L2A+

Enhanced Aeolus L2A for depolarizing targets and impact on aerosol research and NWP

-

Description of the Algorithm Developments (ALGO)

Deliverable Item 03 [DI03]
(Version 2)

Submitted to: Christian Retscher (ESA)

	Name	Function	Date
Prepared by:	E. Proestakis	WP1000 – NOA	08/2023
	A. Gkikas	WP3000 – NOA	08/2023
	K. Rizos	WP3000 – NOA	08/2023
	T. Georgiou	WP4000 – NOA	08/2023
	A. Tsikerdekis	WP4000 – KNMI	01/2024
	A. Kampouri	WP4000/5000 – NOA	08/2023
	E. Drakaki	WP4000/5000 – NOA	08/2023
Approved by:	V. Amiridis	PI	08/2023

-

National Observatory of Athens (NOA)
Institute for Astronomy, Astrophysics, Space Applications & Remote Sensing (IAASARS)
Vas. Pavlou & I. Metaxa, 15236 Penteli, Greece
&
Leibniz Institute for Tropospheric Research (TROPOS), Leipzig, Germany
&
European Centre for Medium-Range Weather Forecasts
[ECMWF]
Reading, United Kingdom

[This page is intentionally left blank.]

Table of Contents

1. ESA-L2A+ DI03 – Overview.

This document consists the Deliverable Item 03 (DI03) – ALGO – Version 2 (V2) submitted to the European Space Agency (ESA) by the consortium of the project "Enhanced Aeolus L2A for depolarizing targets and impact on aerosol research and NWP" (L2A+). DI03– ALGO reports on activities and developments related to L2A+ Work Package 3 (WP3000) – "Development of the L2A+ aerosol product" and Work Package 4 (WP4000) "Assimilation of L2A/L2A+ and application of WRF-L experiments". More specifically, DI03 – ALGO aims to document on the algorithms and methods used for the L2A+ retrieval and the data assimilation methodology and routines.

2. Introduction.

The European Space Agency's (ESA) wind mission, Aeolus, hosts the first space-based Doppler Wind Lidar (DWL) world-wide. Its scientific objectives are to improve Numerical Weather Predictions (NWP) and to advance the understanding of atmospheric dynamics and its interaction with the atmospheric energy and water cycle. Aeolus primary data product consists of profiles of horizontally projected line-of-sight winds from the surface up to about 30 km, and spin-off products are profiles of cloud and aerosol optical properties (e.g., Straume-Lindner et al., 2021). The Aeolus optical properties spin-off product (L2A; Flamentet et al., 2021) has been used to study smoke emissions, e.g., from 2019 Californian wildfires (Baars et al., 2021) and the early 2020 Australian fires, enhancement of volcanic ash forecasting systems and the impact of the Tonga eruption on tropospheric and stratospheric loads (e.g., Kampouri et al., 2021), and has been successfully experimentally assimilated in Copernicus Atmosphere Monitoring Service (CAMS) model C-IFS.

While the L2A product has a reasonable quality (see Baars et al., 2021, GRL, Paschou et al. 2021), its full potential for aerosol and cloud studies and for further improving NWP, has not been exploited. This is mainly because L2A is not provided separately for aerosol and cloud targets. Furthermore, the product is underestimated in terms of backscatter because the instrument measures only the co-polar part of the atmospheric backscattered return of the circularly polarised emitted beam (Paschou et al., 2021; Ehlers et al., 2022). Hence, in the case of strongly depolarizing targets (mainly desert dust particles and ice crystals), the signal measured at the detectors is strongly reduced with respect to non-depolarizing targets. Additionally, the expected atmospheric return signal in orbit is a factor of 2.5 to 3 lower than expected before launch due to lower laser output energies than originally intended (45–72 mJ) and decreased instrument transmission by about 30%, which has caused a lower SNR since mission start (Reitebuch et al., 2020).

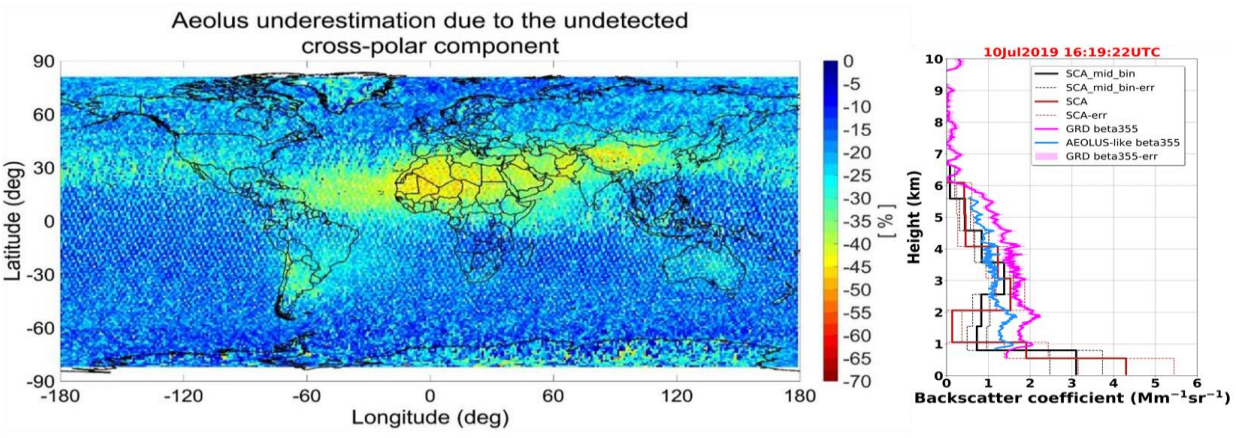

Figure 2.1: Aeolus underestimation due to the missing cross-polar channel (a:) theoretical calculation; (b) observational evidence.

There is a certain scope for delivering an enhanced L2A product based on a more detailed and robust classification, incorporating other data sources such as multi-sensor synergies or CAMS data. Furthermore, the potential of the extinction product should be further exploited as a constraint for the backscatter product for depolarizing targets, since the extinction is not affected by the missing cross-channel.

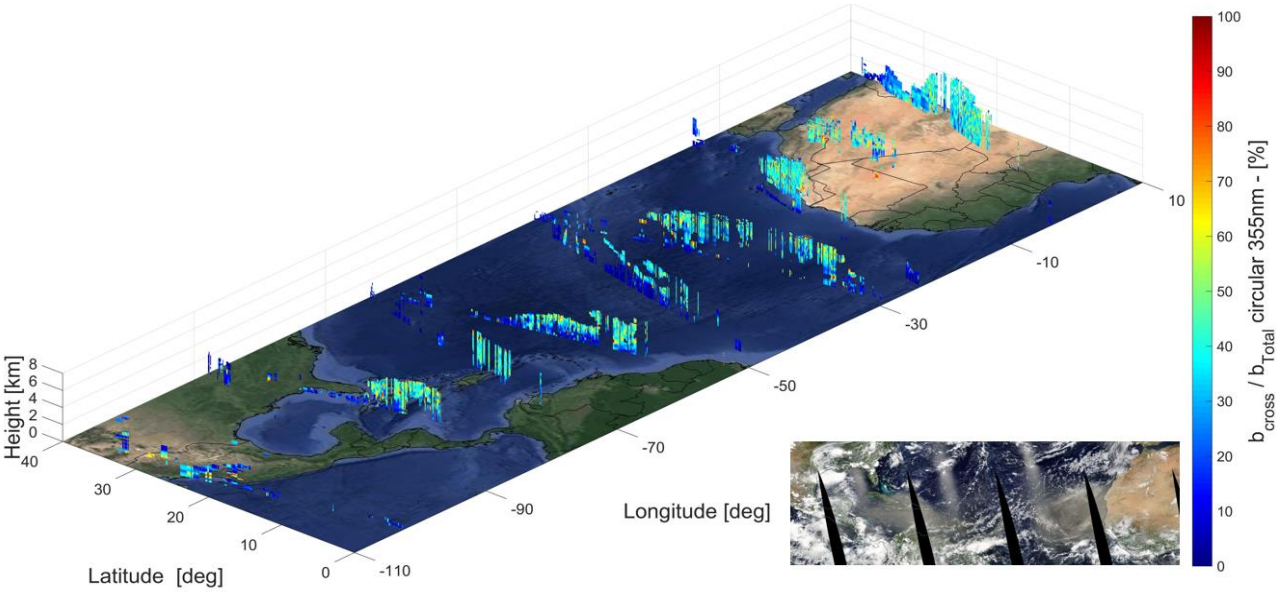

Figure 2.2: Estimates of Aeolus L2A underestimation due to the missing cross-channel using the Aeolus-like profiles retrieved based on CALIPSO for the trans-Atlantic Godzilla dust event on the 23rd of June, 2020.

ESA-L2A+ overarching objectives include the development of a refined Aeolus L2A aerosol product (L2A+), and accordingly test its applications for enhancing aerosol research, contributing to existing climate datasets. Furthermore, L2A+ aims at assessing the impact of the new product on aerosol assimilation towards improved dust transport modelling and for further enhancing NWP.
The following L2A+ individual objectives related to DI03 – ALGO are related to:

1) Objective 1: To develop a refined Aeolus aerosol optical product (L2A+), based on AEL-FM/AEL-PRO algorithms, geostationary AOD products, CAMS, and new AOD retrievals from the Aeolus itself (WP3000). The product will be thoroughly compared with L2A and validated against quality-assured measurements from the ESA-ASKOS/JATAC experiment in Cabo Verde (WP2000).
2) Objective 2: To examine the impact of L2A and L2A+ on aerosol assimilation and dust transport models (WP4000). Constraining aerosol loads through Aeolus assimilation is expected to improve dust emission fluxes over the Sahara Desert and transport simulations of the full particle size range (from fine to giant dust particles).

On the basis of L2A+, DI03 – ALGO aims to provide a consolidated view of the algorithms and methods used for the L2A+ retrieval and the data assimilation methodology and routines associated with the ESA-L2A+ overarching Objectives (1)-(5).

