User's Guide: XAERDT aerosol products in support of MEaSUREs

In response to NASA's 2017 Research Opportunities in Space and Earth Science (ROSES), the Dark Target team proposed to the Making Earth System Data Records for Use in Research Environments (MEaSUREs) Program. Our overall goal was to port the Dark Target (**DT**) aerosol retrieval algorithm to sensors in geostationary (GEO) orbit, so that by combining with retrievals on traditional sensors in low-earth-orbit (LEO), we would be creating a GEO-LEO synergy dataset (known as **XAERDT**) that could help characterize rapid aerosol changes and the aerosol diurnal cycle. The abstract of that proposal is linked here: https://www.earthdata.nasa.gov/esds/competitive-programs/measures/leo-geo-synergy. The XAERDT products created from native resolution observations within the field-of-view of a sensor ("granules" for LEO, "Full Disks" for GEO) are known as Level 2 (L2), whereas gridded products (in time and space) aggregated from L2 are known as Level 3 (L3). This User's Guide accompanies the Version 1.0 L2 products known as **XAERDT_L2**. Currently, XAERDT_L2 includes retrievals from observations made between 2019 and 2022.

This guide will be updated once the Level 3 algorithm (gridding for 0.25°x0.25° at 30-minute intervals) and products (XAERDT_L3) are completed.

1. Introduction

There are several NASA satellite remote sensing aerosol products available for research and public use. This users guide focuses on the L2 aerosol products created using the Dark Target (**DT**) algorithm(s), in response to the MEaSUREs program. To distinguish from standard DT products derived from MODIS (e.g. Collection 6.1) or VIIRS (Version 2.0), the MEaSUREs related products are collectively known as XAERDT. Details of the DT algorithm are being updated within the Algorithm Theoretical Basis Document (ATBD), soon available from NASA's Dark Target website https://darktarget.gsfc.nasa.gov/.

All satellite remote sensing aerosol products must contend with the problem of separating the aerosol signal from the surface signal observed by the sensor. The DT algorithm relies on the phenomena that aerosols over a dark surface target generally brighten the observed scene. This contrast, if observed across multiple wavelength bands, offers the possibility of retrieving the total column aerosol optical depth (AOD) and some information about relative aerosol size via spectral dependence. Therefore, DT requires Visible (VIS), near-infrared (NIR), shortwave-infrared (SWIR), and thermal infrared (TIR) observations. In fact, the DT-algorithm is made up of two sub-algorithms, one for retrieving aerosols over ocean (DT-O), the other for retrieval over dark surfaces (vegetation, soils) over land (DT-L). In general, because deep water absorbs nearly all radiation in wavelengths red and longer (e.g. NIR and SWIR), DT-O is more sensitive to aerosol than DT-L.

Where the surface is bright, such as over ocean glint, desert areas, or snow/ice, the DT algorithm will not retrieve aerosol properties. Other algorithms such as Deep Blue (DB) or Multi-Angle implementation of Atmospheric Correction (MAIAC) may be applicable, but are not discussed here.

The DT aerosol retrieval algorithm was originally developed for deriving aerosol properties over land and ocean using the along-orbit observations of the Moderate-resolution Imaging Spectroradiometer (MODIS), aboard Terra (since 2000) and Aqua (since 2002) satellite platforms. Both of these platforms are in polar, sun-synchronous, low-earth-orbit (LEO), passing over the equator at 10:30 local-solar-time (LST) and 13:30 LST, respectively. Each MODIS observes from 705 km, so that with ~14.5 orbits/day, MODIS samples a 2300 km wide swath that nearly covers the whole Earth each day. Based on concerns over the aging of Terra and Aqua, the DT team began porting the algorithm to newer sensors with similar and sufficient VIS/NIR/SWIR/TIR wavelength bands. This includes the Visible-Infrared Imaging Radiometer Suite (VIIRS) on the Suomi-NPP (SNPP, since 2011) platform and the NOAA-20 (since 2017; formerly known as JPSS-1) platform. Both of these platforms also pass over the equator near 13:30 LST, but by observing from 820 km and ~14 orbits/day, VIIRS samples a 3000 km wide swath that entirely covers the whole Earth each day.

With 1x day sampling for each MODIS or VIIRS, leading to at most 2x or 3x combined samplings at any midlatitude or tropical site, NASA's LEO imagers are insufficient for estimating diurnal cycles and watching rapid development/transport of severe dust or smoke events. However, with regional views that cover thousands of km, sufficient VIS/NIR/SWIR/TIR wavelengths, and 10-minute observing cadence, modern imagers in GEO fill the gaps in some regions. Since 2017, Advanced Baseline Imagers (ABI) are flying on NOAA's Geostationary Operational Environmental Satellite – R series (GOES-R), with one in GOES-East position (over equator at nominal -75.0° longitude), and one in GOES-West (-137.0°). The GOES-East slot is currently occupied by GOES-16 (formerly GOES-R), and between 2018-2022 the GOES-West slot was occupied by GOES-17 (formerly GOES-S). Advanced Himawari Imagers (AHI) are flying on Japan's Himawari series, located near +140.7°, where Himawari-9 replaced Himawari-8 on 13 December 2022.

Together, 6 sensors (MODIS on Terra and Aqua, VIIRS on SNPP, ABI on GOES-East and West, and AHI on Himawari) observe the entire day over most of the western hemisphere, and some of the eastern (Asia Pacific region). A schematic of such sampling is given by Fig 1. Where one sensor's aerosol retrieval may be obscured by clouds or sunglint, another may fill in the gaps. Together, they maximize sampling, leading to a GEO-LEO view of the globe.

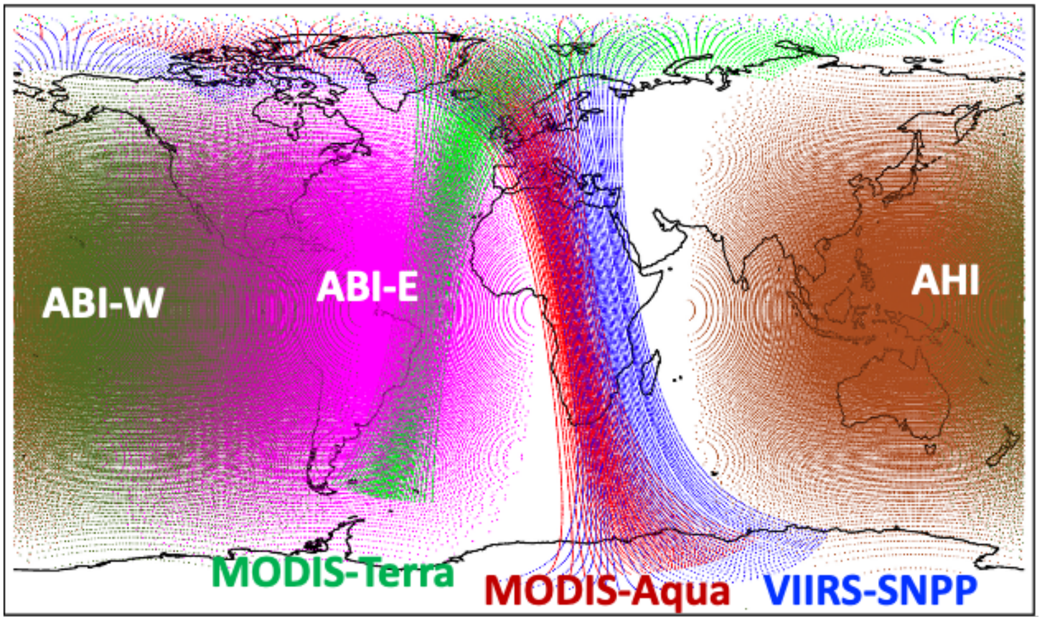


Figure 1: GEO-LEO view of the globe, created by joining ABIs on GOES-West and GOES-East, AHI on Himawari, MODIS on Terra and Aqua, and VIIRS on Suomi-NPP.

