

The definition of an atmospheric database for Aeolus

G. J. Marseille, K. Houchi, J. de Kloe, and A. Stoffelen

Weather Research Department of the Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

Received: 12 April 2010 – Published in Atmos. Meas. Tech. Discuss.: 18 May 2010

Revised: 3 January 2011 – Accepted: 6 January 2011 – Published: 21 January 2011

Abstract. The definition of an atmospheric database is an important component of simulation studies in preparation of future earth observing remote sensing satellites. The Aeolus mission, formerly denoted Atmospheric Dynamics Mission (ADM) or ADM-Aeolus, is scheduled for launch end of 2013 and aims at measuring profiles of single horizontal line-of-sight (HLOS) wind components from the surface up to about 32 km with a global coverage. The vertical profile resolution is limited but may be changed during in-orbit operation. This provides the opportunity of a targeted sampling strategy, e.g., as a function of geographic region. Optimization of the vertical (and horizontal) sampling strategy requires a characterization of the atmosphere optical and dynamical properties, more in particular the distribution of atmospheric particles and their correlation with the atmospheric dynamics. The Aeolus atmospheric database combines meteorological data from the ECMWF model with atmosphere optical properties data from CALIPSO. An inverse algorithm to retrieve high-resolution particle backscatter from the CALIPSO level-1 attenuated backscatter product is presented. Global weather models tend to underestimate atmospheric wind variability. A procedure is described to ensure compatibility of the characteristics of the database winds with those from high-resolution radiosondes. The result is a high-resolution database of zonal, meridional and vertical wind, temperature, specific humidity and particle and molecular backscatter and extinction at 355 nm laser wavelength. This allows the simulation of small-scale atmospheric processes within the Aeolus observation sampling volume and their impact on the quality of the retrieved HLOS wind profiles. The database extends over four months covering all seasons. This allows a statistical evaluation of the mission components under investigation. The database is currently used for the development of the Aeolus wind processing, the definition of wind calibration strategies and the optimization of the Aeolus sampling strategy.

1 Introduction

A general methodology to test potential future (spaceborne remote-sensing) instruments for their ability of observing atmospheric quantities is by simulation studies that include the instrument hardware, a description of the atmospheric state and data processing algorithms to retrieve the objective quantity from the simulated measurement data. A prerequisite for the realistic simulation of all aspects of data processing is a realistic atmospheric description that includes atmospheric processes at all scales as observed by the instrument. The definition of the atmospheric database discussed in this paper is related to ESA's atmospheric dynamics mission, Aeolus, that aims to measure HLOS wind profiles from space with a Doppler wind lidar (DWL) from atmosphere scattering particles and molecules that move with the ambient wind (Stoffelen et al., 2005). Aeolus is a sun-synchronous dawn-dusk polar orbiting satellite with a scheduled lifetime of three years thus yielding global wind coverage for all seasons. The Aeolus DWL has a single fixed beam operated perpendicular to the satellite track direction and pointing at a 35 degrees off-nadir angle towards the atmosphere. The return signal from the atmosphere is divided in sequential time intervals that determine the vertical (range gate) resolution of the retrieved wind profile. The instrument thus measures a profile of single wind components along the laser beam line-of-sight (LOS) rather than the full wind vector. In the original instrument design the laser was operated in so-called burst mode meaning that the laser is turned on and off over a 28 seconds cycle in which the laser is active during 10 s and then turned off for 18 s to yield good-quality measurements along a 50-km track and no data in a subsequent 150-km gap. A single wind profile is then obtained every 200 km from the 50-km along track sampled data. Recently, it was decided to change the laser operation to so-called continuous mode, i.e., a continuous pulsed laser operated at about half the frequency as proposed for burst mode. The Aeolus DWL is a high spectral resolution lidar (HSRL) that operates at 355 nm laser wavelength (Flamant et al., 2008) to enable wind retrieval from both atmospheric particles (aerosols and cloud droplets and



Correspondence to: G. J. Marseille
(gert-jan.marseille@knmi.nl)

ice crystals) and molecules (Tan et al., 2008). In principle two wind profile solutions are obtained; one from the molecular signal on the so-called Rayleigh detection channel and one from the particle signal on the so-called Mie channel (Marseille et al., 2003). In burst mode the laser is operated at 100 Hz pulse repetition frequency yielding a fundamental horizontal resolution of about 70 m. The on-board accumulation for both channels is not fixed at the time of writing but will typically be in the order of 14–50 pulses resulting in a basic horizontal resolution of 1 to 3.5 km, the so-called measurement length. One Aeolus observation is thus obtained from 14–50 measurements that cover the 50 km observation length. This oversampling in the along-track direction offers the possibility to classify atmospheric segments, i.e., either loaded with particles or particle-free, before accumulation to observation level. The 3.5 km measurement length appears a good compromise between sufficient signal-to-noise (SNR) for classification and maximum oversampling, e.g., for the identification of along-track wind gradients.