3. ESA-L2A+ WP3000: Development and validation of the L2A+ aerosol product.

3.1. Description of WP3000.

While the primary Aeolus aerosol product (L2A) has reasonable quality, its applicability in aerosol and cloud studies has not been exploited yet. The main reason is attributed to its deficiency to provide retrievals separately for aerosol and cloud targets. Furthermore, an underestimation of the backscatter coefficient is expected when the aerosol/cloud targets are depolarizing features (non-spherical particles); ALADIN provided only the co-polar part of the atmospheric backscattered return of the circularly polarized emitted light pulses. Towards overcoming the aforementioned inherent deficiencies, WP3000 aims to deliver an enhanced L2A product (with focus on dust) based on a more detailed and robust classification, involving multiple data sources such as multi-sensor synergies in conjunction with reanalysis numerical outputs and reference ground-based measurements. The new L2A product will be utilized on aerosol data assimilation schemes coupled with dust transport models with aim to improve Numerical Weather Prediction (NWP). The study period spans over the entire September 2021 when the JATAC/ASKOS campaigns took place in Cabo Verde. The region of interest refers to an extended tropical domain including the Sahara Desert and the Tropical Atlantic Ocean up to the Caribbean Sea. For the derivation of the enhanced L2A product, a series of processing steps has been designed which are presented in Figure 3.1.

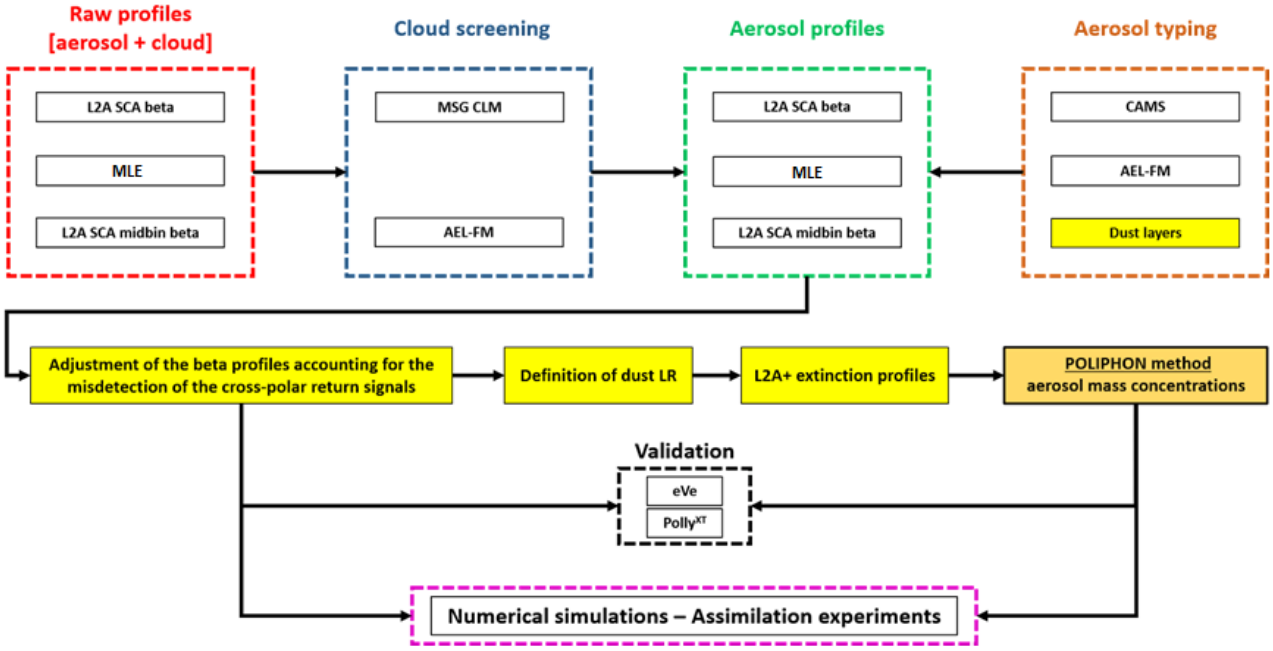

Figure 3.1: A graphical presentation of the WP3000 processing steps.

L2A+ starting point consists of the raw Aeolus L2A retrievals derived using several algorithms, including the Standard Correct Algorithm (SCA), the Standard Correct Algorithm mid-bin (SCA mid-bin) provided at the middle-bin vertical scale (23 instead of 24 vertical bins) and the Maximum Likelihood Estimation (MLE) algorithm (Ehlers et al., 2022). These raw Aeolus L2A optical properties were processed using a synergy of filtering tools. In the first working step a rigorous filtering of the raw Aeolus L2A retrievals will be performed for the elimination of the cloud-contaminated profiles. This cloud-filtering methodology will be performed relying on the AEL-FM (Aeolus Feature Mask) product from the L2A processors II (originated from the developments for

the forthcoming EarthCARE mission but adapted to Aeolus) and the MSG-SEVIRI (Meteosat Second Generation - Spinning Enhanced Visible and Infrared Imager) CLAAS-3 cloud-mask (CM)dataset. Based on AEL-FM retrievals, the features of the probed atmospheric scene are classified either as clouds or as aerosols or as molecular (Rayleigh) atmosphere. Moreover, the MSG-SEVIRI CLAAS-3 cloud dataset is also utilized as an extra or optional cloud-filtering tool for selecting only the pure aerosol Aeolus L2A profiles. The next working step focuses on the identification of the dust layers at the cloud-filtered Aeolus vertically resolved profiles derived from the previous step. At this point, CAMS reanalysis data are used for the assignment of aerosol typing under the absence of an aerosol classification scheme on the raw Aeolus L2A data. In the third working task, the L2A+ aerosol extinction profile will be retrieved by using the Aeolus L2A+ cloud-free backscatter values and an appropriate lidar ratio value for the dust aerosol layers. Before the final product being injected in the WP4000, an assessment of the new L2A+ extinction profiles will be performed in the final working step against the corresponding ground-based retrievals acquired by the eVe and PollyXT lidars, operated in the framework of the JATAC campaign.

In the next sections we will present all the processing steps that have been undertaken up until now as well as some preliminary results.

3.2. Cloud-filtering of the raw Aeolus L2A retrievals.

In the first phase, a rigorous filtering of the raw Aeolus L2A backscatter coefficients retrievals has been performed for the elimination of the cloud-contaminated bins. In order to achieve the optimum cloud screening, the raw Aeolus L2a retrievals provided at BRC level, are jointly processed with the AEL-FM (Aeolus Feature Mask) retrievals and the MSG-SEVIRI (Meteosat Second Generation - Spinning Enhanced Visible and Infrared Imager) CLAAS-3 cloud dataset.

3.2.1. Cloud-filtering methodology based on AEL-FM feature mask product.

At first, the implementation of the cloud screening methodology was based on the AEL-FM retrievals. It is worth mentioning that the AEL-FM feature mask does not separate between different particle types but instead, it detects areas of strong and weak returns or those associated with clear sky conditions (van Zadelhoff et al., 2023). Based on the AEL-FM retrievals, the features of the probed atmospheric scene are classified either as clouds or aerosols or as molecular (Rayleigh) atmosphere. These referenced categories of scatterers, as well as their respective sub-types, are defined relying on image processing techniques and their availability along the Aeolus track facilitates the selection of the cloud-free Aeolus bins on the raw L2A profiles. In Table 3.1, the features that are categorized in the AEL-FM feature mask product are listed. In the first column, the main output of the AEL-FM feature mask product is given and corresponds to a feature detection probability index with values ranging from -3 to 10 and the second column presents how these indexes are defined. Based on the definitions, clear-sky conditions labeled with feature index value 0 are associated with very low signals expected to originate from clear air while stronger signals with values between 6 and 10 are most likely to have originated from liquid or optically thick ice clouds. Additionally, feature retrievals labeled with index value -3 are associated with signals directly affected by the surface.

Table 3.1: Aeolus feature-mask features' definition. The first column provides the feature detection probability index ranging from -3 to 10. The second column shows the definition of each index.

Index	Definition
10	Clouds
9	Most likely clouds
8	Very likely clouds or aerosols
7	More likely clouds or aerosols
6	Likely clouds or aerosols
5	Expected low altitude aerosol
4	Unlikely clouds or aerosol
3	Likely only molecules
2	Very likely only molecules
1	Most likely only molecules
0	Clear sky
-1	Fully Rayleigh attenuated
-2	No retrievals
-3	Surface data

Since the AEL-FM product is provided at the measurement level (~3 km), for the implementation of the cloud-screening procedure a conversion is required to be done in order to match the Aeolus' SCA horizontal and vertical resolution. The conversion process is illustrated in the figure below. In this example and due to space limitations only 3 BRC profiles are presented comprised with 3 measurements each. Depending on the L2A processor version, the total number of measurements accumulated in one BRC profile varies. According to the figure, in step 1 an example of the main output of the AEL-FM product is provided giving the feature detection probability indices with values ranging from -3 to 10. The conversion process (step 2) starts by putting the flag 'cloud' or "1" to the AEL-FM index values between 6 and 10 since according to the definitions given in Table 3.1 the associated returned signals are most likely to have been affected by clouds. The remaining AEL-FM index values from -3 to 5 were flagged with '0' since the associated returned signals are not likely to have been affected by clouds. This process gives us a binary matrix with the cloud-contaminated (1) and non-cloud-contaminated (0) measurements along the Aeolus track (step 3).

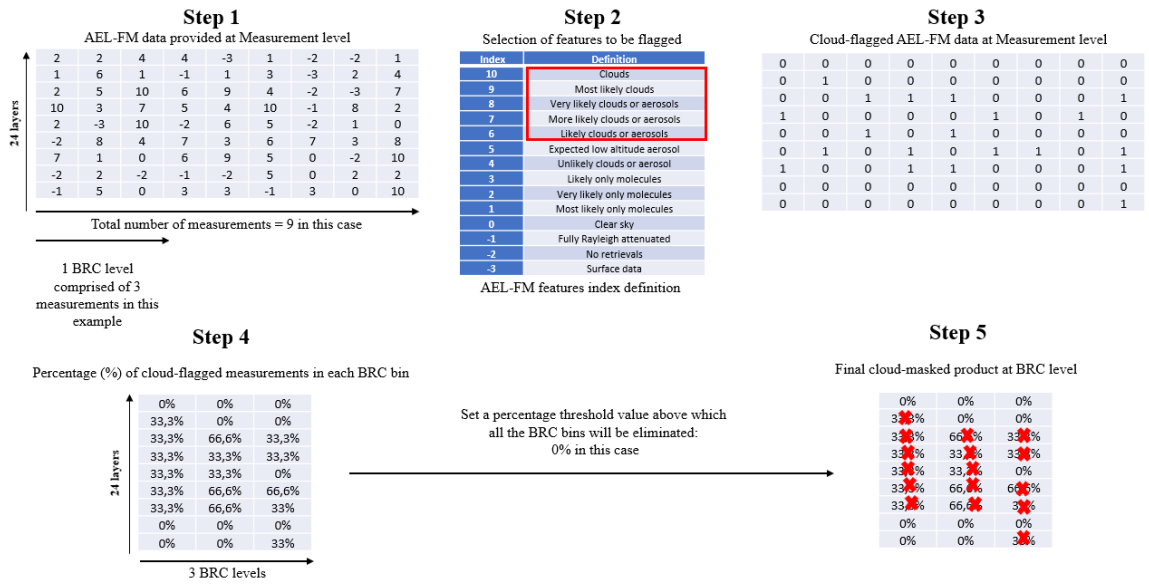

Figure 3.2: A graphical illustration of the cloud-filtering methodology using the AEL-FM feature mask product.