One of the original motivations for this project was to create a dataset suitable for creating meteorological and chemical reanalyses that combine two NASA field campaigns, specifically Fire Influence on Regional to Global Environments and Air Quality (FIREX-AQ; https://csl.noaa.gov/projects/firex-aq/) held over the western U.S. during summer 2019, and the Cloud, Aerosol and Monsoon Processes Philippines Experiment (CAMP2Ex; https://espo.nasa.gov/camp2ex/content/CAMP2Ex), over southeast Asia during the late summer/fall 2019. Thus, an initial goal was to produce a 6-sensor aggregation that included at least the period May-October 2019. As the ability to process data versus the need for consistent cadences and calibration changed over time, plus some interesting events occurring (global wildfire and dust events), led to a goal of covering the four years of 2019-2022. With new sensors (e.g. GOES-18 replacing GOES-17 as GOES-West), and if/when additional sensors and/or their data become available in the near future (e.g. Advanced Meteorological Imagers (AMI) over Korea, Flexible Combined Imager (FCI) over Europe, over India, China, etc.), these sensors can be processed via the DT algorithm, and entered into the GEO-LEO framework.

This guide focuses on XAERDT_L2 products from 6 sensor/platforms (MODIS x 2, VIIRS, ABI x 2 and AHI), each retrieved from that sensor's native spatial/temporal resolution Level 1B data. Created for sensor/platform, they can be searched as XAERDT_L2_Sensor_Platform. For the entire period of 2019-2022, this means retrievals from Sensor_Platform of MODIS_Terra, MODIS_Aqua, VIIRS_SNPP, ABI_G16 and ABI_G17. For AHI, this means AHI_H08 for 1 Jan 2019 through 13 Dec 2022, with AHI_H09 for the remaining 17 days of 2022. As discussed in section XX, in order to use a consistent algorithm for all sensors, there are some differences between XAERDT for MODIS_Terra and MODIS_Aqua, versus the current standard MODIS product (Collection 6.1) known as MOD04_L2 and MYD04_L2 (appendix XX). For VIIRS, the XAERDT_L2_VIIRS_SNPP products are nearly identical to the standard VIIRS

(AERDT_L2_VIIRS_SNPP) products, except for the computer architecture under which they have been processed. Although not originally under the MEaSUREs proposal, XAERDT_L2 products for VIIRS on NOAA20 (XAERDT_L2_VIIRS_NOAA20) are archived for completeness.

Fig 2 shows an example of coverage attained by each sensor, within ±15 minutes of March 26 2020, at 23:45 UTC. The middle panel demonstrates the advantage of merging together all products, which is the purpose of the eventual Level 3 product (XAERDT_L3).

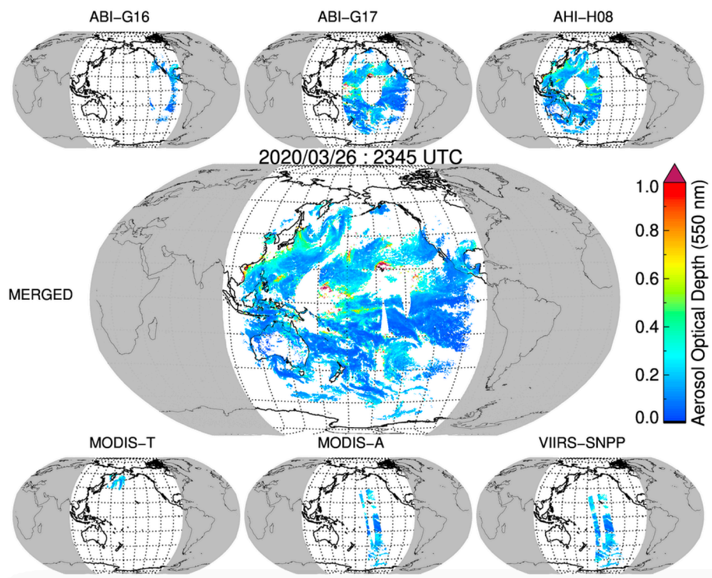


Figure 2: Example of 30 minutes at quarter degree resolution coverage by each sensor and merged AOD datasets on March 26, 2020, at 2345 UTC.

2. Sensors/Platforms

Table 1 provides information characteristics of the sensors used for XAERDT retrieval, including the dates of mission, equator crossing time (if LEO) or the longitude at the equator (if GEO), their swath width, number of wavelengths, spatial resolution of the blue (lower resolution) or red (often higher resolution) wavelength bands, the nominal granule or full disk (FD) pixel size.

Table 1: Properties of Sensors used and DT products archived as XAERDT_L2

	LEO (Granule)		**GEO (Full Disk)**	
Sensor	**MODIS**	**VIIRS**	**ABI**	**AHI-H08/H09**
Dates processed	2019-2022	2019-2022	2019-2022	H08: 2019 - 13 Dec 2022 H09: 14 - 31 Dec 2022
Platform Eq crossing Time for LEO or Longitude for GEO	Terra: 10:30 (descend) Aqua: 13:30 (ascend)	SNPP: 13:30 (ascend) NOAA20: 13:30 (ascend)	G16 (East): 75.0°W G17 (West): 137.0°W	H08: 140.7°E H09: 140.7°E
Altitude (km)	705	820	36K	36K
Orbits/day	14.5	14	1	1
# obs of Ground target at equator	<1/ day	1/day	144/day	144/day

Swath width (km)	2330	3040	Full Disk	Full Disk
Field of view	±55°	±56°	±8°	±8°
Number of Wavebands	36	22	16	16
Blue / Red band pixel size at nadir (km)	0.5 / 0.25	0.75 / 0.375	1.0 / 0.5	1.0 / 0.5
Granule or FD X × Y size in pixels	2708 × 4080	3200 × 3232	10884 × 10884	11000 × 11000
Granule or FD time length (mins)	5	6	10	10
Scan lines per scan	20	8	1000	1000
BowTie?	Yes	Corrected	No	No
N×N aggregation (of blue band)	20×20	8×8	10×10	10×10
XAERDT_L2 Product resolution at nadir (km)	10	6	10	10
XAERDT_L2 Product X × Y size	135 × 204	400 x 404	1088 x 1088	1110 x 1110

2.1. MODIS

MODIS flies on the Earth Observation System's (EOS) Terra and Aqua satellites. Both satellites are in polar-orbiting, sun-synchronous, Low Earth Orbit (LEO), with Terra descending (southward) over the equator about 10:30 local sun time (LST), and Aqua ascending (northward) over the equator about 13:30 LST. From a vantage 705 km above the surface and a ±55° view scan, each MODIS instrument views a swath about 2330 km. Each day, MODIS makes ~14.5 orbits (99 minutes per orbit), observing nearly the entire globe. There are small gaps at the equator, so full coverage takes about 3 days. Orbit patterns repeat every 16 days.