The vertical resolution is fixed to 24 bins for both the Mie and Rayleigh channels with minimum and maximum bin sizes of 250 m and 2000 m respectively. Intermediate bin sizes must be multiples of 250 m. The maximum bin altitude is 32 km which is mainly driven by SNR considerations to yield Rayleigh channel winds that meet the mission LOS wind quality requirement. The distribution of these bins may be changed with a maximum of on average 8 times per orbit. The switching moments can not be decided instantaneously, but are preprogrammed and uploaded to the satellite every week. This provides the opportunity of optimizing (targeting) the sampling for different climate regions. The definition of optimal sampling strategies thus requires climatological knowledge on the distribution of atmospheric particles and their correlation with the atmospheric dynamics, more in particular the occurrence (geographical location and intensity) of heterogeneous scenes with large optical and dynamical variability within the Aeolus sampling volume. Such scenes are most challenging for Aeolus to retrieve representative winds because they are prone to biases (height assignment errors) that could be detrimental when used in numerical weather prediction (NWP) models. An adequate sampling may reduce such biases for instance by increasing the Mie channel resolution in areas of relative large variability of particle density.

From the considerations above the following requirements apply for the atmospheric database:

- global coverage and covering all seasons,
- inclusion of the following atmospheric meteorological parameters: zonal, meridional and vertical wind, temperature and optionally specific humidity all as a function of height above mean sea level and pressure,
- inclusion of the following optical parameters: molecular and particle backscatter and extinction at 355 nm

wavelength all as a function of height above mean sea level and pressure,

- the parameter profiles must extend from the earth surface up to an altitude of at least 32 km,
- the database horizontal resolution should be in the order of a few km and less than a few hundred meters in the vertical to simulate small-scale processes within the Aeolus sampling volume of typically 50 km along track and 1 km in the vertical.

An atmospheric database fulfilling these requirements enables the testing of the Aeolus ground processing and to generate global statistics of the occurrence of heterogeneous atmospheric scenes as a function of season, climate region and altitude. These statistics serve as input to define strategies for an optimal positioning of the bins for the Mie and Rayleigh channels. In addition to the primary geophysical parameters defined above, ground albedo and surface elevation are included enabling the assessment of opportunities for instrument zero-wind calibration at the surface.

As such the database would provide valuable input to assess the quality and coverage of winds obtained from the Mie and Rayleigh channel in heterogeneous atmospheric conditions. For instance, increasing the resolution of the Mie channel (to 250 or 500 m) in regions of large wind shear and cloud/aerosol variability will reduce height assignment errors for Mie channel winds and improve the quality of winds from the Rayleigh channel through a more accurate cross-talk correction, i.e., to correct for reflections from the Mie channel's Fizeau interferometer that enter the Rayleigh channel (Tan et al., 2008), (Dabas et al., 2008). Statistics of the occurrence and location of such events is obtained from the database and used as input for optimizing the vertical sampling strategy. Aeolus simulation tools like LIPAS (Marseille et al., 2003) can simulate the impact of various sampling scenarios on the quantity and quality of retrieved winds from the database input. Another aspect is improving the characterization of the stratospheric flow with Aeolus, i.e., either through an improved tropospheric flow and upward propagating waves into the stratosphere or through directly sampling the stratosphere. This aspect has recently been addressed in an operational NWP context through experiments at ECMWF for various sampling scenarios with a realistic distribution of simulated Aeolus winds, based on the database input as described in the remainder of this paper.

Existing databases or observing systems that measure the combined dynamical and optical parameters with the specified requirements are not available. Alternatively the database may be constructed by combining data from different sources. Available data sets that may serve as input for the atmospheric database include high vertical-resolution (30–60 m) radiosonde data (Houchi et al., 2010) for the meteorological part of the database, but these have limited spatial coverage in particular in the Tropics and Southern

Hemisphere. Radiosonde data also do not provide explicit information on the atmosphere optical parameters. Cloud resolving models, e.g. (Blossey et al., 2007) fulfill the high-resolution requirements but are limited in time and generally limited to the tropics and thus cannot fulfill the global coverage requirement. The EARLINET network (Mattis et al., 2004) of ground-based lidars over Europe provides high resolution data of the atmosphere optical properties over several years, but is limited to Europe and more in particular to the locations of the lidar instruments and lacks high-resolution collocated wind and temperature observations. Many field campaigns over the last decades have resulted in combined high-resolution lidar and wind data, the latter from occasional radiosonde launches. Although valuable for the study of typical phenomena, field campaign data are limited to the observation site and moreover radiosonde drift from the launch location hinders exact collocation of atmospheric optics and dynamics for ground campaigns. Advances in lidar technology enable ship and aircraft campaigns to measure typical phenomena such as boundary layers winds. An aircraft campaign issued by the Deutschen Zentrums für Luft- und Raumfahrt (DLR) near Iceland and Greenland yield combined wind and aerosol/cloud measurements from a 2 μm lidar, but limited to the field campaign period in November 2003 (Weissmann et al., 2007).