Then, in the next step (step 4) for each BRC bin comprised of three measurements each in this example, the total percentage of cloud-flagged measurements was computed. Based on the converted AEL-FM product (at the Aeolus' BRC level) giving the total percentage of "cloud-flagged" measurements for each BRC, the observations with values exceeding a threshold value (0% in this case) were considered as cloud-contaminated observations and the associated Aeolus SCA optical properties observations were eliminated (step 5). Following the aforementioned cloud-filtering process, the Aeolus L2A cloud-free backscatter profiles are derived.

For the first cloud-filtering methodology some preliminary results are presented for an indicative study case on 17th September 2021. Figure 3.3 shows the region of interest and the Aeolus satellite orbit path for the given case with the thick blue line indicating the ALADIN's measurement track.

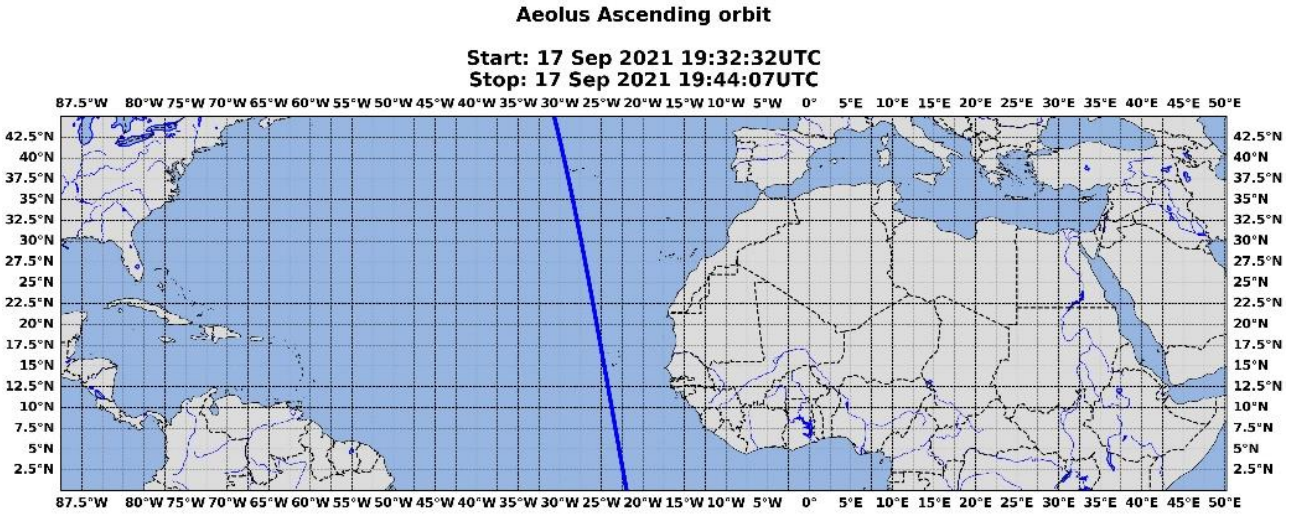

Figure 3.3: Region of Interest and the Aeolus overpass (orbit id: 017790) on 17th September 2021.

The cloud-filtering results for the specific Aeolus orbit are presented in the next figure (Figure 3.4). At first, the raw (unprocessed) Aeolus L2A profiles of the extinction and backscatter coefficient retrieved with the SCA algorithm are presented. Next, the primary AEL-FM feature-mask output along the Aeolus measurement track is shown where the classified features of the probed atmospheric scene can be observed. It can be seen that features associated with "strong" returns mainly attributed to clouds or high optically thick aerosol layers are colorized in brown and red respectively while those associated with the molecular atmosphere or clear sky conditions are colorized in green and cyan respectively. Then, for each BRC profile comprised of 30 measurements in this case the total percentage values of cloud-contaminated measurements are computed and presented in the next figure. Finally, based on the converted AEL-FM dataset given at the BRC level, the cloud-covered BRC bins exceeding 0% are detected and eliminated with the associated primary L2A Aeolus retrievals. In the last line of the figure the derived cloud-filtered extinction and backscatter SCA profiles are presented.

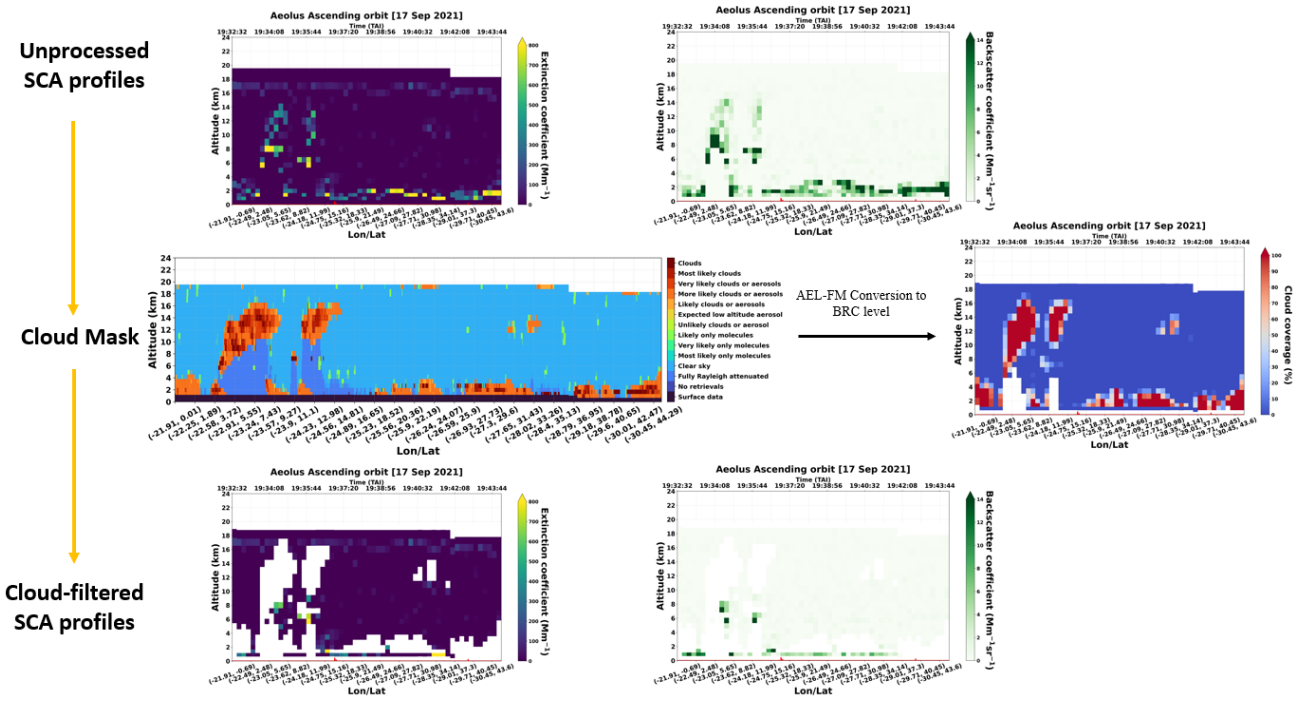

Figure 3.4: Unprocessed profiles of L2A SCA backscatter coefficient; AEL-FM Feature Mask product at measurement level and cloud-covered (in%) BRC bins; Cloud-Filtered profiles of SCA extinction and backscatter coefficient for the Aeolus overpass of orbit id: 017790 on 17 September 2021.

3.2.2. Cloud-filtering methodology based on MSG-SEVIRI CLAAS-3 cloud dataset.

For the second cloud screening process of the raw (unprocessed) Aeolus L2A profiles, the CLAAS-3 cloud dataset derived directly from the scene analysis of the MSG-SEVIRI (Meteosat Second Generation - Spinning Enhanced Visible and InfraRed Imager) imagery, will be processed. CLAAS-3 is the latest edition of CLAAS with previous editions documented in Stengel et al. (2014) and Benas et al. (2017). CLAAS-3 includes the following cloud properties: cloud mask/type, cloud top temperature/pressure/height, cloud thermodynamic phase, cloud optical thickness, cloud particle effective radius and cloud water path. Additionally, cloud droplet number concentration and cloud geometrical thickness are provided for liquid clouds. In cloud mask product, it is merely provided a categorization of the clear and cloudy pixels rather than a detailed information of clouds' macrophysical properties (i.e., cloud coverage), phase (i.e., ice, water, mixed) or types (i.e., low, middle, high). For the needs of the WP3000, the dataset has been provided for the whole study period of September 2021 from N. Benas - KNMI and S. Martin – DLR. An example of the MSG SEVIRI cloud dataset is provided in the figure below (Figure 3.5) for an indicative study case on 17th of September 2021 at 19.45 UTC where the grey and blue areas declare the cloud-covered and clear-sky areas. Additionally, an Aeolus ascending orbit on the specific date and time is also depicted.