MODIS performs measurements in 36 spectral channels (or bands) that cover the solar to thermal infrared spectrum region between 0.41 to 14.2 µm (Salomonson et al., 1989). Nominal pixel resolutions (at nadir) are 0.25 km (for 2 bands), 0.5 km (for 5 bands) and 1 km (for 29 bands). Detailed specifications and components can be found at http://modis.gsfc.nasa.gov. For the DT aerosol retrieval, we rely on window (small gas absorption) Bands 1 through 7 (B1-B7), with "centroid" wavelengths approximately 0.47, 0.55, 0.65, 0.86, 1.24, 1.64 and 2.11 µm, respectively. B1 and B2 (red and NIR) are observed at 0.25 km resolution, with B3-B7 (blue, green, SWIR1, SWIR2 and SWIR3) at 0.5 km resolution. For cloud masking and other purposes, DT uses 1 km "cirrus" band (B29 near 1.38 µm) as well as wavelengths in the TIR.

Note that MODIS is not a "camera", rather it makes use of a continually rotating scan mirror. Each scan of the mirror images 10 lines of 1 km pixels, (20/40 lines of 0.5/0.25 km pixels). Because of the 55° swath convolved with Earth's curvature, 1 km pixels grow to approximately 4.8 by 2.0 km at swath edge. This gives rise to the geolocational oddity known as the panoramic "bow-tie" effect that means the scans are partially overlapping towards swath edge. This phenomena is not corrected for during XAERDT retrieval.

MODIS observation data are split into 5-minute granules, which represent an along-track path of 2030 km (203 "scans"). For XAERDT Version 1.0, MODIS observation inputs include the Collection 6.1 versions of Level 1B (L1B) reflectance/radiance, known as MxD021KM, MxD02HKM and MxD02QKM for 1km, 0.5km and 0.25km resolution channels, respectively, where the x=O for MODIS on Terra, and x=Y for MODIS on Aqua. Geolocation information are provided in files known as MxD03. All L1B data are calibrated and produced via NASA's MODIS Adaptive Processing

System (MODAPS) and can be found in standard archives at NASA's Land and Atmosphere Distribution System (LAADS).

2.2. VIIRS

The first VIIRS was launched aboard Suomi-NPP in 2011, on NOAA-20 in 2017, on NOAA-21 in 2022, and on future NOAA Joint Polar Satellite System (JPSS-3 and 4) satellites. All VIIRS copies are at elevation of ~825 km with ascending equator crossings around 13:30 LST. With a higher altitude and slightly wider view ±56°, each VIIRS observes a 3040 km wide swath which entirely covers the globe (no gaps) with ~14 orbits per day. NOAA-20 is placed a half-orbit behind SNPP (50 minutes), so crosses the equator at the same LST but to the west.

Like MODIS, VIIRS is multi-spectral, but with only 22 bands (covering 0.41 to 12.3 µm; https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/viirs/). Nominal spatial resolution for the 5 "Imagery bands" (I-bands; I1-I5) are at 0.375 km, with the remaining "Moderate bands" (M-bands; M1-M16) and the DayNightBand (DNB) at 0.75 km. The DT retrieval (in daylight) makes use of wavelength bands analogous to those on MODIS. These include window bands (M3, M4, M5, M7, M9, M10 and M11, near 0.49, 0.55, 0.67, 0.86, 1.24, 1.61 and 2.26 µm), cirrus reflective band (M8 near 1.38 µm), and TIR bands. Note that for VIIRS cloud masking, DT makes use of the I2 (red wavelength I-band near 0.64 µm). Note that unlike MODIS on Terra and Aqua, the copies of VIIRS on SNPP and NOAA-20 differ slightly in their wavelength spectral response, leading to differences of 3-5 nm in some bands.

VIIRS is similar in technology to MODIS in that it uses a scanning mirror, and is hampered by bow-tie effect and pixel overlap. However, VIIRS onboard processing is such that it partially compensates for the bow-tie effect, by deleting bow-tie influenced scans through a "pixel trim" (https://www.star.nesdis.noaa.gov/jpss/documents/AMM_All/VIIRS_SDR/Provisional/VIIRS_USERS_GUIDE_TechReport142FINAL_cc_rdlnCmts_cc_020192013.pdf). For SNPP these data cannot be recovered, but they can be estimated via interpolation.

NASA's VIIRS L1B data are provided as 6-minute granules, also archived at LAADS. The M-band resolution (0.75 km) geolocation and radiance are known as Vx03MOD and Vx02MOD (x=VNP for SNPP and x=VJ1 for JPSS1/NOAA20) and the I-band files (0.37 km) are Vx03IMG and Vx02IMG. Note that these L1B data differ than those offered at NOAA's Comprehensive Large Array-Data Stewardship System-CLASS, due to following NASA protocols for calibration and granularity.

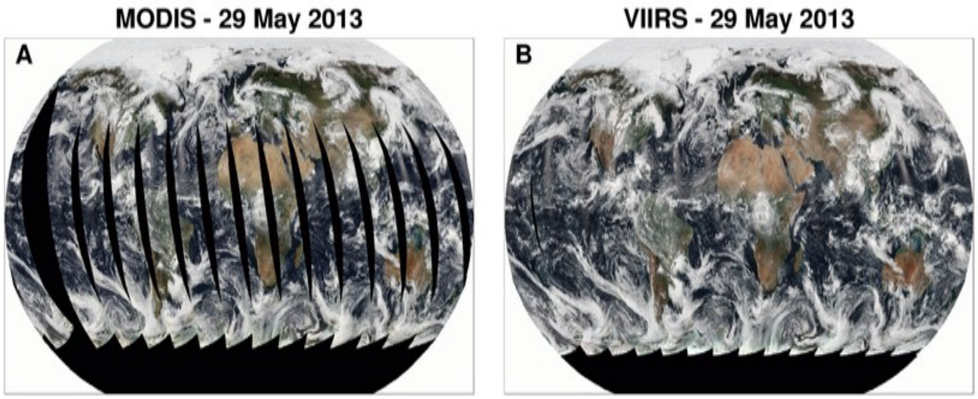


Figure 3: The true color images above taken from each sensor show one full day of data coverage. MODIS (on Aqua) has gaps in its coverage near the equator. VIIRS' wider swath width eliminates these gaps.