The Lidar-In-space Technology Experiment (LITE) in September 1994 was the first experiment to yield a dataset of lidar backscatter that nicely revealed the vertical structures of aerosols and clouds at high resolution and with a large coverage (Winker et al., 1996). LITE was however limited to a 2-week period in September and its coverage limited to ± 52 degrees latitude, thus missing the polar regions and large parts of the mid-latitude (storm-track) regions. The successor of LITE the Geoscience Laser Altimeter System (GLAS) carried on the Ice, Cloud and land Elevation Satellite (ICESat), launched on 13 January 2003, does provide a global coverage of atmospheric aerosol and cloud data extending from the surface up to 41 km with a 172 m horizontal and 77 m vertical resolution (Spinhirne et al., 2005). Instrument degradation however limits the dataset of high-quality measurements to the period 20 February until 18 November 2003.

With the advent of the successful Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) mission it was decided to use these data for the optical part of the atmospheric database. CALIPSO is part of the Aqua satellite constellation (or A-train) and was launched in April 2006 (Winker et al., 2007). It is operational for almost 4 years thus covering all seasons. CALIPSO combines an active lidar instrument, the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), with passive infrared (IIR) and visible imagers (WFC) to probe the vertical structure and properties of thin cloud and aerosol over the globe at a nominal horizontal and vertical resolution of 333 m and 30 m respectively. CALIPSO is positioned in a sun-synchronous orbit at

a nominal altitude of 705 km and inclination of 98.2 degrees thus yielding atmosphere optical properties with global coverage. The level-2 cloud and aerosol products were not available at the time of the database development. Moreover, the level-2 horizontal resolutions of 5 km (cloud) and predominant 80 km (aerosol) (Omar et al., 2009) are too coarse to simulate small-scale atmospheric processes (heterogeneity) within the Aeolus measurement volume. Therefore we developed an algorithm to retrieve particle (aerosol and cloud) backscatter and extinction at 3.5 km horizontal and 125 m vertical resolution from the high resolution CALIPSO level-1 calibrated attenuated backscatter product. This is further discussed in section 2.

The ECMWF global model was selected to complete the atmospheric database with meteorological parameters. These include the three wind components, temperature and humidity. To generate a database of combined CALIPSO optical and ECMWF meteorological parameters the latter were interpolated to the CALIPSO orbit track. This is further discussed in section 3. A general limitation of NWP models is the lack of spatial variability as compared to the real atmosphere, i.e., they are a smooth representation of the real atmosphere and the effective resolution is generally substantially smaller than the model grid size (Skamarock, 2004), (Frehlich et al., 2008). We come back to this issue in section 4.1 where we compare model and high-resolution radiosonde winds and present a method to add variability to the model winds.

2 CALIPSO atmosphere optics

The CALIOP instrument on-board the CALIPSO satellite is a two-wavelength polarization-sensitive lidar that provides high-resolution vertical profiles of aerosol and clouds. CALIOP utilizes three receiver channels: one measuring the 1064 nm backscatter intensity and two channels measuring orthogonally polarized components of the 532 nm backscattered signal. CALIPSO travels 14.55 orbits per day with a 24.7 degrees longitudinal separation. The orbit repeat cycle is 16 days. The nominal off-nadir angle of CALIOP is 0.3 degrees. The lidars are operated at 20 Hz yielding a nominal horizontal resolution of about 333 m. The fundamental vertical sampling resolution of the lidars is 30 meters, however on-board averaging reduces the horizontal/vertical resolution at high altitudes to 1000 m/60 m between 8.2 and 20.2 km, 1667 m/180 m between 20.2 and 30.1 km and 5000 m/300 m between 30.1 and 40 km. The CALIPSO level-2 cloud and aerosol products were not available at the time of the database development. In addition their horizontal resolution is too coarse to simulate small-scale atmospheric processes at Aeolus measurement resolution of typically 1–3.5 km. Therefore, we used the high-resolution CALIPSO level-1 532 nm total calibrated attenuated backscatter (i.e., the sum of the 532 nm parallel

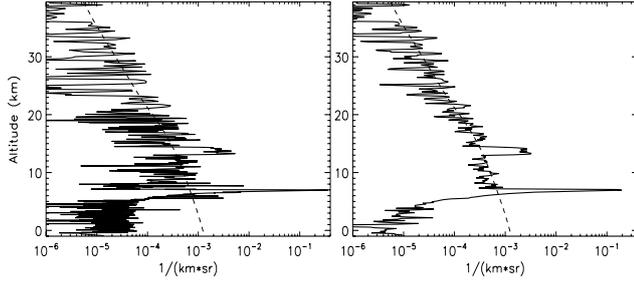


Fig. 1. CALIPSO raw calibrated level-1 532 nm attenuated backscatter profile (left) and the corresponding database profile after averaging the raw data to 3.5 km horizontal and 125 m vertical resolution (right) for a tropical scene showing an optically thin cirrus cloud layer at 14 km and an optically thick convective cloud between 6 and 8 km. The dashed reference signal corresponds to a synthetic particle-free atmosphere.

and perpendicular return signals). Before insertion into the database and to be used for Aeolus studies we retrieve molecular and particle backscatter at 532 nm followed by conversion of these quantities to 355 nm, i.e., the Aeolus DWL laser wavelength. A lidar inversion algorithm has been developed for this purpose which is the subject of the next section.