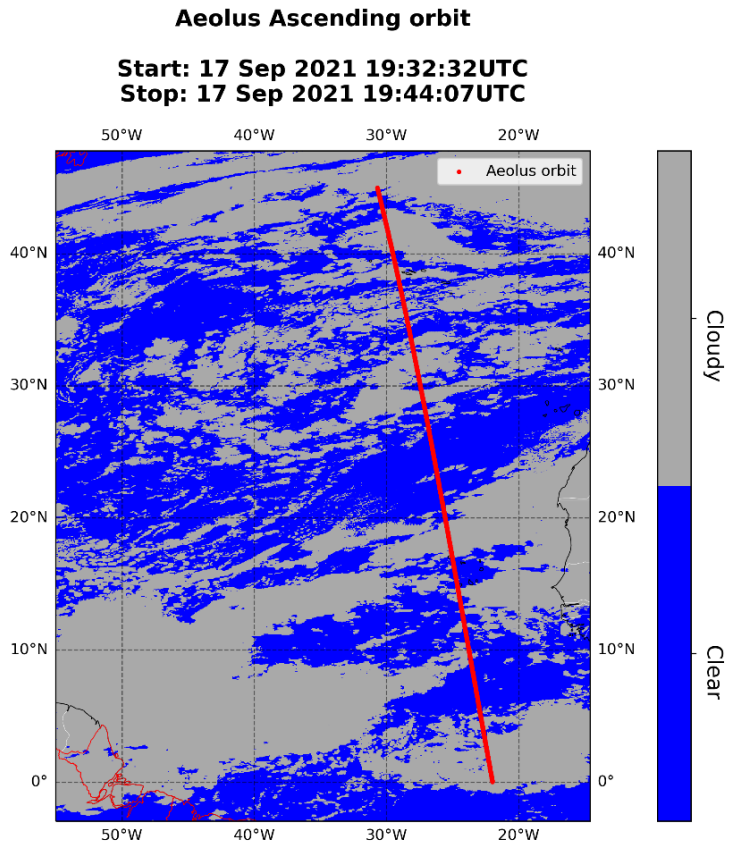

Figure 3.5: SEVIRI CLAAS-3 cloud mask product for an indicative study case on 17th of September 2021 at 19.45 UTC and Aeolus orbit.

For the identification of the cloud-contaminated Aeolus profiles, from the SEVIRI CLAAS-3 cloud mask dataset, the closest grid cells to the Aeolus ascending steps were selected. Then, for each grid cell with the flag "cloud" the associated Aeolus measurement step was flagged with the value of "1". Once the cloud-flagging process was completed, for each BRC profile the total percentage of cloud-contaminated measurements was computed. In the figure below the raw Aeolus L2A profiles of the extinction and backscatter coefficients retrieved with the SCA algorithm are presented. Accordingly, based on the MSG SEVIRI cloud dataset, the cloud-contaminated BRC profiles of the primary L2A Aeolus retrievals are detected and eliminated taking into account different test cases with different percentage values of cloud-covered BRC profiles.

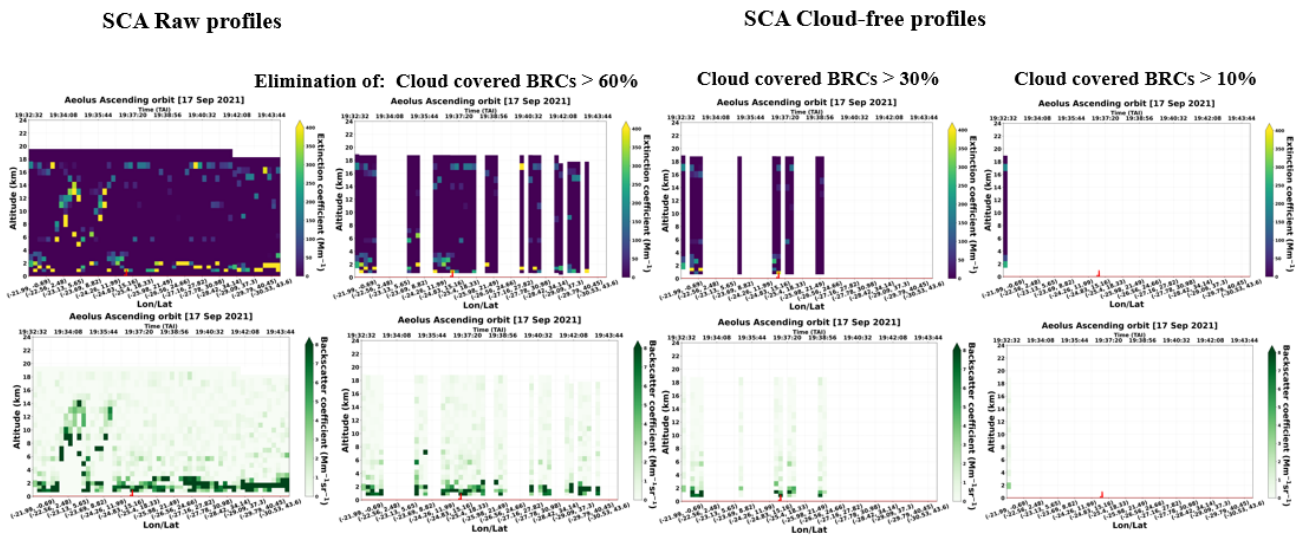

Figure 3.6: SCA raw profiles of extinction and backscatter coefficients for a study case on 17th of September 2021. The associated cloud-free profiles for different percentage values of cloud-covered BRC profiles are given.

3.2.3. Cloud-filtering methodology combining AEL-FM and MSG-SEVIRI CLAAS-3 cloud dataset products.

To achieve more rigorous filtering of the raw Aeolus L2A retrievals, both cloud-filtering approaches based on AEL-FM and MSG SEVIRI retrievals were combined and the pure aerosol profiles were derived. Note that for the cloud-filtering methodology based on the MSG-SEVIRI cloud dataset which has been described in detail in Section 3.2.2, after counting the total number of cloud-contaminated measurements belonging to each BRC profile, we excluded only those BRCs with cloud-contaminated measurements exceeding the threshold value of 60%. This is not quite a strict filtering but taking into account that both filtering methodologies have been combined at this time and a large number of BRC bins have already been excluded based on the AEL-FM feature mask product, we selected an appropriate threshold value in order to have a sufficient number of BRCs to work with for the derivation of the final L2A+ dust product. In the figure below, the totally cloud-filtered Aeolus L2A backscatter profiles derived from the SCA, SCA mid-bin and MLE algorithms are presented for an indicative Aeolus overpass on 17th September 2021 (orbit id: 017790).

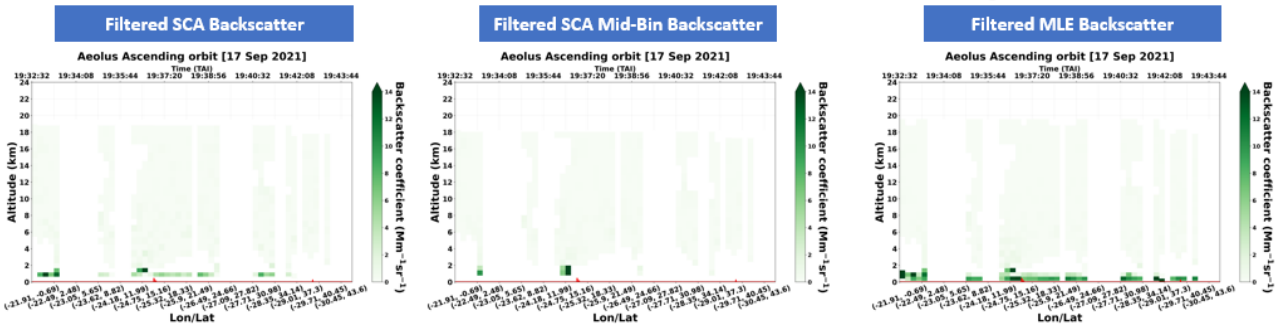

Figure 3.7: SCA, SCA mid-bin and MLE cloud-filtered profiles of backscatter coefficients for a study case on 17th of September 2021. The cloud-filtered profiles were derived using both AEL-FM and MSG-SEVIRI retrievals.

3.3. Dust discrimination using CAMS reanalysis dataset

In this phase, the altitude ranges of the dust aerosol layers are defined on the derived cloud-free Aeolus L2A+ backscatter coefficient profiles from the previous working phase. Due to the absence of an aerosol classification scheme on the raw Aeolus L2A data, numerical outputs from the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis (Inness et al., 2019) are implemented towards assignment of aerosol typing. The CAMS reanalysis, produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), consists of three-dimensional time-consistent atmospheric composition fields, including aerosols and chemical species (Inness et al., 2019) The CAMS aerosol model component of the IFS includes twelve prognostic tracers, consisting of three bins for sea salt of different size (0.03–0.5, 0.5–5 and 5–20 μm), three bins for dust (0.03–0.55, 0.55–0.9 and 0.9–20 μm), hydrophilic and hydrophobic organic matter and black carbon, sulphate aerosols plus its precursor trace gas of sulfur dioxide (SO2) (Morcrette et al., 2009). For aerosol typing, some initial processing steps of CAMS reanalysis outputs have been undertaken. Specifically, for the study period of September 2021, gridded data for dust aerosols were extracted for the entire Region of Interest (RoI). The data have a horizontal resolution of 0.5°, 60 hybrid sigma-pressure model levels on the vertical scale and a temporal resolution of 3h. An indicative study case is illustrated in Figure 3.8, presenting a Saharan dust outbreak on the 17th of September 2021, when dust-abundant air masses originating from N. Africa crossed the Tropical Atlantic Ocean with the red-coloured line presenting the closest Aeolus's ascending orbit on the specific time.

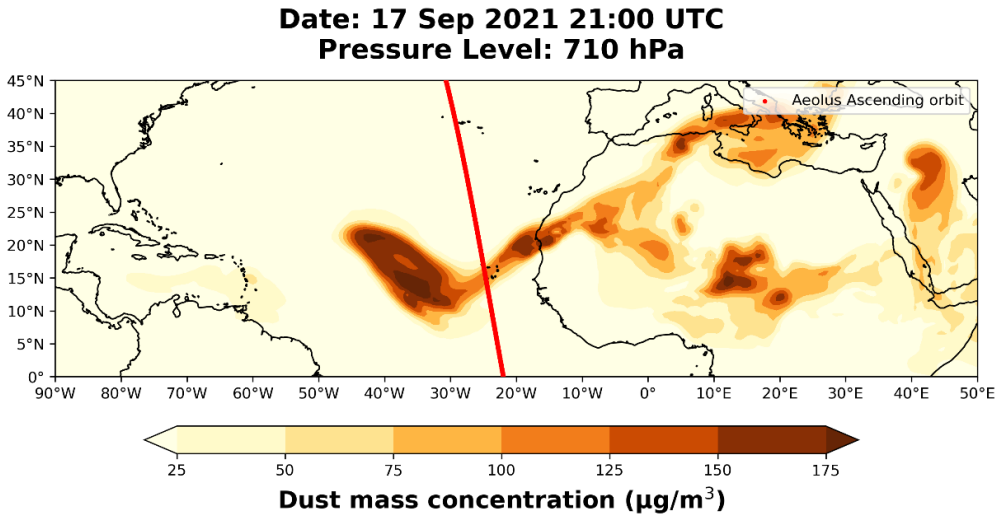

Figure 3.8: CAMS dust mass concentration over L2A+ RoI for an indicative study case on the 17th September 2021. The red-coloured line presents the closest Aeolus orbit on the referenced time.

