Mineral Aerosol Profiling from Infrared Radiances (MAPIR) version 5.1

File content and user information

We request that any potential user read this document to avoid misuse or misinterpretation of the data and especially its information content. We highly recommend to also reading the reference publication. Profile comparison with reference data and profile assimilation should be done according to the state-of-the-art literature on the subject and usually involves using the averaging kernels.

Note: version 5.1 (described here) is currently under evaluation

The user interested only in level 3 (average on a grid, daily and/or monthly, morning, evening or all day) dust AOD and/or dust mean altitude can find the data in the Climate Data Store (cds.climate.copernicus.eu) - search term "aerosol" then in the data set containing all aerosol data, select the MAPIR algorithm to access our data. The user interested in level 2 (at satellite resolution) dust AOD and/or dust mean altitude and/or the full vertical profiles of dust concentration should use the data set available here.

Contact information

Sophie Vandenbussche

EMail: sophie.vandenbussche@aeronomie.be

Phone: +32 2 373 03 86

Reference publication

Currently, the new version 5.1 publication is under preparation, here is the link to the previous version of the algorithm. Until the publication of version 5.1 is ready, users may ask additional information through the contact details above.

Callewaert et al., AMT 2019 : doi.org/10.5194/amt-12-3673-2019

Short algorithm description

The MAPIR algorithm retrieves **dust aerosols vertical profiles** (**10 layers** centred at 0.5km to 9.5km by steps of 1 km) following the well-known **Optimal Estimation** (OE) method with **Levenberg Marquardt** implementation (on the **logarithm** of the dust concentration), using three spectral windows in the so-called atmospheric window (800-1250cm-1). The surface temperature is also adjusted in the retrieval process as it is a very important parameter in the thermal infrared, but it is not considered a product and therefore not included in the files. Currently MAPIR is applied to data from **IASI/Metop-A** and **IASI/Metop-C** (sun-synchronous polar orbit with local solar time of about **9h30 and 21h30**). When available (currently: for IASI/Metop-A – the data available for IASI/Metop-C remains consistent with that reprocessing), the reprocessed IASI data was used, to obtain a time-consistent data set.

No prefilters for clouds or dust presence are done prior to the retrieval. Only the quality of the IASI spectra and corresponding temperature and humidity profiles is checked.

The **total column** of dust aerosols is obtained by adding the partial columns from each layer, themselves obtained simply by multiplying the retrieved concentration in the layer by the layer thickness. The dust **AOD** is obtained by multiplying the aerosol total column by the extinction cross-section at the desired wavelength. See Table 1 for the exact values at 10µm and 550nm.

The **dust mean altitude** is obtained from the vertical profile as the altitude for which half the aerosols are below and half are above. It is the altitude cutting the aerosol total column in two equal partial columns. The dust mean altitude is only provided if the 10µm AOD is larger than 0.05 and the information content (degrees of freedom in the dust vertical profile, in the OEM) is larger than 1.25. Lower information content means there is no sufficient information in the vertical range to provide more than just an integrated vertical column. Lower AOD would lead to very high uncertainty in the computed mean altitude. When the mean altitude is not calculated, the fill value is used.

Table 1: Dust extinction cross-sections for AOD computation

Wavelength	550nm	10µm
Dust extinction cross-section	6.93367e-08 cm2	3.895323e-08 cm2

An **uncertainty** on the retrieved parameters is calculated during the retrieval, following the OEM formalism. That uncertainty comprises the propagation of the spectral noise, the so-called smoothing error, and the propagation of uncertainty linked to temperature and humidity profiles. The assumed spectral noise is between 0.2 and 0.5K (depending on the spectral window). The assumed uncertainty in temperature and humidity profiles were respectively 1K and 10%, which are the target accuracy for those products.

Information content and averaging kernels

The sensitivity of a vertical profile retrieval is characterized by the **Averaging Kernel** (AK), which is calculated during the retrieval. The AK is a **square matrix of the size of the state vector** (all retrieved parameters). In our case, we remove the parts linked to the surface temperature retrieval, leaving the AK to be a matrix of 10*10 linked to the vertical profile of dust in 10 layers.

The **diagonal** elements represent the specific **sensitivity** of the retrieval to **each layer**, while the **other** elements represent the **cross-sensitivity** between the different layers (i.e. AK_{ij} is the sensitivity of the retrieved dust in layer i to "real" dust in layer j).

The **trace** (sum of diagonal elements) of the averaging kernel is called the "**Degrees Of Freedom**" (DOF) of the retrieval. For MAPIR, the DOF is **usually about 2** in good conditions (temperature not too low, and reasonable amount of dust aerosols) and may be even higher in the best cases.

Given that the DOF is about 2 while the number of retrieval layers is 10, it is obvious that **the precise dust concentrations retrieved in each layer are partly depending both on the a priori** (when the diagonal AK for that layer is significantly lower than 1) **and on the dust in other layers** (when the cross-diagonal elements are significantly higher than 0). Each layer's retrieved concentration is not independent, as only 2 independent pieces of information are available.

Note: The common assumption that those 2 pieces of information are the total AOD and the mean altitude is wrong here, as those are correlated in the case of Thermal Infrared aerosol retrievals. Therefore 2 DOF means that 2 independent layers can be retrieved.