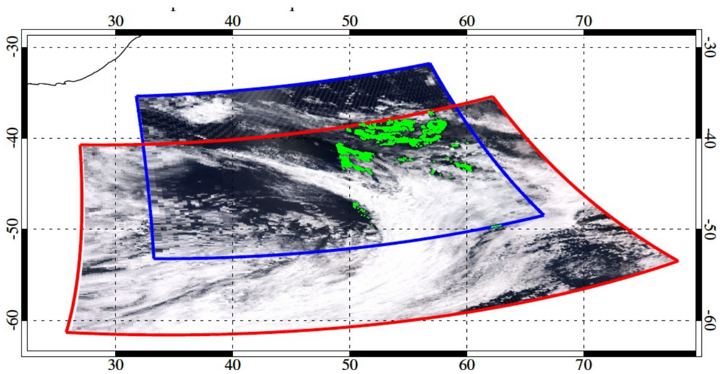


Figure 4: . Example of a MODIS 5-minute (blue outline) vs VIIRS 6-minute (red outline) granule coverage. (Adopted from Sayer et al., 2017; https://doi.org/10.5194/amt-10-1425-2017)

2.3. ABI

ABIs were launched upon GOES-R (became GOES-16 on orbit) in 2017, GOES-S (GOES-17) in 2019, and GOES-T (GOES-18) in 2022. GOES-16 became NOAA's operational GOES-East (over equator at 75.0°W), whereas GOES-17 became GOES-West (137.0°W). After a period of overlap at GOES-West position between June 2022 and December 2022, GOES-18 has been the operational GOES-West since Jan 2023. From an altitude of ~36,000 km, each ABI carves out a scan pattern that includes mesoscale, continental, and full-disk (FD) defined areas. Prior to April 2019, the scan pattern took 15 minutes, but is now 10 minutes (https://www.goes-r.gov/users/abiScanModeInfo.html). Each ABI makes 144 FD scans per day, observing hemisphere with radius approximately 82° in longitude/latitude (local zenith angle). Currently, DT is run only on FD.

ABI observes at 16 wavelength bands, covering a range of VIS though TIR. Analogous to MODIS or VIIRS, ABI has blue, NIR and SWIR2 bands (near 0.47, 0.86 and 1.61 μm) at 1 km resolution, red band (near 0.64 μm) at 0.5 km resolution, and SWIR3 band (near 2.24 μm) at 2 km spatial resolution. There are also 1.37 μm "cirrus" and TIR bands used for cloud masking and pixel selection. ABI has no Green or SWIR1 bands, so the DT algorithm must include compensations. Each FD image is approximately 10884 × 10884 pixels (for the 1 km resolution). Pixel-overlap (bow tie) is not a problem for a GEO imager, however, pixel sizes toward FD limbs increase greatly.

For XAERDT, DT retrieval uses the L1B files as offered through NOAA's CLASS. Rather than named by resolution as MODIS or VIIRS, they are named as 'OR_ABI-L1b-RadF-M6Ccc_Gxx' where 'cc' is the GOES-ABI band number and 'xx' is '16' or '17' (or '18' after 2023). Note that there is no specific "geolocation" file like there is for MODIS or VIIRS, so that the geolocation (latitude/longitude, observation geometry/angles) are calculated for each FD image. Note that that DT algorithm uses the data "as is", with no additional calibration adjustments; DT products may be impacted by discrete updates or adjustments on the operational side (https://www.star.nesdis.noaa.gov/GOESCal/goes_SatelliteAnomalies.php).

2.4. AHI

AHIs were launched on the Japanese Meteorological Agency's (JMA) Himawari-8 (H08) satellite in 2015 and on Himawari-9 (H09) in 2016. Both are located near 140.7°E. H08 was operational until December 2022 when it was replaced by H09. H08 is currently in standby orbit. Like ABI, AHI carves out a scan pattern that includes FD imagery every 10 minutes. AHI also observes 16 bands similar to ABI in wavelength and spatial resolution, however AHI replaces the 1.37 μm "cirrus" with a 0.51 μm Green band. The FD is slightly larger, covering 11000 × 11000 pixels (for 1km resolution).

AHI data are processed via JMA's ground system, leading to files in Himawari Standard Data (HSD) format. Through agreements between NOAA, NASA and JMA, these HSD are available for U.S. research and forecasting. Based on readers provided by JMA, the University of Wisconsin's Space Science and Engineering Center (SSEC) has developed a tool (nominally known as HSD2NC), that creates files analogous to the ABI files provided via NOAA-CLASS. These look like 'OR_AHI-L2-CMIPF-M1Ccc_GHx' (where the cc is the AHI band number and 'x' is '8' or '9'). Using the Himawari projection longitude (140.7°E) a code is used to derive the geolocation information. Like ABI data, DT products are created from AHI L1B "as is", and they may be impacted by any updates or adjustments on the operational side (https://www.data.jma.go.jp/mscweb/en/oper/event_H8.html).

3. DT-Package, inputs, ancillary files, and processing environments

The XAERDT_L2 products are created using a consistent deployment of the DT retrieval algorithm, known as the **DT-Package**. Details of ancillary files, lookup tables (LUTs), and cloud masking are being updated via a generalized ATBD that will be available soon. However, all instances of DT-Package for XAERDT_L2 share common themes, including

- All aerosol LUTs are derived using the same combination of particle scattering and Radiative Transfer (RT) codes. Aerosol optical and physical properties are assumed the same regardless of sensor being used, so that assumed size and refractive index of a given aerosol "type" represent the wavelength region (e.g. "blue", "red", "NIR", "SWIR3", etc.). However details, like the details like absorption by gases (water vapor, ozone, trace gases) and Rayleigh (molecular) scattering, represent the exact wavelength range of a given sensor.
- Maps describing expected aerosol type over land (seasonal at 1°x1°) are invariant between

sensors.

- All retrievals use the highest spatial resolution red band imagery for cloud masking, along with consistent tests using other bands. If an upstream cloud masking (known as the Wisconsin cloud mask) exists (as for MODIS and VIIRS), then use some of the internal tests.
- For ancillary meteorological assumptions (gas concentration columns, wind speed over water), use the Goddard Modeling and Assimilation Office (GMAO) provided re-analysis from the Goddard Earth Observing System (GEOS), valid at the time of observation.

The DT-Package is modular and generally free of system-dependent libraries. It has been tested under multiple architectures, and efficient processing of 3 LEO + 3 GEO sensors required that it be run across multiple platforms. DT-Package can read files in HDF format (used MODIS) as well as the now-standard Network Common Data format version 4 (NetCDF4) that are used for VIIRS, ABI and AHI. For all sensors (including MODIS), DT-Package writes as NetCDF. For XAERDT_L2, Version 1.0, MODIS and ABI retrievals are run at NASA's MODAPS, VIIRS run at the University of Wisconsin-Madison's Atmosphere- Science Investigation Processing and Support (A-SIPS), and AHI run on a separate system within the UW-Madison's SSEC.

4. XAERDT Aerosol Products

XAERDT products are considered to be under NASA's "Atmosphere" discipline, so they are archived at the Level-1 and Atmosphere Archive & Distribution System - Distributed Active Archive Center (LAADS-DAAC). They can be accessed via LAADS, as well as NASA's EarthData portal using different search-and-order tools. All XAERDT_L2 products are stored in NetCDF4 format. A downloadable tutorial for LAADS DAAC site and several other remote sensing resources is available on the NASA ARSET website. Another easy way to visually search and find files for an individual event is to use the NASA Worldview site. Currently, Worldview displays MODIS and VIIRS imagery and retrievals, and selected ABI-GOES imagery. Worldview does not yet display ABI and AHI aerosol retrievals.