Next, for the dust-typing, CAMS data were spatiotemporally collocated and vertically integrated to derive an Aeolus-type dust product from CAMS. In Figures 3.9 a,c, the vertical profiles of dust mass concentration from CAMS along the specific Aeolus orbit (id: 017790) on the 17th of September 2021 are presented separately for the regular (24 vertical bins) and middle bin (23 vertical bins) vertical scale where we can clearly observe specific BRC bins over the latitudinal band of 5°- 15°N up to 6km with elevated dust mass concentration values exceeding in many cases the value of 50 μg/m^3. For a more robust and detailed discrimination of the pure dust layers, the proportion of dust in the total aerosol mass concentration was also used by computing the dust-to-total aerosol mass concentration ratio. Figures 3.9 b and d present the vertical profiles of the dust-to-total mass concentration ratio values (expressed in percentage) along the specific Aeolus overpass (orbit: 017790). Then, the correction of the Aeolus cloud-free L2A backscatter profiles was focused only on the dust-contaminated BRC bins with dust mass concentrations above the median value of 1.3 μg/m^3 and dust-to-total mass concentration values exceeding 50%.

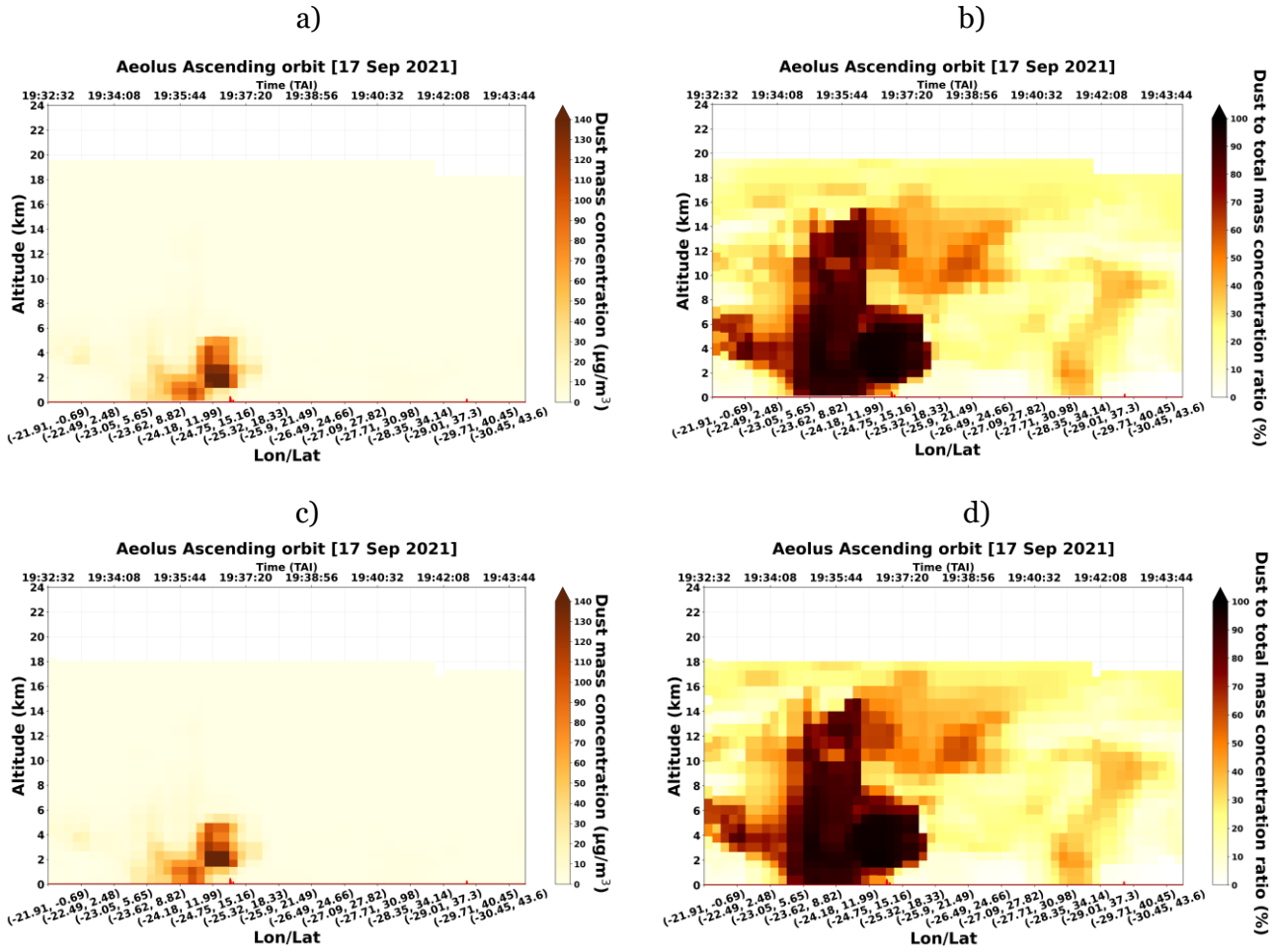

Figure 3.9: Vertical profiles of CAMS dust mass concentration and dust-to-total aerosol mass concentration ratio along the Aeolus orbit (id: 017790) provided in the regular (a, b) and middle-bin (c, d) vertical scales on the 17th of September 2021.

3.4. Reconstruction of the L2A extinction profiles

In this part, the Aeolus L2A cloud-free extinction profiles were reconstructed focusing on the pure dust profiles identified from CAMS reanalysis dataset (Section 3.3). For the derivation of the new L2A+ extinction profiles, the Aeolus L2A cloud-free backscatter profiles (co-polar part) were converted to total backscatter by accounting at first the missing cross-polar component. The formula used to convert between the 355 nm co-polar backscatter and total backscatter is the following:

$$\beta_{355}^{co} = \frac{\beta_{355}^{total}}{1 + \delta_{355}^{circ}} \qquad (1)$$

where $\beta_{355\,nm}^{co}$ is the measured Aeolus 355 nm co-polar backscatter coefficient, $\beta_{355\,nm}^{total}$ the 355 nm total backscatter coefficient and δ_{355nm}^{circ} is the particle circular depolarization ratio at 355 nm. It is noteworthy to mention that since the particle circular depolarization ratio at 355 nm (δ_{355nm}^{circ}) is not

directly measured, it can be computed indirectly using the particle linear depolarization ratio at 355 nm via the following conversion formula:

$$\delta_{355}^{\text{circ}} = \frac{2\,\delta_{355}^{\text{linear}}}{1 - \delta_{355}^{\text{linear}}} \qquad (2)$$

where the value of particle linear depolarization ratio at 355 nm ($\delta_{355}^{\text{linear}}$) for mineral dust particles was selected according to the DeLiAn database (Floutsi et al., 2023). Finally, the new L2A+ extinction profiles were computed using the estimated total backscatter coefficient and a lidar ratio of 53.5 sr at 355 nm for Saharan dust (Floutsi et al., 2023). Figure 3.10 gives a graphical illustration of the reconstruction process of the L2A pure-dust extinction profiles including the pure dust Aeolus SCA co-polar backscatter profiles, the converted total SCA backscatter profiles after adjusting the mis detected cross-polar component of the backscattered lidar signal and a lidar ratio for Saharan dust (53.5 sr). Note that the aforementioned methodology has been implemented for all the available Aeolus retrieval algorithms and specifically for the SCA, SCA mid-bin and MLE algorithms (all included in the latest processor version of Baseline 16) and for each of them, the L2A+ extinction profiles have been derived.

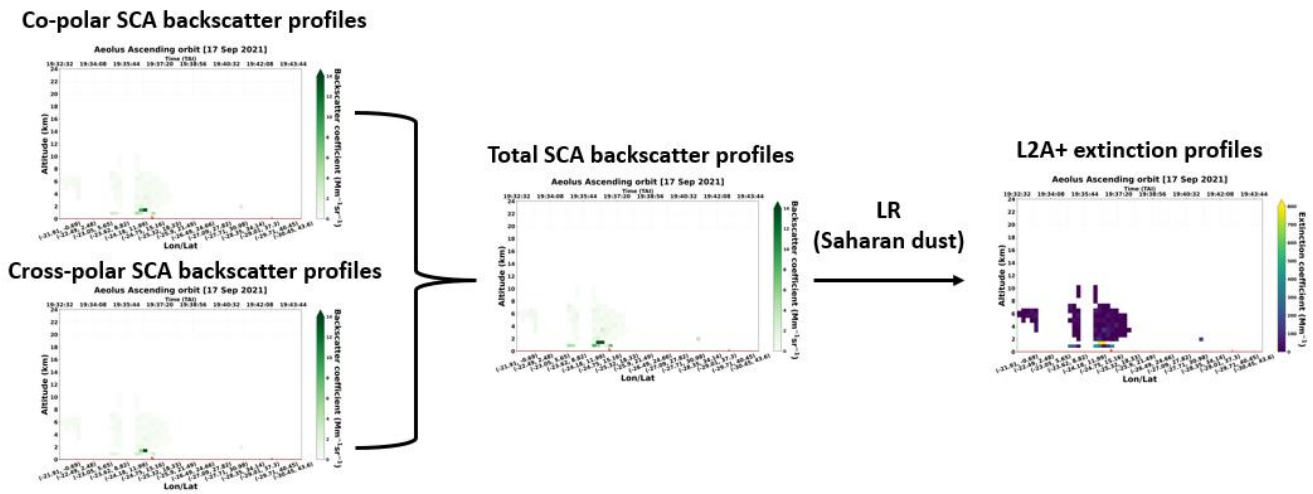

Figure 3.10: Illustration of the L2A extinction profiles reconstruction process

3.5. Derivation of the dust mass concentration using the POLIPHON method

This is the last step of the WP3000 which includes the conversion of the L2A+ extinction profiles to dust mass concentration using the POLIPHON method. A more detailed description of the POLIPHON method can be found in Mamouri and Ansmann (2014; 2015; 2016; 2017), Ansmann and Mamouri (2019), and concerning the INP concentration retrieval in Marinou et al. (2019). The core concept of the methodology is briefly outlined in the current section.