2.1 CALIPSO lidar inversion algorithm

The objective of the lidar inversion algorithm presented in this section is not to generate a level-2 aerosol and cloud product, but to retrieve the atmosphere particle properties (backscatter and extinction) from the CALIPSO level-1 attenuated backscatter product. The result is needed and sufficient for the database as a data source to simulate the distribution of particles and their optical properties inside the Aeolus sampling volume and their correlation with the atmospheric dynamics. A further characterization of cloud and aerosol types, their particle size distribution, etc is of less relevance for simulating Aeolus winds. Although the discrimination between cloud and aerosol is not essential for the database contents, the detection of particle layers from the attenuated backscatter data is an important component of the algorithm. This is discussed in the next section.

As a first step the resolution of the raw level-1 CALIPSO 532 nm total attenuated backscatter is reduced by averaging horizontally and vertically to yield a dataset at 3.5 km horizontal and 125 m vertical resolution, i.e., the resolution of the database, see Fig. 1. These values are a compromise between reducing the measurement noise and maintaining the high resolution needed to simulate Aeolus sub-volume processes. When mentioning signal in the remainder of this paper we mean the signal at database resolution unless otherwise stated.

We further elaborate on the noise characteristics in the raw and averaged signals that are needed later for the detection of particle layers as part of the lidar inversion algorithm. The

signal noise is a superposition of instrument noise and photon noise that is related to the signal strength. The signal noise mean and standard deviation is obtained from the difference of the raw signal and a reference noise-free signal representing a synthetic particle-free atmosphere. Details of the latter are given in the next section. The left panel of Fig. 1 shows an example of the raw and reference signals. The atmospheric region above 30 km is generally free of particles. The noise standard deviation in the raw signal, σ_n , above 30 km is obtained from a large sample of data between 30 and 34 km and equal to $7.07 \times 10^{-8} \text{ (m}^{-1} \text{ sr}^{-1})$ for the periods considered in this study and discussed in section 2.1.4. This value is not constant throughout the profile because of the altitude-dependent on-board averaging and the non-constant photon noise level as is clear from the raw signal in the left panel of Fig. 1 that shows an increasing noise level when going downwards in the profile (note the logarithmic scaling of the x -axis). The $3500 \times 125 \text{ m}$ volume average is larger for the averaged signal than for the raw signal at all altitude regions except above 30 km. As a result, the noise standard deviation of the averaged signal is reduced relative to the raw signal at almost all altitude regions, yielding noise levels throughout the profile that are proportional to the signal strength only, see the right panel of Fig. 1 (and note the logarithmic scaling of the x -axis).

The resulting profiles after averaging are processed individually without using information from neighboring profiles. Attenuated backscatter, denoted β' ($\text{m}^{-1} \text{ sr}^{-1}$), at height z above mean sea level is related to backscatter, β ($\text{m}^{-1} \text{ sr}^{-1}$), one-way transmission, τ , and extinction, α (m^{-1}), through

$$\beta'(z) = [\beta_m(z) + \beta_p(z)][\tau_m(z)\tau_p(z)\tau_{O_3}(z)]^2 \quad (1)$$

$$\tau_k(z) = e^{-\int_z^{z_{\text{sat}}} \alpha_k(z') dz' / \cos(\hat{\phi})} \text{ with } k = \{p, m, O_3\} \quad (2)$$

with the indices p, m, O_3 denoting particle, molecule and ozone respectively, where particles can be either aerosol or cloud, and z_{sat} the satellite altitude. In agreement with the vertical coordinate used in the CALIPSO attenuated backscatter data set we prefer for the vertical coordinate the altitude above mean sea level, z , at the location where the laser beam intersects the earth surface, rather than the range distance from the satellite to the atmosphere, r . These are simply related through $r = z_{\text{sat}} - z$ when ignoring the laser off-nadir pointing angle of the close to zero 0.3 degrees for CALIOP, i.e., an underestimation of the range from the satellite to the earth surface by $1.5 \times 10^{-3}\%$ or 10.6 m for a 705 km satellite altitude. Also, the cosine of the local incidence angle at the earth surface is very close to 1 and therefore not applied in the transmission equation Eq. (2).

Molecular scattering falls within the Rayleigh regime. Molecular backscatter and extinction at a given wavelength, λ , are obtained from Rayleigh scattering theory and atmospheric pressure and temperature. Extinction through scattering is a function of altitude z above mean sea level and

obtained from the product of the molecular number density, $n(z)$ (m^{-3}), and the molecular absorption cross-section, $\sigma_m(\lambda, z)$ (m^2). The latter is a function of laser wavelength, a depolarization factor of 1.061 and a refractive index, e.g. (Penndorf, 1955), (Bodhaine et al., 1999). A fixed value for σ_m of $5.2262419 \times 10^{-31}$ (m^2) at 532 nm was used in the retrieval algorithm. The molecular number density is obtained from knowledge of the atmospheric temperature, T (K), and pressure, p (Pa) through the ideal gas law, $p(z) = n(z)kT(z)$ with $k = 1.381 \times 10^{-23}$ (JK^{-1}) the Boltzmann constant. Pressure, temperature and molecular number density profiles are available from the CALIPSO level-1 dataset. Molecular backscatter, β_m , and extinction, α_m , are related through a wavelength independent constant factor of $3/8\pi$. In formulas:

$$\alpha_m(\lambda, z) = \sigma_m(\lambda, z) \frac{p(z)}{kT(z)}; \beta_m(\lambda, z) = \frac{3}{8\pi} \alpha_m(\lambda, z) \quad (3)$$