4.1. XAERDT File Naming

As with all Atmosphere discipline products, LAADS groups datasets by "Archive Sets", and the Version 1 XAERDT_L2 files can be found under AS-5019. In general, filenames look like:

```
XAERDT_L2_SENSOR_PLATFORM.AYYYYDDD.HHMM.001.YearDayHrMnSc.nc
```

where:

- SENSOR_PLATFORM = ABI_G16, ABI_G17, AHI_H08, AHI_H09, MODIS_Aqua, MODIS_Terra, VIIRS_SNPP or VIIRS_NOAA20
- YYYY, DDD and HHMM are the four-digit year, three-digit Julian day, and time of day (hours and minutes in UTC) of the observation
- CCC is the collection/version (001),
- YearDayHrMnSc is year, Julian day, hour, minute and second when the file was processed.

- .nc = NetCDF format.

As noted above, MODIS granules correspond to 5-minute segments and VIIRS granules are 6-minute segments. Geolocation information is contained within the product files so there are no separate data and metadata files as are found in some satellite products. The Terra, Aqua, S-NPP and NOAA20 satellites have a 16-day orbital cycle. Granules from an individual sensor with the same time designation which are 16 days apart should be observing the same geographic location with the same geometry unless an orbital maneuver has been performed in the intervening time period. All AHI have a 10-minute cadence, except there are never files at 02:20 or 14:20 daily due to orbital maneuvers. ABI on both GOES-16 and GOES-17 began 2019 with 15-minute observation cadence, but switched to 10-minute cadence to match with AHI starting at 16:00 UTC on April 2, 2019 (DDD = 093). There may be missing files due to any number of reasons.

Product file dimensions are a function of the granule size, swath width, pixel size and viewing geometry. XAERDT_L2 products' approximate spatial dimensions of retrieval boxes (km at nadir) and retrieval box/ file pixel size (X × Y) are given in Table 1. For MODIS and VIIRS granules, "Y" means "Along swath" which refers to the path of the satellite and "X" means "across swath" which is perpendicular to this direction. While MODIS is 10 km x 10 km near nadir, its pixels expand in the across swath direction. The further from nadir the greater the expansion due to the sensor's viewing angle coupled with the curvature of the Earth. At the edge of the swath the product may be closer to 40 x 20 km. For VIIRS, the product is about 6 x 6 km at nadir, but due to updates in reduced in VIIRS sensor design and data handling choices, product dimensions at the edge only approximately double. For ABI and AHI, the FD includes the limb, as well as imagery "off" the edge of the Earth! The DT product keeps only when view angles are less than 72°. The relative increase in size (function of viewing angle) is presented as Fig 5, where there is some variability due to approximations in calculating ground-area based on assumed latitude and longitude "corners".

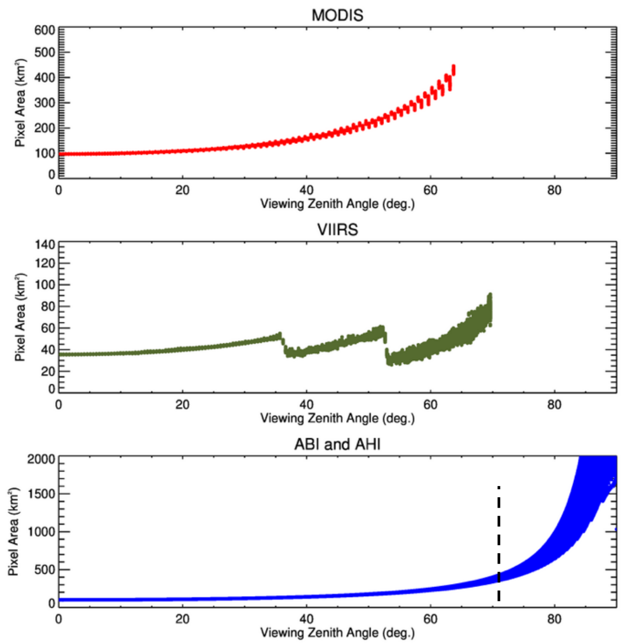


Figure 5: Pixel area (km2) as a function of sensor viewing geometry for all six sensors. The corresponding maps (Figure 1) show the locations of the individual swath. The figure used the same data as visualized in Figure 1.

4.2. XAERDT File Structure

XAERDT_L2 products are provided as NetCDF4 files, and there are numerous resources that include example codes (Python, Fortran, C, R, etc.) for reading and displaying NetCDF4 files. There are standalone command line tools known as NetCDF Operators (NCO). Finally, NASA maintains a free NetCDF display tool known as Panoply.

A main convenience of NetCDF is that they are considered to be "self-describing" via their metadata. To be archived and searchable at the LAADS datacenter, XAERDT_L2 files required the inclusion of metadata information such as text regarding the provenance of the data (which L1B, ancillary, and LUT files were used for the retrieval), bounding coordinates (corners if a LEO granule), information about where and what architecture the file was created, as well as information about the spacecraft and sensor.

Each retrieval parameter within a product file is termed a Scientific Data Set (SDS), and for XAERDT products, there are 44 SDSs in two groups (listed in Table 2). "Geolocation_data" include coordinates on the Earth's surface ("longitude" and "latitude") as well as sun/sensor observation angle geometry. The first four angles are interpolated from the native-resolution L1B file inputs, the Scattering and Glint angles are calculated during the DT retrieval. SDS group "Geophysical_data" includes the retrieved aerosol properties, further derived properties, diagnostics, and quality assurance information. Most SDSs are of two dimensions (function of X × Y), however some may be of three (third dimension may be as a function of wavelength band or something else). Since "Topographic_Altitude_Land" and "Land_Sea_Flag" are derived by interpolating ancillary information (not directly from the sensor's L1B), they are kept with Geophysical_data.

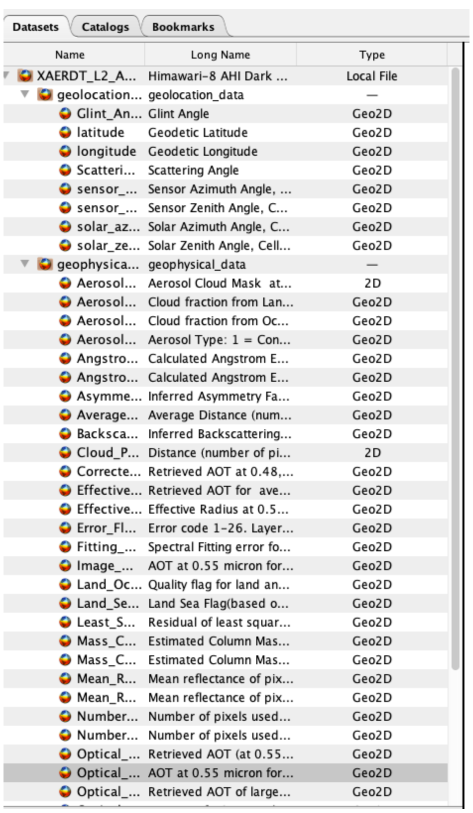
Table 2: List of SDSs within all XAERDT_L2 product files