POLIPHON (Polarization Lidar Photometer Networking) is a powerful technique that allows the separation of mineral dust and non-dust aerosols by combining the unique capabilities of the polarization lidar with the well-established global aerosol climatology of aerosol optical and microphysical properties provided by AERONET (Aerosol Robotic Network) (Holben et al., 1998). The first step of the POLIPHON technique includes the analysis of the polarization lidar observations based on the different polarization properties of the different aerosol types to obtain vertically-

resolved profiles of dust and non-dust backscatter coefficients, and thus the backscatter-related dust fraction (Ansmann et al., 2019). Then, the derived dust and non-dust backscatter coefficients are converted to respective dust and non-dust extinction coefficients using appropriate dust extinction-to-backscatter ratios or lidar ratios. In the second step of the POLIPHON technique, the vertically-resolved profile of the dust mass concentration (both fine and coarse dust particles) is derived using the respective coarse- and fine-dust extinction coefficients and extinction-to-volume conversion factors determined from AERONET observations (Ansmann et al., 2019). Additional products that can be retrieved using the POLIPHON technique include the particle number, surface area, volume concentration for dust and non-dust aerosol components as well as meteorological parameters such as cloud condensation nuclei (CCN) and ice nucleating particle (INP) concentrations.

For our study, the POLIPHON method was implemented and the mass concentration of dust was derived using:

$$M_d = p_d * v_d \quad (3)$$

where p_d is the dust particle density (2.6 g cm^{-3}) and v_d the dust volume concentration. The latter was estimated using

$$v_d = c_{d,532} * \sigma_{d,532} \quad (4)$$

with $\sigma_{d,532}$ expressing the L2A+ extinction coefficient at 532 nm and $c_{d,532}$ the extinction-to-volume conversion factor at 532 nm derived from the AERONET observations. All the AERONET-based conversion parameters required in the conversion of particle extinction coefficient at 532 nm into particle number, surface area and volume concentration are described in Table 4 of Ansmann et al. (2019). Before using Eq. (4) for the derivation of the dust volume concentration, a previous step was the conversion of the measured L2A+ extinction coefficient from 355 nm to 532 nm using the empirical relation introduced by Ångström (1929):

$$\frac{\sigma_{\lambda 1}}{\sigma_{\lambda 2}} = \left(\frac{\lambda 1}{\lambda 2}\right)^{-\alpha} \quad (5)$$

where λι and λ2 the two wavelengths of 355 and 532 nm respectively and α with value 0.1 is the selected Ångström exponent for Saharan dust derived from DeLiAn database presented in (Floutsi et al., 2023). Using Eq. (5) the L2A+ extinction coefficient at 532 nm was derived and then it was introduced in Eq. (4) for the derivation of the dust volume concentration before the final estimation of the dust mass concentration (Eq. 3). Figure 3.11 illustrates the results of the final step on which the pure dust Aeolus L2A+ extinction profiles are converted from 355 to 532 nm before being converted to mass concentration using the POLIPHON method.

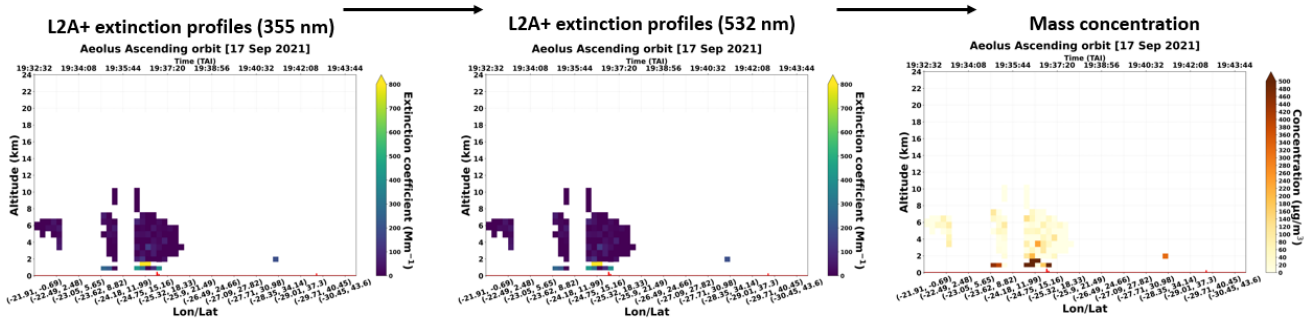

Figure 3.11: L2A+ extinction profiles at a) 355 nm, b) 532 nm, and c) the final dust mass concentration profiles.

4. ESA-L2A+ WP4000: Assimilation of L2A/L2A+ and application of WRF-L experiments

4.1. Description of WP4000.

The aim of this work package is to carry out Data Assimilation (DA) experiments to study the impact of joint Aeolus wind & aerosol information in a regional model, as well as determine the improvements brought by the development of the L2A+ product. As such, work was focused on setting up a regional dust transport model (WRF-CHEM) for assimilation (with NCAR DART) and incorporating the necessary tools and algorithms to assimilate Aeolus data. In this section we begin by describing our assimilation approach and how variables are transformed from model space to observation space and conclude by briefly describing the 3rd-party tools used and presenting their configuration.

4.2. Data Assimilation Overview.

The overarching goal of any assimilation system is to estimate the probability distribution function (PDF) of the Earth system state at the initial time. To achieve this goal, a forecast is corrected with information from observations to provide a better estimate of the system state, called the analysis. This general idea is presented visually in Figure 1.

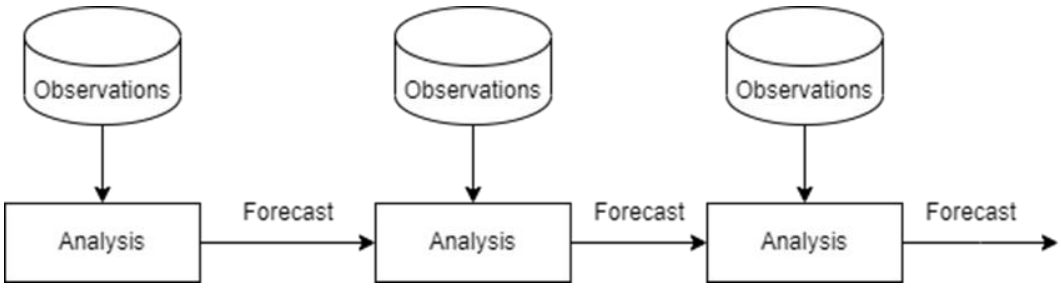

Figure 4.1: A graphical presentation of the WP3000 processing steps.

Assimilation systems can be intuitively grasped by looking at the problem from a Bayesian perspective: we aim to update our prior knowledge of the Earth system state distribution (prior, also called the background), to produce a posterior of the distribution (also called the analysis) as the conditional probability of the prior given the observations. There are two major schools of thought for tackling this problem. The first, Variational Data Assimilation, includes popular algorithms such as 3D-Var and 4D-Var, work by determining the best possible model trajectory (in a "least square" sense) that fits the available observation data. The second school of algorithms, called Ensemble Kalman Filters (EnFK), operate by using an ensemble of forecasts at each time step, which is treated as a random draw from the probability distribution of the model state. This probability distribution is then updated using information from the observations. Because the quantities that we can observe, especially using space-borne instruments like ATLID, are not necessarily part of the model state, assimilation systems use "observational operators" that calculate the expected observation from the model state. These operators are also known as the "forward model" in retrieval problems.

While most operational Numerical Weather Prediction systems rely on variational methods, in L2A+ the assimilation experiments are conducted using EnKF because of a distinct advantage. Variational systems require linearized versions of both the forecast model and the observational operators, something that requires significant development effort to achieve. In contrast, EnKF methods do not have this requirement, allowing rapid prototyping of a system that can assimilate new and novel

observations. An important disadvantage of EnKF is the requirement to use an ensemble of N independent forecasts (members), which is computationally demanding. However, this issue is less of an obstacle in research scenarios where there is no "real-time" requirement (Bannister, 2017).

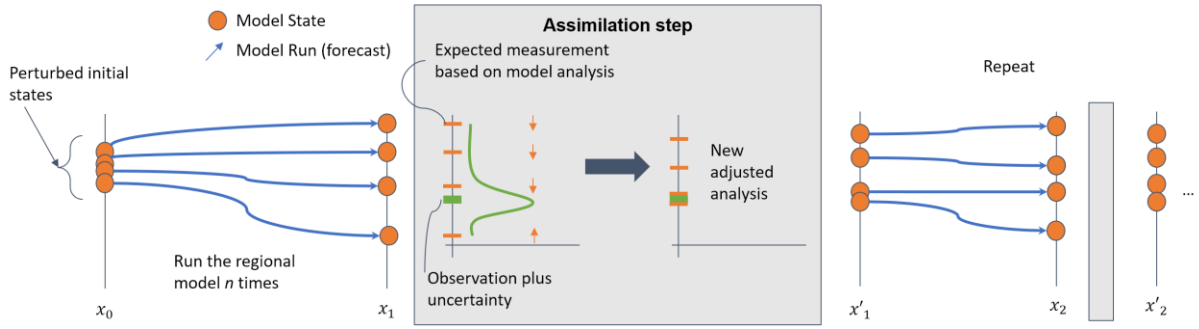

Figure 4.2: A schematic showing how an Ensemble Kalman Filter data assimilation system operates.

Figure 4.2 shows a high-level overview of how EnKF works. Let x_t be the model state vector at time t and y_t denote the available observations, close to time t (exactly how close is a tunable parameter). An intuitive way to imagine the EnKF method is to think of it as an update of the PDF of x through the information in the PDF of y, like the Bayes theorem. So, to compute the posterior (analysis) x'_t, we would compute at all grid points:

$$x'_t = x_t + K(y - H(x_t)) \quad (1)$$

As aforementioned, since our observations do not concern physical variables that are part of the model state, it is necessary to compute the expected observation from our model state. This observation operator is denoted with $H(\cdot)$.

For each kind of observation, the appropriate operator must be developed to map the model state to observation space. For Aeolus, this is developed during the project and described in a later subsection. At each timestep, the analysis (x'_t) is used to initialize the model for the next timestep, a technique known as cycling. The above description was an introduction into the intuition of the EnKF method, the exact formulation can be studied in J. Anderson (2001).