It is noted that ozone extinction can not be ignored at 532 nm wavelength. In a similar way as for molecules above, ozone extinction is obtained from the product of the ozone number density (m^{-3}) and the ozone absorption cross-section (m^2). The former is a function of altitude and part of the CALIPSO level-1 data set. A fixed value for the ozone absorption cross-section of 2.728461×10^{-25} (m^2) was used as obtained from the MODTRAN 3.7 database. Figure 2 shows typical ozone extinction and transmission profiles.

The ozone absorption cross-section at 355 nm is roughly two orders of magnitude smaller than at 532 nm, meaning that ozone extinction at 355 nm, i.e., the Aeolus laser wavelength, is negligible.

To solve Eq. (1) it is discretized in equidistant bins of size $\Delta z = 125$ m. Quantities inside the bins are assumed constant. For bin number i , centered at altitude z_i , this yields:

$$\beta'(i) = [\beta_m(i) + \beta_p(i)] T_m^2(i-1) T_p^2(i-1) \tau_m^2(i) \tau_p^2(i) \quad (4)$$

with bin i positioned below bin $i-1$, i.e., bin 1 is at the top of the atmosphere. $T_m(i-1)$ is the total molecular and ozone one-way transmission and $T_p(i-1)$ the total one-way particle transmission from the top of the atmosphere until the bottom of bin $i-1$ and $\tau_m(i)$ and $\tau_p(i)$ are the one-way molecular and particle transmission in the center of bin i . To discriminate between particle-rich and particle-free parts of the atmosphere the measured attenuated backscatter profile is related to an approximate attenuated molecular backscatter profile, β'_m , for a synthetic molecules-only atmosphere:

$$\beta'_m(i) = \beta_m(i) T_m^2(i-1) \tau_m^2(i) \quad (5)$$

This profile is well determined by the atmospheric temperature, pressure and ozone parameters. Defining the scattering ratio, ρ , as the ratio of total backscatter and molecular backscatter, $\rho(i) = [\beta_m(i) + \beta_p(i)] / \beta_m(i)$, we can write for the ratio, β'' , of Eqs. (4) and (5)

$$\beta''(i) = \beta'(i) / \beta'_m(i) = \rho(i) T_p^2(i-1) \tau_p^2(i) \quad (6)$$

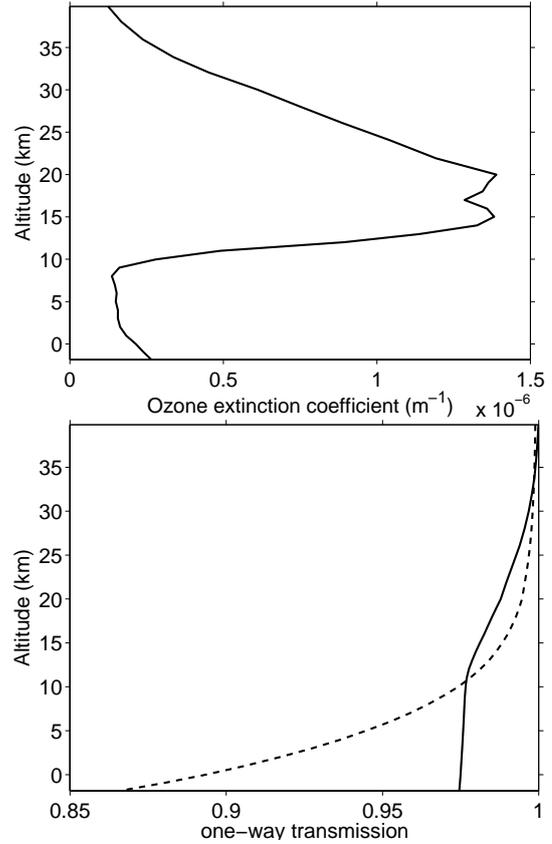


Fig. 2. Extinction (top) and one-way transmission (bottom) for ozone (solid line) and a reference molecular profile (dashed line) at 532 nm wavelength.

Eq. (6) is the basic equation for the detection of particle-rich layers in the measured profile as explained in the following subsection.

2.1.1 Particle layer optical properties

The second step of the retrieval algorithm, after averaging the raw data, is to detect particle-rich layers within the measured profile using Eq. (6) and to determine the layer optical properties. The discussion closely follows (Young, 1995). Assume a particle-rich layer with the layer top and bottom positioned in the bins denoted i_{top} and i_{bot} respectively. Above the layer top and below the layer bottom the atmosphere is assumed particle-free over a certain range, i.e., $\rho(i) = \tau_p(i) = 1$ for $i < i_{\text{top}}$ and $i > i_{\text{bot}}$. From Eq. (6) $\beta''(i)$ has a constant value of $T_p^2(i_{\text{top}} - 1)$ above the particle layer, i.e., the total transmission of particle layers aloft, and a constant value of $\bar{T}_p^2 T_p^2(i_{\text{top}} - 1)$ below the particle layer with \bar{T}_p^2 the two-way transmission of the particle layer under investigation, i.e., the ratio of β'' below and above the particle layer. Inside the particle layer β'' is related to the distribution of particles that is generally non-constant. The layer top and bottom can thus be found from the gradient of the β'' profile

that is close to zero outside and non-zero inside particle layers.