Type	Name	Dims	Valid_Range or values after scaling (units)
	geolocation_data		
float	longitude	X x Y	-180° to 180° (degrees)
float	latitude	X x Y	-90° to 90° (degrees)
short	solar_zenith_angle	X x Y	0° to 180° (degrees)
short	solar_azimuth_angle	X x Y	-180° to 180° (degrees)
short	sensor_zenith_angle	X x Y	0° to 180° (degrees)
short	sensor_azimuth_angle	X x Y	-180° to 180° (degrees)
short	Scattering_Angle	X x Y	0° to 180° (degrees)
short	Glint_Angle	X x Y	0° to 180° (degrees)
short	Land_Sea_Flag	X x Y	0 = Ocean, 1 = Land and ephemeral water, 2 = Coastal
	geophysical_data		
short	Aerosol_Cldmask_Land_Ocean	XN x YN	0 = Cloud 1 = Clear
short	Aerosol_Cloud_Fraction_Land	X x Y	0.0 to 1.0
short	Aerosol_Cloud_Fraction_Ocean	X x Y	0.0 to 1.0
short	Aerosol_Type_Land	X x Y	1 = Continental, 2 = Moderate Absorption Fine, 3 = Strong Absorption Fine, 4 = Weak Absorption Fine, 5 = Dust Coarse
short	Angstrom_Exponent_1_Ocean	X x Y	-0.1 to 5.0
short	Angstrom_Exponent_2_Ocean	X x Y	-0.1 to 5.0
short	Asymmetry_Factor_Average_Ocean	X x Y x WA	0.0 to 3.0
short	Average_Cloud_Pixel_Distance_Land_Ocean	X x Y	0 to 60 (pixels)
short	Backscattering_Ratio_Average_Ocean	X x Y x WA	0.0 to 3.0
short	Cloud_Pixel_Distance_Land_Ocean	XN x YN	0 to 60 (pixels)
short	Corrected_Optical_Depth_Land	X x Y x WL	-0.1 to 5.0
short	Effective_Optical_Depth_Average_Ocean	X x Y x WA	-0.1 to 5.0
short	Effective_Radius_Ocean	X x Y	0.0 to 5.0 μm
short	Error_Flag_Land_And_Ocean	X x Y x Q	Error code 1-26. Layer Q=1 ocean, Layer Q=2 land. Ask for Documentation
short	Fitting_Error_Land	X x Y	0.0 to 1.0
short	Image_Optical_Depth_Land_And_Ocean	X x Y	-0.1 to 5.0
short	Land_Ocean_Quality_Flag	X x Y	0 = Bad 1 = Marginal 2 = Good 3 = Very Good
short	Least_Squares_Error_Ocean	X x Y	0.0 to 1.0
float	Mass_Concentration_Land	X x Y	0.0 to 1.0 (1.0e-6g/cm^2)
float	Mass_Concentration_Ocean	X x Y	0.0 to 1.0 (1.0e-6g/cm^2)
short	Mean_Reflectance_Land	X x Y x WA	0.0 to 1.0
short	Mean_Reflectance_Ocean	X x Y x WA	0.0 to 1.0
short	Number_Pixels_Used_Land	X x Y x WL	1 to 400
short	Number_Pixels_Used_Ocean	X x Y x WA	1 to 400
short	Optical_Depth_By_Models_Ocean	X x Y x S	-0.1 to 5.0
short	Optical_Depth_Land_And_Ocean	X x Y	-0.1 to 5.0
short	Optical_Depth_Large_Average_Ocean	X x Y x WA	-0.1 to 5.0
short	Optical_Depth_Ratio_Small_Land	X x Y	0.0 to 1.0
short	Optical_Depth_Ratio_Small_Ocean_0p55micron	X x Y	0.0 to 1.0
short	Optical_Depth_Small_Average_Ocean	X x Y x WA	-0.1 to 5.0
float	PSML003_Ocean	X x Y	0.0 to 1E10 (Particles/cm^2)
short	STD_Reflectance_Land	X x Y x WA	0.0 to 2.0
short	STD_Reflectance_Ocean	X x Y x WA	0.0 to 2.0
short	Surface_Reflectance_Land	X x Y x WL	0.0 to 1.0
short	Topographic_Altitude_Land	X x Y	0.0 to 10.0 (km)
short	Wind_Speed_GMAO_Ocean	X x Y	0.0 to 80.0 (m/s)

Type: short = 2-byte integer, float = 4 byte
Dimensions: X x Y = size of most products at retrieval resolution; XN x YN = size of some products which are at native pixel resolution; WA = wavelength indices for all channels; WL = wavelength indices of bands used for Land retrieval; S = Solution indices of Ocean retrieval, Q = Quality Flag indices. (note that actual values of X, Y, XN, YN may vary between sensors.

As noted in Table 2, XAERDT_L2 geophysical_data group includes SDSs referring to retrievals by DT-L (noted with _Land) and retrievals by DT-O over Ocean (noted with _Ocean). The Land SDSs may have a third dimension of WL, which refers to the 4 wavelengths (Blue, Green, Red, and SWIR3) used within the DT-L retrieval. Some Land and Ocean SDSs have a 3rd dimension of WA, referring to All wavelengths used in the retrieval. In the case of MODIS and VIIRS, this means all 7 (also includes NIR, SWIR1 and SWIR2). For ABI/AHI, while listed as dimensions, values of the SDSs for the missing WA wavelength may not be reported. A few SDSs have different 3rd dimensions, such as "S" which refers to the "Solution Index" of 9 aerosol modes fitted during the DT-O retrieval, and "Q" referring to "Quality_Flag_Num" of 2 sets (Land and Ocean) of Error codes determined during retrieval. A few variables (relating to Cloud Mask or Cloud Distance) are at the native resolution of the sensor (XN × YN).

For each SDS, there is additional metadata information that helps the user (or a tool) to properly scale and interpret the values of the data. For example, from XAERDT_L2_AHI_H08 file, we use Panoply tool to analyze the SDS named "Optical_Depth_Land_And_Ocean", and can see the kind of metadata which describes that SDS (Fig 6)



Variable "Optical_Depth_Land_And_Ocean"

In file "XAERDT_L2_AHI_H08.A2019207.0200.001.2022359145825.nc"

Variable full name: geophysical_data/Optical_Depth_Land_And_Ocean

```
short Optical_Depth_Land_And_Ocean(number_of_lines_10x10=1100, number_of_pixels_10x10=1100);
  :valid_range = -100S, 5000S; // short
  :_FillValue = -9999S; // short
  :long_name = "AOT at 0.55 micron for both ocean (Average) (Quality flag = 1, 2, 3) and land (corrected) (Quality flag = 3)";
  :units = "None";
  :scale_factor = 0.001; // double
  :add_offset = 0.0; // double
  :Parameter_Type = "Output";
  :Geolocation_Pointer = "Internal geolocation arrays";
  :coordinates = "/geolocation_data/longitude /geolocation_data/latitude";
  :_ChunkSizes = 1U, 1100U; // uint
```

Figure 6: Screenshot of Panoply window (top), highlighting the metadata within the SDS named "Optical_Depth_Land_And_Ocean" (bottom)

We see the following information about the SDS named Optical_Depth_Land_And_Ocean :

short – variable type specification

number_of_lines_NxN and **number_of_pixels_NxN** - the dimension of this parameter as returned by the algorithm in the Y × X directions. For AHI case, N = 10, and the values are 1100 and 1100.

valid_range – the minimum and maximum raw values returned by the algorithm

_FillValue – the value returned when there is no algorithm retrieval

long_name – a brief text descriptor of the variable. When multiple wavelengths are mentioned in the long name there will be a third dimension for this SDS parameter.

units – unit descriptor or name applied to this variable

scale_factor – the scaling value applied to the raw numbers returned by the algorithm **add_offset** – the offset value added to the raw numbers returned by the algorithm

Parameter_Type – Input or Output

coordinates– This tells a tool such as Panoply to consider this SDS as a "Geo2D" array, so that it can be plotted as a function of latitude and longitude.