4.3. Experiment setup in L2A+.

In L2A+ the regional numerical weather prediction model used is WRF-CHEM (Advanced Research Weather version of the Weather Research and Forecasting coupled with Chemistry) (Grell et al., 2005; Skamarock et al., 2008), Georgia Institute of Technology-Goddard Global Ozone Chemistry Aerosol Radiation and Transport (GOCART) aerosol model and the Air Force Weather Agency (AFWA) dust emission scheme (LeGrand et al., 2019) for studying the life cycle of mineral aerosols. The Data Assimilation Research Testbed (DART) by the National Center for Atmospheric Research (NCAR) is used for its EnKF implementation (Anderson et al., 2004). Specifically, the Ensemble Adjustment Kalman Filter (EAKF) (Anderson et al., 2001) implementation is selected for its ability to avoid the filtering divergence problem.

In technical terms, the experiment setup requires running N independent forecasts with WRF-CHEM for a period of time (a cycle), at which point DART is used to assimilate Aeolus observations. The resulting analysis is used as initial conditions for starting the forecasts again for the next cycle. The number of ensemble members (N) and the cycle length are yet to be finalised. This procedure is shown visually in Figure 4.3.

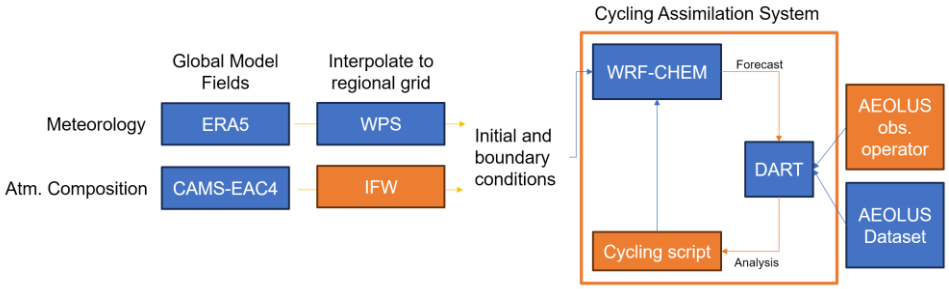

Figure 4.3: Experiment setup with WRF and DART. Blocks and arrows colored orange denote developments that happened in L2A+, specifically the chemistry pre-processor (IFW), the cycling assimilation system (which includes management of the model ensembles, experiment configuration, the cycling script and HPC automations), and the AEOLUS observational operators.

The model domain is shown in Figure 4.4. It covers the general North Atlantic region, parts of Africa which are important for their dust emissions and parts of North and South America which are receptor points for dust transport. The exact domain boundaries and the exact configurations for WRF and DART are subject to change as the experiments continue.

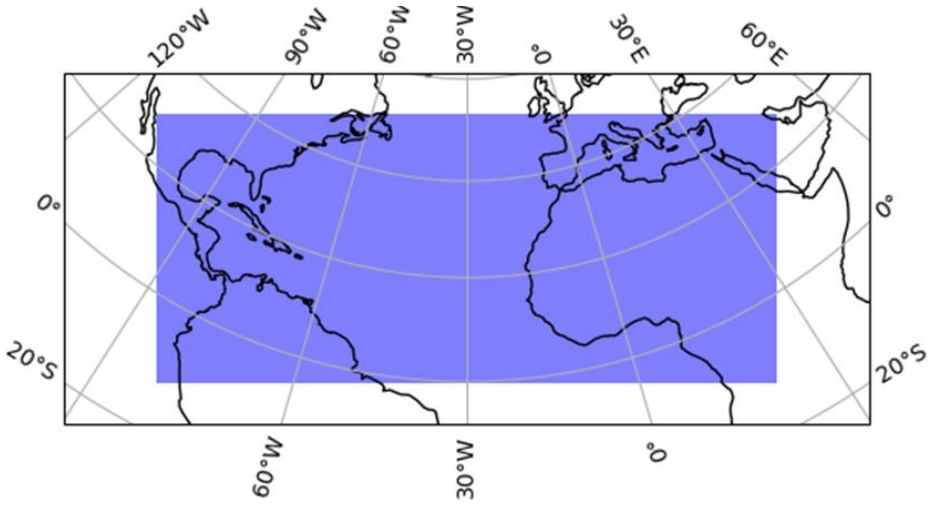

Figure 4.4: Model domain shaded in blue.

The ensemble of forecasts used in EnKF reassembles a random draw from the Earth system state probability distribution. Thus, we expect the forecasts of the ensemble member to differ more in areas of the Earth at which the model is less certain about the prediction. The numerical models used are deterministic, i.e. they give a specific output for a given input. In order to encourage the forecasts to "spread" appropriately and give us an estimate of the uncertainty, perturbations are applied to the initial conditions of each member. The algorithm for creating spatially correlated perturbation fields is based on a technique described in Tsikerdekis et. al. (2021, supplement) and consists of the following steps:

1. For each variable to be perturbed, generate a random field of the same dimensions by sampling from a Gaussian distribution with known mean and standard deviation
2. For n steps, apply a Gaussian filter (i.e. rolling average window) of size 1 (3 × 3).
3. Standardize the distribution of the perturbation field as to have a mean value of 1, then scale it to get the desired mean perturbation.
4. Add the perturbation field to the variable being perturbed.

Each ensemble member gets different perturbation fields, which are also stored for diagnostic reasons. When comparing between two experiments (e.g., a control run against a run with Aeolus assimilation), the same perturbation fields are used. The list of perturbed variables (beyond the wind elements U and V) is still under consideration.

The forecast model is initialized using ERA5 re-analysis meteorological fields (Hersbach et. al, 2020) and CAMS global reanalysis (EAC4) aerosol fields (Inness et. al, 2019). The initial condition fields must be interpolated to the regional model's grid before use and while the ERA5 meteorological fields are supported by WRF's Preprocessing System (WPS), there is no support for aerosol fields. In addition, EAC4 has three dust aerosol size bins compared to AWFA-GOCART's (the dust scheme used in WRF-CHEM) five. To enable use of EAC4, an application was developed to 1) interpolate the EAC4 fields onto the regional model grid, and 2) map the EAC4 dust species to AFWA-GOCART's species. The application is developed in Python and is generic enough to allow interpolation of any global chemistry model fields onto a WRF grid if you can implement a file reading routine in Python. The application is called "interpolator-for-wrfchem" (henceforth IFW) and is available on Github.

The methodology of IFW is based on the WPS, with the additional step of doing the species mapping. The distinct steps are:

1. For each variable of the source aerosol fields
    a. Horizontally interpolate each CAMS level (2D) to the WRF lat-lon grid.
    b. Vertically interpolate each CAMS column to the WRF vertical levels.
2. Map the source model species to the target model species.

The above procedure is performed once for the initial fields and one time for each of the regional model boundaries.

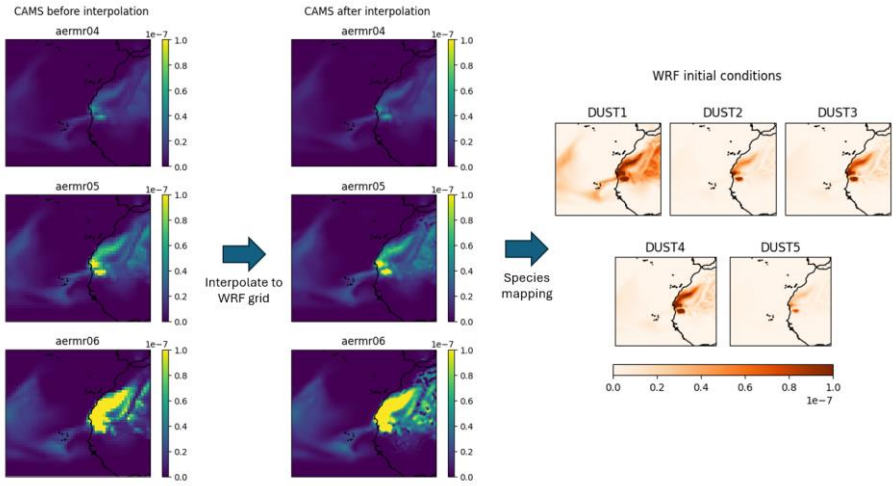

Figure 4.5: The aerosol field preprocessing methodology shown visually. All values refer to mixing ratios (kg/kg).

The mapping of WRF-CHEM species to AFWA-GOCART is constructed using a mass size distribution derived from the mean lognormal number size distribution, which is fitted to the in-situ size distributions measurements inside the Saharan Air Layer (SAL), during the AER-D campaign (Table 5 in Ryder et al., 2018, Table XX1 herein). The conversion factors (Table XX2) for each WRF-Chem model size bin ($D_1 = 0.2 - 2.0$ μm, $D_2 = 2.0 - 3.6$ μm, $D_3 = 3.6 - 6.0$ μm, $D_4 = 6.0 - 12.0$ μm, $D_5 = 12.0 - 20.0$ μm) are calculated as the fraction of the mass contained in each CAMS model size bin range ($D_{CAMS1} = 0.06 - 1.1$ μm, $D_{CAMS2} = 1.1 - 1.8$ μm, $D_{CAMS3} = 1.8 - 40$ μm). The density of dust particles is assumed equal to 2610 kg m^{-3}.

Table 4.1: Lognormal mode properties for the mean number size distribution, fitted to the in-situ measurements, taken within the SAL during AER-D campaign. d_{pg} stands for the median diameter, σ_g for the geometric mean, and N for the total aerosol concentration.

	MODE 1	MODE 2	MODE 3	MODE 4
D_{PG}	0.150	0.851	1.850	32.527
Σ	2.2	1.181	1.928	1.528
N	$3.91x10^{2}$	$8.39x10^{0}$	$1.16x10^{1}$	$1.38x10^{-4}$

Table 4.2: Calculation of mixing ratios at the AFWA-GOCART size bins using linear combinations of components from CAMS-EAC4.