The detection of particle layers is done successively starting from the top of the atmosphere. The first region in the profile where β'' exceeds a pre-specified threshold value is marked as a particle region. The layer top is found as the first bin of a number of consecutive bins on top of the marked region where β'' is constant and equal to $T_p^2(i_{\text{top}} - 1)$. Going from the layer top through the layer, the layer bottom is found where β'' becomes constant again and with a value less than $T_p^2(i_{\text{top}} - 1)$.

The magnitude of the threshold value specifies the type of particle layer that is detected. The current implementation of the algorithm focuses on the detection of cloud layers. Ice clouds in the upper troposphere are generally optically thin with signal strength of the same order of magnitude as aerosol layers in the lower troposphere. To discriminate between these layer types the atmosphere is split in two parts: above 8 km where most of the ice clouds reside and below 8 km where most aerosol layers are found.

Threshold values for automatic detection of the particle layer boundaries need to take into account the signal noise. This was discussed in the previous section. For each profile the signal noise standard deviation, σ_n , is obtained from its difference with the reference signal in the region between 30 and 34 km. For the detection of particle layers the following thresholding profile, ρ_t , applied to β'' is defined

$$\rho_t(i) = 1 + f(i)/\beta_m(i); f(i) \text{ in } (\text{m}^{-1} \text{sr}^{-1}) \text{ with} \quad (7)$$

$$f(i) = 6\sigma_n (z_i > 8 \text{ km}) \text{ and } 5 \times 10^{-6} (z_i < 8 \text{ km}) \quad (8)$$

Bins for which β'' exceeds the threshold value are in the cloud layer. Eq. (8) translates in the upper troposphere to a signal exceeding the reference signal plus 6 times the noise standard deviation σ_n above 30 km. This choice is motivated by the assumption that photon noise is the dominating contributor to the noise in the signal. From the reference signal in Fig. 1 it is seen that the signal strength increases by a factor of about 16 between 30 km and the upper tropospheric region between 11 and 17 km, where most tropical cirrus clouds reside. Thus, the noise standard deviation in the upper troposphere is about $4\sigma_n$. The value of 6 in Eq. (8) was found as a compromise of maximum probability of detecting (thin) tropical cirrus clouds and to minimize the false alarm rate from spurious (noise) peaks. Isolated bins exceeding the threshold value are assumed noise outliers and not treated as a particle layer. The value of 5×10^{-6} ($\text{m}^{-1} \text{sr}^{-1}$) for the lower troposphere enables the detection of (generally optically thick water) clouds while ignoring aerosol layers. More advanced thresholding strategies may be developed for the detection and characterization of isolated lofted aerosol layers. We stress again that the objective is not to generate a level-2 aerosol and cloud product, but to retrieve particle backscatter and extinction for adoption in the Aeolus database. The retrieval of aerosol optical properties is further discussed in the next section.

Having determined the particle layer boundaries and transmission, the next step is to compute the backscatter and extinction inside the layer. Hereto, the lidar ratio parameter $S_p(i)$ (sr) for particles is introduced that is defined as the ratio of particle extinction and backscatter, $S_p(i) = \alpha_p(i)/\beta_p(i)$. The lidar ratio does not depend on the particle density but on particle properties such as the size distribution and shape. These properties are assumed homogeneous inside the layer and thus the layer lidar ratio is assumed constant. Rearranging terms in Eq. (6) and substituting for $\tau_p(i) = \exp(-\alpha_p(i)\Delta z/2) = \exp(-S_p\beta_p(i)\Delta z/2)$ (the factor 1/2 referring to the one-way transmission at the bin center rather than the bin bottom) yields the so-called retrieval function for bin i :

$$[\beta_m(i) + \beta_p(i)]e^{-\eta S_p\beta_p(i)\Delta z} - c(i) = 0 \quad (9)$$

with known $c(i) = \beta_m(i)\beta''(i)/T_p^2(i - 1)$.

For the high altitude CALIPSO orbit of about 705 km multiple scattering effects cannot be ignored (Winker, 2003). Even for the very narrow field of view of $130 \mu\text{rad}$ for CALIPSO, the large distance between the lidar and the troposphere generates a relatively large spot size at the target altitudes. To account for this effect the multiple scattering factor η is introduced in Eq. (9). The value of η can vary between 1 and 0.5 thus η has the effect of reducing the effective optical depth such that these photons are treated as if they had not been scattered at all (Hogan, 2008). The exact magnitude of the multiple-scattering effect depends on the number density, size distribution and shape of the cloud particles, and also on the lidar field of view and the distance of the scattering volume from the lidar (Young, 1995). Typical values used in the retrieval algorithm are detailed below.