A tool such as Panoply understands to use the "scale_factor" and "add_offset" to compute the actual value of the SDS. For example, if the AOD described by this SDS is reported as a short integer with value of x = 525, then AOD = x * scale_factor + add_offset = 0.525. This effectively converts a two-byte integer (saving space in the file), into a 4-byte float variable, which then can be used in additional calculations.

An example of a 3D SDS is "Effective_Optical_Depth_Average_Ocean", presented in Fig XX. Here the long_name would be updated to reflect the wavelengths of the given sensor.

Variable "Effective_Optical_Depth_Average_Ocean"

In file "XAERDT_L2_AHI_H08.A2019207.0200.001.2022359145825.nc"

Variable full name: geophysical_data/Effective_Optical_Depth_Average_Ocean

```
short Effective_Optical_Depth_Average_Ocean(number_of_lines_10x10=1100, number_of_pixels_10x10=1100, Wavelength_Used_ALL=7);
  :valid_range = -100S, 5000S; // short
  :_FillValue = -9999S; // short
  :long_name = "Retrieved AOT for  average  solution at 0.48, 0.55, 0.67, 0.86, 1.24, 1.60, 2.25 microns";
  :units = "None";
  :scale_factor = 0.001; // double
  :add_offset = 0.0; // double
  :Parameter_Type = "Output";
  :Geolocation_Pointer = "Internal geolocation arrays";
  :coordinates = "/geolocation_data/longitude /geolocation_data/latitude";
  :_ChunkSizes = 1U, 1100U, 7U; // uint
```

Figure 7: Metadata for the 3-dimensional SDS named Effective_Optical_Depth_Average_Ocean"

Note that while the structure of all SDS metadata remains the same, values may depend on the sensor. Other tools (such as the command-line NCO tool) may present this information in a different way. Finally, note that if writing your own code to operate on SDSs, one must be careful to spell names correctly, including underscores.

4.3. Aerosol SDSs and use cases

As its overall goal, the DT algorithm compares satellite-measured spectral reflectance (in up to 7 wavelength bands) with Lookup Tables (LUTs). The assumptions for making the LUTs are identical between sensors, except for accounting for wavelength shifts. The spectral fitting is a minimizing procedure that results in aerosol conditions that best match the spectral observations. For each retrieval box (NxN aggregations of the native resolution pixels), the result includes **aerosol optical depth (AOD),** and the **fine model fraction (FMF).** AOD and FMF are retrieved using different assumptions whether over Land or over Ocean. From these retrieved parameters, one can further derive properties of the aerosol, such **Angstrom Exponent (AE)**.

From these retrieved and derived parameters, DT estimates additional parameters, which we will not discuss in this User Guide.

Each retrieval come with an indicator of the quality of the retrieval referred to as the Quality Assurance/Confidence (QA) flag, with a value at each (retrieval) pixel location. QA flag values are found in an SDS that is separate from the product value SDSs. QA flag values range from the lowest confidence of 0 to the highest confidence of 3. For Land-based products we suggest using only QA≥2. For Ocean based products we suggest using QA≥1 .

Aerosol Optical Depth (AOD)

The aerosol optical depth (AOD or τ; also called aerosol optical thickness or AOT) refers to the optical loading of the aerosols in the atmospheric column. AOD is a unitless value. It is most closely related to the total surface area of the aerosol that is interacting with a light source. AOD is spectrally dependent, meaning it varies by wavelength. In general many remote sensing studies looking at AOD are reporting at a mid-visible wavelength. Therefore, for all XAERDT_L2, we are reporting AOD values at 0.55 microns.

The most commonly used SDS's pertaining to AOD are listed here and explained below:

- Corrected_Optical_Depth_Land
- Effective_Optical_Depth_Average_Ocean
- Land_Ocean_Quality_Flag
- Image_Optical_Depth_Land_And_Ocean
- Optical_Depth_Land_And_Ocean

All XAERDT products allow for small negative AOD retrieval values in order to avoid an arbitrary negative bias at the low AOD and in long-term statistics. This is because there is insufficient sensitivity over land (due to poorly constrained surface reflectance) to retrieve aerosol to better than within +/-0.05. The consequence is that in very clean conditions the algorithm cannot truly distinguish between AOD values in the range of -0.05 to 0.05. If we eliminate all the negative numbers and keep only the positive numbers, we will introduce an artificial bias to the long-term statistics, therefore we allow negative retrievals down to -0.05. For end users: If you are calculating long-term statistics simply include the negatives in your analysis. If you are looking at individual retrievals then count negative retrievals as 'very clean'. You could force them to be AOD = 0, for example. It really depends on the application. However, these small negative AOD values are valid retrievals and do contain useful information.

Corrected_Optical_Depth_Land – AOD retrieved over Land at 4 wavelengths (Blue, Green=0.55 µm, Red, SWIR3).

> The term "Corrected" in the SDS name is a legacy from earlier products. There is no "Uncorrected" version of this SDS. Except for the green wavelength always reported at 0.55 µm, other AODs are reported at the sensor's wavelength.

Effective_Optical_Depth_Average_Ocean – AOD retrieved over Ocean at 7 wavelengths.

> This SDS averages all retrievals within the algorithm's acceptable error limits. AOD at 0.55 is reported for all sensors, however, the other wavelengths represent specific sensor.

Land_Ocean_Quality_Flag – Quality assurance (QA) or Confidence for the Ocean and Land

> For each retrieval path Land or Ocean, QA values are subjectively assigned by the algorithm team based on numerical standards such as number of input pixels used for the retrieval, proximity to bright land or ocean glint and error fitting values. QA values range from 0 – 3 where zero is lowest confidence and 3 is highest confidence.

Image_Optical_Depth_Land_And_Ocean - AOD land plus ocean at 0.55 µm, not filtered for QA.

This SDS joins the AOD retrievals for land and ocean, with no requirements for QA confidence levels. It is reported only at 0.55 µm, and provides a snapshot with the greatest amount of retrieval coverage, suitable for general overviews of the scene.

Optical_Depth_Land_And_Ocean – AOD land plus ocean reported at 0.55 µm, *filtered* for QA

This SDS uses QA to "filter" before joining the AOD retrievals for land and ocean. Specifically, it requires QA≥2 for land, and QA≥1 for ocean. We expect this SDS should be the default for most quantitative studies. It is reported only at 0.55 µm.