Dust Bin	CAMSD1	CAMSD2	CAMSD3
DUST1	0.96	1.0	0.16
DUST2			0.19
DUST3			0.30
DUST4			0.35
DUST5			0.11

4.4. Operators.

4.4.1. L2B Wind HLOS observations.

Aeolus measures horizontal line-of-sight (HLOS) wind speeds, which must be compared with the model's U and V wind vectors. The following observational operator is used (Šavli et al., 2018, Krisch et al. 2022), where a is the azimuth angle, defined clockwise from north:

$$\mathrm{HLOS} = \mathrm{U}\sin a + \mathrm{V}\cos a \tag{3}$$

This operator is implemented in DART, alongside a program that reads AEOLUS DBL files and converts them to DART intermediate format, as required for assimilation. These developments are available on a Github fork and will be contributed upstream when validation is complete.

4.4.2. L2A and L2A+ aerosol observations.

In order to assimilate the extinction coefficient profiles of the L2A and L2A+ products, an operator must be used to calculate extinction coefficient profiles from the model state. In our case, the model state consists of dust mixing ratio at each model grid point, at distinct size bins. For every cell, the dust extinction coefficient EC is calculated as:

$$\mathrm{EC} = \sum_{1}^{k} \frac{3}{2\rho_k D_{\mathrm{eff}}} \mathrm{M_k Q}$$

where $\mathrm{M_k}$ is the dust concentration for size bin k in $\mathrm{g\,m^{-3}}$ (calculated as the product of the corresponding dust mixing ratio and air density), ρ_k is the particle density ($\mathrm{g\,m^{-3}}$), D_{eff} is the effective diameter (m), and Q is the extinction efficiency factor.

The extinction efficiency factor Q is wavelength-dependent and thus must be provided at 355nm, corresponding to the operation wavelength of ATLID. This calculation is performed using the Mie scattering code (Mie, 1908), assuming spherical particles and a refractive index of $1.55 + 0.005i$, which is considered representative of dust (Dubovik et al., 2002). While the spherical assumption for dust is not perfect, as dust particles are known to have irregular shapes, the focus on this study is not to push the boundaries of dust representation in operators.

List of Figures

List of Tables

References

Ångström, A.: On the atmospheric transmission of sun radiation and on dust in the air, Geograf. Ann., 11, 156–166, 1929.

Ansmann, A., Mamouri, R.-E., Hofer, J., Baars, H., Althausen, D., & Abdullaev, S. F. (2019). Dust mass, cloud condensation nuclei, and ice-nucleating particle profiling with polarization lidar: updated POLIPHON conversion factors from global AERONET analysis. Atmospheric Measurement Techniques, 12(9), 4849–4865. https://doi.org/10.5194/amt-12-4849-2019

Benas, N., Finkensieper, S., Stengel, M., van Zadelhoff, G.-J., Hanschmann, T., Hollmann, R., & Meirink, J. F. (2017). The MSG-SEVIRI-based cloud property data record CLAAS-2. Earth System Science Data, 9(2), 415–434. https://doi.org/10.5194/essd-9-415-2017.

C. Skamarock et al., 'A Description of the Advanced Research WRF Version 3', 2008, doi: 10.5065/D68S4MVH.

Dubovik, O., B. N. Holben, T. Lapyonok, A. Sinyuk, M. I. Mishchenko, P. Yang, and I. Slutsker. "Non-Spherical Aerosol Retrieval Method Employing Light Scattering by Spheroids." Geophysical Research Letters 29, no. 10 (2002): 54-1-54–4. https://doi.org/10.1029/2001GL014506.

Ehlers, F., Flament, T., Dabas, A., Trapon, D., Lacour, A., Baars, H., & Straume-Lindner, A. G. (2022). Optimization of Aeolus' aerosol optical properties by maximum-likelihood estimation. Atmospheric Measurement Techniques, 15(1), 185–203. https://doi.org/10.5194/amt-15-185-2022

Floutsi, A. A., Baars, H., Engelmann, R., Althausen, D., Ansmann, A., Bohlmann, S., Heese, B., Hofer, J., Kanitz, T., Haarig, M., Ohneiser, K., Radenz, M., Seifert, P., Skupin, A., Yin, Z., Abdullaev, S. F., Komppula, M., Filioglou, M., Giannakaki, E., … Wandinger, U. (2023). DeLiAn – a growing collection of depolarization ratio, lidar ratio and Ångström exponent for different aerosol types and mixtures from ground-based lidar observations. Atmospheric Measurement Techniques, 16(9), 2353–2379. https://doi.org/10.5194/amt-16-2353-2023

G. A. Grell et al., 'Fully coupled "online" chemistry within the WRF model', Atmospheric Environment, vol. 39, no. 37, pp. 6957–6975, Dec. 2005, doi: 10.1016/j.atmosenv.2005.04.027.

Hersbach, Hans, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, et al. "The ERA5 Global Reanalysis." Quarterly Journal of the Royal Meteorological Society 146, no. 730 (2020): 1999–2049. https://doi.org/10.1002/qj.3803.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – a federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, 1998.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., & Suttie, M. (2019).

The CAMS reanalysis of atmospheric composition. Atmospheric Chemistry and Physics, 19(6), 3515–3556. https://doi.org/10.5194/acp-19-3515-2019.

J. L. Anderson, T. Hoar, K. Raeder, and N. Collins, 'Data Assimilation Research Testbed'. UCAR/NCAR - Computational and Information Systems Laboratory (CISL), 2004. doi: 10.5065/D6WQ0202.

J. L. Anderson, 'An Ensemble Adjustment Kalman Filter for Data Assimilation', Monthly Weather Review, vol. 129, no. 12, pp. 2884–2903, Dec. 2001, doi: 10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

Krisch, Isabell, Neil P. Hindley, Oliver Reitebuch, and Corwin J. Wright. "On the Derivation of Zonal and Meridional Wind Components from Aeolus Horizontal Line-of-Sight Wind." Atmospheric Measurement Techniques 15, no. 11 (June 10, 2022): 3465–79. https://doi.org/10.5194/amt-15-3465-2022.

Mamouri, R. E. and Ansmann, A.: Fine and coarse dust separation with polarization lidar, Atmos. Meas. Tech., 7, 3717–3735, https://doi.org/10.5194/amt-7-3717-2014, 2014.

Mamouri, R. E. and Ansmann, A.: Estimated desert-dust ice nuclei profiles from polarization lidar: methodology and case studies, Atmos. Chem. Phys., 15, 3463–3477, https://doi.org/10.5194/acp-15-3463-2015, 2015.

Mamouri, R.-E. and Ansmann, A.: Potential of polarization lidar to provide profiles of CCN- and INP-relevant aerosol parameters, Atmos. Chem. Phys., 16, 5905–5931, https://doi.org/10.5194/acp-16-5905-2016, 2016.

Mamouri, R.-E. and Ansmann, A.: Potential of polarization/Raman lidar to separate fine dust, coarse dust, maritime, and anthropogenic aerosol profiles, Atmos. Meas. Tech., 10, 3403–3427, https://doi.org/10.5194/amt-10-3403-2017, 2017.

Marinou, E., Tesche, M., Nenes, A., Ansmann, A., Schrod, J., Mamali, D., Tsekeri, A., Pikridas, M., Baars, H., Engelmann, R., Voudouri, K.-A., Solomos, S., Sciare, J., Groß, S., Ewald, F., and Amiridis, V.: Retrieval of ice-nucleating particle concentrations from lidar observations and comparison with UAV in situ measurements, Atmos. Chem. Phys., 19, 11315–11342, https://doi.org/10.5194/acp-19-11315-2019, 2019.

Mie, Gustav. "Beiträge Zur Optik Trüber Medien, Speziell Kolloidaler Metallösungen." Annalen Der Physik 330, no. 3 (1908): 377–445. https://doi.org/10.1002/andp.19083300302.

Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J. W., Razinger, M., Schulz, M., Serrar, S., Simmons, A. J., Sofiev, M., Suttie, M., Tompkins, A. M., & Untch, A. (2009). Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling. Journal of Geophysical Research, 114(D6), D06206. https://doi.org/10.1029/2008JD011235.

M. Šavli, N. Žagar, and J. L. Anderson, 'Assimilation of horizontal line-of-sight winds with a mesoscale EnKF data assimilation system', Quarterly Journal of the Royal Meteorological Society, vol. 144, no. 716, pp. 2133–2155, 2018, doi: 10.1002/qj.3323.

Ryder, Claire L., Franco Marenco, Jennifer K. Brooke, Victor Estelles, Richard Cotton, Paola Formenti, James B. McQuaid, et al. "Coarse-Mode Mineral Dust Size Distributions, Composition and Optical Properties from AER-D Aircraft Measurements over the Tropical Eastern Atlantic." Atmospheric Chemistry and Physics 18, no. 23 (December 6, 2018): 17225–57. https://doi.org/10.5194/acp-18-17225-2018.

R. N. Bannister, 'A review of operational methods of variational and ensemble-variational data assimilation', Quarterly Journal of the Royal Meteorological Society, vol. 143, no. 703, pp. 607–633, 2017, doi: 10.1002/qj.2982.

Šavli, Matic, Nedjeljka Žagar, and Jeffrey L. Anderson. "Assimilation of Horizontal Line-of-Sight Winds with a Mesoscale EnKF Data Assimilation System." Quarterly Journal of the Royal Meteorological Society 144, no. 716 (2018): 2133–55. https://doi.org/10.1002/qj.3323.

Stengel, M., Kniffka, A., Meirink, J. F., Lockhoff, M., Tan, J., & Hollmann, R. (2014). CLAAS: the CM SAF cloud property data set using SEVIRI. Atmospheric Chemistry and Physics, 14(8), 4297–4311. https://doi.org/10.5194/acp-14-4297-2014.

S. L. LeGrand, C. Polashenski, T. W. Letcher, G. A. Creighton, S. E. Peckham, and J. D. Cetola, 'The AFWA dust emission scheme for the GOCART aerosol model in WRF-Chem v3.8.1', Geoscientific Model Development, vol. 12, no. 1, pp. 131–166, Jan. 2019, doi: 10.5194/gmd-12-131-2019.

Tsikerdekis, Athanasios, Nick A. J. Schutgens, and Otto P. Hasekamp. "Assimilating Aerosol Optical Properties Related to Size and Absorption from POLDER/PARASOL with an Ensemble Data Assimilation System." Atmospheric Chemistry and Physics 21, no. 4 (February 23, 2021): 2637–74. https://doi.org/10.5194/acp-21-2637-2021.

van Zadelhoff, G.-J., Donovan, D. P., & Wang, P. (2023). Detection of aerosol and cloud features for the EarthCARE lidar ATLID: the A-FM product. EGUsphere, 2023, 1–29.

[End of ESA-L2A+ DI03 - ALGO]