The retrieval function has two unknowns, $\beta_p(i)$ and the effective lidar ratio $S^* = \eta S_p$. For a given value of S^* , $\beta_p(i)$ can be solved in principle although an analytical solution is not available and the retrieval function is therefore solved iteratively. Properties of the retrieval function are discussed in the Appendix. To solve Eq. (9) for the particle layer an exhaustive search is applied for integer values S^* in the range 1 to 120. For each value the retrieval function is solved for $\beta_p(i)$ inside the layer $i_{\text{bot}} \leq i \leq i_{\text{top}}$. Next the corresponding layer transmission, $\hat{T}_p(S^*) = \prod_{i=i_{\text{bot}}}^{i_{\text{top}}} \exp[-S^*\beta_p(i, S^*)\Delta z]$ is computed. The solution for the particle layer is found for S^* with corresponding $\beta_p(i, S^*)$ for which $\hat{T}_p(S^*)$ is closest to the predetermined particle layer transmission \bar{T}_p . This procedure of layer detection and retrieval of its particle optical properties is repeated until the PBL is reached to enable the detection of multiple particle layers in the measured profile.

The procedure described above is applied only to isolated particle layers above the PBL for which the particle-free range above and below the layer is sufficiently large to determine the layer transmission with sufficient accuracy. Despite the averaging of the raw data in the first step of the

algorithm the remaining noise in the resulting profiles is still substantial, see the right panel of Fig. 1. It was found that a minimum 1 km particle-free range is needed, but more accurate transmission estimates are obtained for larger ranges. For non-isolated particle layers the algorithm can not compute the layer transmission from the measured signal. Non-isolated particle layers are found in atmospheric scenes with (i) layers inside the PBL that is generally rich of aerosols, (ii) layers that extend to the surface (e.g., through multiple scattering in optically thick PBL clouds), (iii) too small SNR below the layer (below optically thick clouds); in the algorithm a threshold of 10% of the reference signal is used, or iv) too small particle-free ranges either above or below the layer, i.e., less than 1 km (nearby underlying particle layers or when the layer bottom is close to the surface). For these scenes that allow no calculation of the particle layer transmission the algorithm uses default lidar ratio values for clouds. It is then assumed that the layer is a cloud layer which is generally true for the selected threshold values in Eqs. (7, see also 8). We adopted the lidar ratio values as described in the CALIOP algorithm theoretical basis documents (ATBD), i.e., 18 sr for water clouds, Pinnick et al. (1983), 24 sr for ice clouds and 21 sr for mixed water/ice clouds. Discrimination between cloud type is based on atmospheric temperature, i.e., water clouds above 0 degrees Celsius, mixed water/ice clouds between -20 and 0 degrees Celsius and ice clouds below -20 degrees Celsius. For ice clouds η is well approximated by a value of 0.7 (Winker, 2003). For water clouds exact knowledge of η is of less relevance because the laser beam will generally only penetrate the cloud top because of its large optical thickness. The SNR below isolated water clouds in the free troposphere is therefore generally too small for cloud transmission calculation. Also, multiple scattering of water clouds in the PBL gives substantial signal below the cloud bottom until the earth surface. For the algorithm it is sufficient to detect water clouds and to set the effective lidar ratio to a value that avoids divergence of the algorithm, see below. The exact backscatter and extinction in the lower part of the cloud and below the cloud is of less relevance for Aeolus because the lidar beam can not penetrate deeply into water clouds.

The retrieval algorithm is prone to divergence if the selected value for the lidar ratio is too large. In that case particle transmission is underestimated (or extinction overestimated) initially yielding excessive large backscatter values further down into the profile (for increasing i) and ultimately no solution for the retrieval function. This becomes clear from Eq. (9). If particle transmission is underestimated then $T_p(i)$ will eventually converge to zero and $c(i)$ to infinite, with no valid solution for β_p . This is called divergence of the retrieval function solution. The algorithm checks for divergence by comparing the ratio of the retrieved particle backscatter and the signal attenuated backscatter against a pre-specified threshold value. A threshold value of 5 is used in the algorithm for cloud layers that is not hampering large

backscatter solutions for optically thin ice clouds but is conservative for optically thick water clouds with small signal values at the cloud bottom. The selected default lidar ratio values above are tested on divergence inside the particle layer. If divergence occurs then the lidar ratio value is decreased by 20% until divergence is avoided. The lidar ratio is thus merely a tuning parameter to ensure robustness of the algorithm. The SNR below optically thick clouds such as cumulus or stratus cloud is generally too small to yield a valid solution for the cloud lidar ratio, cloud transmission and resulting cloud backscatter. This potential limitation is not detrimental for the assessment of the 355 nm Aeolus laser because the laser beam will not penetrate the lower part of the cloud and the region below the cloud at this wavelength.