Parameters Related to Particle Size and Model choices

Fine and Coarse Mode

In general, atmospheric aerosols are observed to have multi-modal distributions. Smaller particles are referred to as "fine mode" (also known as accumulation mode), and have effective radii between 0.1 and 0.25 micrometers. Larger particles ("coarse mode"), have effective radii between 1.0 and 2.5 microns. The DT algorithm assumes properties of each size range (fine model and coarse model) and attempts to mix them together when matching the satellite-observed spectral reflectance. Details on the specific assumptions and how the algorithm computes (whether Land or Ocean) are given in the ATBD.

The aerosol fine mode fraction (FMF; also known as Fine Mode Weighting or FMW) is the proportion of fine mode aerosols to the total, and is defined in terms of total AOD. Therefore, it then be multiplied by the total AOD to determine proportional fine and coarse -AOD (coarse = total – fine). The proportional AOD is only reported for the ocean product. We feel that the algorithms' ability to distinguish fine or coarse mode over land is not accurate enough to permit us to make a fine mode AOD calculation, even when the retrieval algorithm is providing FMF.

Optical_Depth_Ratio_Small_Land: FMF over land

Optical_Depth_Ratio_Small_Ocean_0p55micron: FMF over ocean

Optical_Depth_Small_Average_Ocean: AOD * FMF over ocean, reported at 7 bands.

Optical_Depth_Large_Average_Ocean: AOD * (1-FMF) over Ocean, reported at 7 bands.

Optical_Depth_By_Models_Ocean: The Ocean retrieval retrieves its best fit by *choosing* among combinations of fine (4 choices) and coarse (5 choices) aerosol models. Essentially, Optical_Depth_Small gets assigned to its model choice (index 1 to 4), whereas Optical_Depth_Large gets assigned to its model choice (index 5 – 9).

Aerosol_Type_Land: The Land retrieval retrieves its best fit by *assuming* aerosol models based on season and location. The values are denoted by *1 = Continental, 2 = Moderate Absorption Fine, 3 = Strong Absorption Fine,4 = Weak Absorption Fine, 5 = Dust Coarse*

Ångström Exponent

The Ångström Exponent (AE) relates to spectral dependence of AOD, and is often used as a qualitative indicator of mean particle size. As a rough guideline Ångström Exponent values in the range of 2 indicate small particles which might be associated with pollution or biomass burning. Values in the range of 1 or less indicate the presence of large particles such as sea salt or dust. For the MODIS algorithm Ångström Exponent is not a true measurement but is a derived value. Note, due to signal-to-noise issues, AE is only reported when aerosol loading (total AOD) is moderate or larger (AOD>0.2). Note that AEs are calculated based on the nominal wavelengths of the sensor, and may not be entirely comparable between sensors.

Angstrom_Exponent_1_Ocean: AE over Ocean, computed by using Effective_Optical_Depth and comparing AOD for 0.55 and NIR (e.g. 0.87 µm) bands; relates to spectral dependence of fine mode

Angstrom_Exponent_2_Ocean: AE over Ocean, computed by using Effective_Optical_Depth and comparing AOD for NIR (e.g. 0.87) and SWIR3 (e.g. 2.25) bands; relates to spectral dependence of coarse mode

Angstrom_Exponent_Land: AE over Land, computed by using Corrected_Optical_Depth, and comparing AOD for Blue (e.g. 0.49) and Red (e.g. 0.67).

Other Parameters

There are a number of other parameters reported within the DT aerosol product. Some of these are "derived" or "diagnostic" products; reported because they have been useful for some applications. None of these are validated. More details can be found in the full ATBD.

For example, there is an estimate of **Mass_Concentration** over both **_Land** and **_Ocean**, which represents the total integrated columnar mass per square area. It is not a retrieved parameter, but rather is a function of aerosol models assumed during the retrieval (and assumed density). Other similar SDSs include **Assymetry_Parameter** and **BackScattering_Ratio**.

Other diagnostic parameters include SDSs relating to **Mean_Reflectance**. These are the averaged values of reflectance used for either the **_Land** or **_Ocean** retrievals.

There are SDSs related to the algorithm's estimate of clouds in the retrieval box. This includes the **Aerosol_Cldmask_Land_Ocean**, which is an NxN grid within each retrieval box, of pixels marked "cloud" or "clear", along with an estimate of the **Aerosol_Cloud_Fraction** within the retrieval box.

Finally, an interesting diagnostic is the **Cloud_Pixel_Distance_Land_Ocean**, which attempts to indicate the distance (in pixels) between each pixel and the nearest detected "cloud" pixel.

Appendix: XAERDT difference from "standard" MODIS and VIIRS products

As noted in the introduction, MODIS and VIIRS have been flying much longer than since 2019. NASA satellite product developers are constantly working to evaluate, update and improve their products. From time to time after significant changes are made to the algorithm and/or instrument calibration, the entire data set for a single or related group of instrument products are reprocessed. Development of MODIS and VIIRS standard aerosol products, including those derived using the DT algorithm, is performed at different places, and updates to algorithm and products are guided by their own funding and timeline vis-à-vis needs of the MODAPS and/or SIPS schedules. For this version of the XAERDT product, we decided to "freeze" the DT algorithm, and port as a "lowest common denominator", thus leading to as consistent of a MEaSUREs dataset as possible. Therefore, there are some key differences with standard DT products for MODIS and VIIRS.

For VIIRS, the set of similarly-derived products is known as a Version, with the current Version (as of Nov 2023) known as V2.0. These products are known as AERDT_L2_VIIRS_Sensor, where Sensor = SNPP or NOAA20. In fact, there are only minor differences between XAERDT_L2_VIIRS V1.0 and AERDT_L2_VIIRS V2.0. The only expected differences are in the top-level metadata, and possibly some machine calculation precision, as the two sets of products were produced for different purposes and on different systems at Wisconsin's SSEC.

For MODIS, the set of common products is known as a "Collection", with the current Collection C6.1 (or C61) being in production since 2017. The Level 2 products for MODIS are known as MxD04_L2, where the x=O for MODIS on Terra, and x=Y for MODIS on Aqua. There are a few key differences with XAERDT_L2_MODIS_Sensor versus MxD04_L2 including:

- XAERDT is based on only the DT algorithm, whereas MxD04 includes retrievals from both DT and Deep Blue retrieval algorithms
- XAERDT is in NetCDF4 format versus MxD04 being in HDF4. Geophysical_data and geolocation_data are separated into groups.
- XAERDT drops some heritage SDSs, deemed either too confusing or unnecessary.
- XAERDT computes cloud mask based on the entire granule, whereas MxD04 uses a heritage "scan by scan" method
- XAERDT uses highest resolution (Red) band for visible-band cloud masking over ocean, compared to MxD04 using green band.
- XAERDT uses SWIR2 band rather than SWIR1 wavelength band for sediment mask, to be consistent with missing SWIR1 on GEO sensors.
- Differences in code structure and "modularity"

The net result is that XAERDT_L2_MODIS V1 and MxD04_L2 C61 are qualitatively similar, however there are some decision making (to retrieve or not to retrieve) in cloudy situations. For a future MODIS "Collection 7", we intend to adopt the DT-Package structure (and NetCDF4 outputs) from XAERDT.