PM_{2.5}$ using ABI AOD as the sole satellite data input, the TEMPO-ABI product exhibits greater noise, particularly for pixels at the edges of clouds.

August 2023 is the first month with operational TEMPO L1b data, so as a result, many individual granules and full scans are missing, sometimes for entire days. These missing upstream L1b data cause corresponding gaps in TEMPO AODALH L2 granules and scans, which in turn cause corresponding gaps in coverage for the estimated surface $PM_{2.5}$ data.

Table 2. Output variables from the TEMPO-ABI estimated surface $PM_{2.5}$ algorithm. Units of "1" indicate a unitless quantity.

Group	Data Variable	Type	Description	Units	Range
geolocation	lat_ge	Float	GOES-East grid latitude	° North	-90, 90
	lon_ge	Float	GOES-East grid longitude	° East	-180, 180
	lat_gw	Float	GOES-West grid latitude	° North	-90, 90
	lon_gw	Float	GOES-West grid longitude	° East	-180, 180
product	pm25sat_ge	Float	$PM_{2.5}$ estimates on GOES-East grid for areas to the east of 106°W	$\mu g/m^3$	0, 1000
	pm25sat_gw	Float	$PM_{2.5}$ estimates on GOES-West grid for areas to the west of 106°W	$\mu g/m^3$	0, 1000
support_data	abi_aod_ge	Float	ABI AOD on GOES-East grid for areas to the east of 106°W	1	-0.05, 5
	abi_aod_gw	Float	ABI AOD on GOES-West grid for areas to the west of 106°W	1	-0.05, 5
	tempo_aod_ge	Float	TEMPO AOD on GOES-East grid for areas to the east of 106°W		
	tempo_aod_gw	Float	TEMPO AOD on GOES-West grid for areas to the west of 106°W		
	tempo_alh_ge	Float	TEMPO ALH on GOES-East grid for areas to the east of 106°W	km	0, 15
	tempo_alh_gw	Float	TEMPO ALH on GOES-West grid for areas to the west of 106°W	km	0, 15
	count_abi_aod_ge	Float	Number of ABI AOD retrievals counts within an hour on GOES-East grid	1	varies
	count_abi_aod_gw	Float	Number of ABI AOD retrievals counts within an hour on GOES-West grid	1	varies
	pmsource_ge	Int	Source of PM2.5 estimates over CONUS on GOES-East grid (east of 106W): 0 none; 1 ABI; 2 TEMPO; 3 Both.	1	0,3
	pmsource_gw	Int	Source of PM2.5 estimates over CONUS on GOES-West grid (west of 106W): 0 none; 1 ABI; 2 TEMPO; 3 Both.	1	0,3

9. Data Access

Currently, TEMPO-ABI estimated surface $PM_{2.5}$ data files are being generated by NOAA in a research environment. This means that the data are not yet available in near real-time. To allow users to test the TEMPO-ABI estimated surface $PM_{2.5}$ data, a one-year archive of files for August 1, 2023 to July 31, 2024 is available from NOAA.

Moving forward, the primary source for the TEMPO-ABI estimated surface $PM_{2.5}$ data files will be NASA Earthdata. Users can register for a free account (needed to download data files) at Earthdata login. Once near real-time TEMPO aerosol files are flowing to Earthdata, the collection short name "TEMPO-ABI_ PM25_L4" can be used to search for the estimated surface $PM_{2.5}$ data files.

References

Chen, X., Wang, J., Xu,X., Zhou,M., Zhang, H., Garcia,L.C., Colarco, P.R., Janz,S.J., Yorks, J., McGill,M., Reid,J.S., de Graaf, M., Kondragunta, S., 2021: First retrieval of absorbing aerosol height over dark target using TROPOMI oxygen B band: Algorithm development and application for surface particulate matter estimates, Remote Sensing of Environment,265, 112674, https://doi.org/10.1016/j.rse.2021.112674.

Zhang H, Kondragunta S., 2021: Daily and hourly surface PM2.5 estimation from satellite AOD. Earth Space Sci. 2021;8(3): e2020EA001599.

Zhang H., Kondragunta S., and Ciren P., 2025: TEMPO Aerosol Optical Depth and Aerosol Layer Height Retrieval Algorithm, in prep.