Additional information on the averaging kernels, some examples and the implications of the OEM-LM implementation in the logarithmic mode retrieval are in the reference publication.

Additional information on how to **use dust concentration in a single layer** (for example, the layer closest to surface) are in another publication, in its section 3.1: doi.org/10.5194/acp-2020-130 [1]. One should **ensure a sufficient information content in the target layer**, using the corresponding **diagonal element from the AK**. If combining more than one layer in a partial column, again one should care that a sufficient amount of information is contained in those layers using the sum of the AK diagonal elements for those layers.

[1] North African mineral dust sources: new insights from a combined analysis based on 3D dust aerosols distributions, surface winds and ancillary soil parameters, Sophie Vandenbussche, Sieglinde Callewaert, Kerstin Schepanski, and Martine De Mazière, ACPD 2020

Quality flags and data filtering

Two flags are available in the data files.

A **quality flag**, reporting on the quality of the retrieval, which is assessed based on the following criteria :

- Root Mean Square of the Spectral Residuals (RMSSR) < 0.75K
- Retrieved 10µm AOD < 5
- Retrieved Ts > 200 K
- Retrieved Ts < 350 K
- Retrieved Ts - a priori Ts > -2.5K

The spectral residuals are the differences between the modelled spectrum with the final state vector and the observed spectrum, in the retrieval windows only. High RMSSR are mostly due to a bad retrieval situation, but it could also be linked to a problem in the surface emissivity or in the vertical profile of temperature, or to the presence of a cloud.

The second and third tests are meant to remove (thick) clouds, for which a very high AOD and/or a very low surface temperature are retrieved.

Retrievals with an unphysically high surface temperature are also flagged as bad (fourth test). They refer to most probably a problem with either the spectrum or the ancillary data.

The last test is meant to eliminate relatively thin clouds. It may be used only if the a priori surface temperature comes from the PWLR IASI level 2 data (which is the case here). It is based on the fact that the PWLR retrieval, being a statistical retrieval, outputs a rather good surface temperature even under a cloud or aerosols. In MAPIR, for cloudy scenes the retrieval will output a rather high AOD and a surface temperature biased low with respect to reality. Under dust aerosols situations, MAPIR will output the dust AOD and an unbiased surface temperature. The precise threshold of this test was decided empirically, analysing data from selected events and comparing with MODIS imagery data to “humanly” discriminate clouds from dust.

A **cloud flag** is also included, produced using the EUMETSAT IASI cloud fraction (maximum 10% for clear sky) and, when available, cloud flag (maximum value of 1 for clear sky). **This cloud flag is however not used to filter data** before the retrieval, because it has been shown to also flag some dust events. It is included only for information.

Known shortcomings of the algorithm / data set

- **Imperfect cloud filtering**: it is difficult to discriminate thin clouds from thin dust plumes, and some misclassification may occur, leading to the removal of some dust events, or to wrong dust AOD due to an actual cloud that was not filtered
- **Uncertainties regarding the dust optical properties**: a generic set of dust optical properties is used, while in reality those may vary with time and space. The consequence is an increased uncertainty mostly on the vertical distribution of the dust aerosols.
- **The validation is undergoing**; a quick first evaluation shows that MAPIR seems to overestimate the lowest dust AODs, and underestimate the high dust AODs.

Data files content

The data is stored in **netcdf4** format. The data provided is the level 2 data, at satellite resolution. The files contain **all IASI scenes**. They do contain also results of dubious retrievals and of cloudy scenes, which should be screened out using the **quality_flag**.

Field name	**Short description**
latitude	Latitude in degrees North, range -90 to 90
longitude	Longitude in degrees East, range -180 to 180
time	UTC Time in seconds since 01/01/1970, 0h
altitude	Mid-layer altitudes (km) for the dust vertical profiles (fixed)
D_vertical_profile	Retrieved dust vertical profile (concentration in particles / cm^3) on the vertical range given in "altitude")
D_vertical_profile_uncertainty	Uncertainty on the dust vertical profile (concentration in particles / cm^3)
D_a_priori	A priori profile (particles / cm^3)
D_averaging_kernel	Full averaging kernels. See previous section.
D_AOD10000	Dust AOD at 10µm (column integration of the vertical profile, multiplied by the cross-section)
D_AOD10000_uncertainty	Partial uncertainty on the dust AOD at 10µm
D_AOD550	Dust AOD at 550nm (D_AOD550= D_AOD10000 x 1.78)
D_ALT	Mean dust altitude (altitude at which half the dust column is below and half is above)
cloud_flag	Flag based on the EUMETSAT IASI cloud products (see previous section) – this flag is NOT USED, is included for information only, for expert users 0: clear; 1: cloudy
surface_type_number	0: sea; 1: land
satellite_zenith_at_center	Satellite viewing angle at the centre of the field of view (given with respect to the vertical at the ground)
solar_zenith_at_center	Solar zenith angle at the centre of the field of view (given with respect to the vertical at the ground)
quality_flag	Flag referring to the post retrieval quality screening, ensuring good quality of the data. See previous section. 0: bad; 1: good
orbit_flag	Part of the satellite orbit 0:ascending; 1:descending
cell	The dimension variable for the netcdf file (no other use)