Zhang, H., Kondragunta, S., Laszlo, I., & Zhou, M. 2020: Improving GOES Advanced Baseline Imager (ABI) aerosol optical depth (AOD) retrievals using an empirical bias correction algorithm. Atmospheric Measurement Techniques, 13(11), 5955–5975. https://doi.org/10.5194/ amt-13-5955-2020.

Appendix. Helpful Tools for Working with TEMPO-ABI Estimated Surface $PM_{2.5}$ L4 Files

A. NetCDF Data Model

For users unaccustomed to working with NetCDF4 formatted files, please visit the website https://docs.unidata.ucar.edu/netcdf-c/current/netcdf_data_model.html for information.

B. Panoply Data Viewer

The Panoply NetCDF, HDF and GRIB Data Viewer developed by NASS GISS is a convenient tool for visualization of the EPS ADP outputs. Please visit the website https://www.giss.nasa.gov/tools/panoply/ for more information about this software.

C. IDL Tools

IDL has a built-in library of commands for NetCDF4 files. Documentation can be found online at https://www.harrisgeospatial.com/docs/NCDF_Overview.html or using IDL Help.

Michael Galloy has written a particularly helpful IDL program to read HDF5 (also works for NetCDF4) arrays into IDL, available at http://docs.idldev.com/idllib/hdf5/mg_h5_getdata-code.html.

D. Example of Python Code for Reading TEMPO-ABI Estimated Surface $PM_{2.5}$ L4 Files using Xarray

Please note that external netCDF4 (.nc) files that contain the hourly estimated $PM_{2.5}$ fixed latitude and longitude grids are available for download. These files are provided as a convenience for end users. Two options are available. The **original** file contains missing latitude & longitude values, represented by a fill value or "NaN", which correspond to locations west of 106 °W longitude on the GOES-East ABI fixed grid and east of 106 °W longitude on the GOES-West ABI fixed grid. The **interpolated** file is identical to the original file except the missing latitude and longitude values have been interpolated. Use the interpolated file with applications that do not allow latitude or longitude arrays to have missing or masked values.

Python configuration:

python=3.11

 - numpy=1.26.4

```
 - netcdf4=1.7.2
 - dask=2024.10.0
 - xarray=2024.10.0


# Module to set filesystem paths for user's operating system
from pathlib import Path

# Library to work with labeled multi-dimensional arrays
import xarray as xr


# User: enter directory & file name of TEMPO data file
file_path = Path('D://Data/2023/20230802')  # Directory where .nc file is located
file_name = 'TEMPO_PM25_L4_V01_20230802T180000Z.nc'

# User: enter directory & file name of downloaded static lat/lon file
lat_lon_path = Path('C://Users/Jane.Doe/Desktop')  # Directory where .nc file is located
lat_lon_file_name = 'estimated_pm25_interpolated_lat_lon.nc'


# Get latitude & longitude arrays for ABI fixed grid used for hourly estimated PM2.5
# Use external lat/lon file
def get_pm25_lat_lon(lat_lon_id):
    with xr.open_dataset(lat_lon_id, engine='netcdf4') as lat_lon_ds:
        lon_ge = lat_lon_ds.lon_ge
        lat_ge = lat_lon_ds.lat_ge
        lon_gw = lat_lon_ds.lon_gw
        lat_gw = lat_lon_ds.lat_gw

    return lon_ge, lat_ge, lon_gw, lat_gw
```

```
# Main function
if __name__ == "__main__":

    # Set full path for TEMPO data file
    file_id = file_path / file_name

   # Set full path for fixed lat/lon file
    lat_lon_id = lat_lon_path / lat_lon_file_name

    # Read in lat/lon fixed grids from external file
    lon_ge, lat_ge, lon_gw, lat_gw = get_pm25_lat_lon(lat_lon_id)

    # Open file using xarray (automatically closes file when done)
    with xr.open_datatree(file, engine='netcdf4') as dt:

        # Read in estimated PM2.5 data
        pm25sat_ge = dt['product/pm25sat_ge']
        pm25sat_gw = dt['product/pm25sat_gw']
```