2.1.2 Aerosol optical properties

After having detected the cloud layers and the determination of their optical properties, the next step is to retrieve the aerosol backscatter and extinction in the remaining regions of the atmosphere. Hereto Eq. (9) is solved for bins that were not detected in the previous step and with indices p now denoting aerosols. For detected cloud bins in the previous step the computed transmission values are used in the computation of $c(i)$. Because the algorithm can not discriminate between different types of aerosol a default value for the effective aerosol lidar ratio was used that is typical for clean marine and continental conditions with lidar ratio values ranging from 20–35 (sr). The value of 35 sr, that is typically used for clean background aerosol was selected as default value (Omar et al., 2009). The aerosol solution is sensitive to signal noise. The upper part of the atmosphere is almost free of aerosol and the solution is dominated by noise rather than aerosol presence with a possible exception in the Tropics where deep convection may transport sea salt and smoke (forest fires) to high altitudes up to the tropopause while stimulating cloud formation. For low signal with zero mean noise the retrieval results in negative aerosol backscatter solutions for about half of the retrievals. These are set to zero, thereby introducing a negligible bias. The positive solutions are dominated by noise rather than aerosol presence. To avoid overestimation of extinction in the upper part of the atmosphere, due to noise rather than aerosol solutions, the aerosol lidar ratio is set to 0 above 8 km where the density of background aerosol is generally negligible. As for the particle layers, the algorithm is checked for divergence but with a different criterion, namely the particle backscatter solution should be smaller than 10^{-3} ($\text{m}^{-1} \text{sr}^{-1}$). In case of divergence the aerosol effective lidar ratio is reduced by 20% until divergence is avoided. Such events are rare and mainly related to aerosol scenes with overlying clouds and an imperfect cloud transmission estimation.

The transmission of lofted aerosol layers with larger effective lidar ratio values than assumed by the algorithm is overestimated. This is the case for smoke (from biomass

burning), dust (e.g. from desert dust outbreaks), polluted dust (dust mixed with biomass burning smoke) and polluted continental (urban industry) aerosol with average lidar ratio values of 70, 40, 65 and 70 (sr) respectively (Omar et al., 2009). The retrieval function always finds a solution for β_p . However, when the aerosol lidar ratio is underestimated the backscatter inside the layer is underestimated and the layer transmission is overestimated. The value for c below the layer is then too small (or the computed signal below this layer is too large) yielding an underestimate and possibly negative aerosol backscatter below the aerosol layer. The algorithm will then compensate by an aerosol-free atmospheric layer below the aerosol layer until the calculated total atmospheric particle transmission (T_p) agrees with the signal. Vice versa, when the aerosol lidar ratio is overestimated, e.g. for loft clean marine layers with an on average 20 (sr) lidar ratio (Omar et al., 2009), the backscatter inside the layer is overestimated and the layer transmission is underestimated with too large values for c below the aerosol layer. The algorithm will then add aerosols below the layer bottom, i.e., the algorithm compensates by increasing the layer thickness. This is done until the calculated total atmospheric particle transmission (T_p) agrees with the signal. But the algorithm might also diverge in which case the lidar ratio is decreased. In general, the magnitude of the retrieved aerosol backscatter is sensitive to the lidar ratio value used in the retrieval algorithm. A smaller/larger value than the inherently unknown true value of the aerosol layer results in an under/overestimate of the retrieved aerosol backscatter that may be in the order of 100% for large discrepancies. The validation in section 2.1.4 will show that such large discrepancies are rare.

2.1.3 Wavelength conversion

For adoption in the Aeolus atmospheric database the retrieved molecular and particle backscatter and extinction at 532 nm needs conversion to the Aeolus laser wavelength of 355 nm. Molecular backscatter and extinction at 355 nm is obtained from Eq. (3) similar as for 532 nm in the lidar inversion algorithm but with a different value for the absorption cross-section. For wavelength conversion of particle backscatter and extinction the Ångström exponent power law that expresses the spectral dependence of aerosol optical thickness on the wavelength of incident light is used:

$$\alpha(\lambda_2) = \alpha(\lambda_1) \left(\frac{\lambda_1}{\lambda_2} \right)^{\mathring{A}} \quad (10)$$

with \mathring{A} the Ångström exponent and here $\lambda_1 = 532$ nm and $\lambda_2 = 355$ nm. The Ångström exponent is often used as a qualitative indicator of aerosol particle size, with values larger than 2 indicating small particles associated with, e.g., combustion by-products, and values less than 1 indicating large particles like sea salt and dust (Schuster et al., 2006). Mattis et al. (2004) determine the Ångström exponent as a

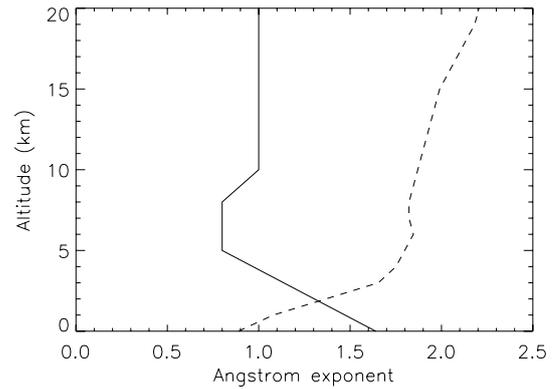


Fig. 3. Ångström exponent as derived from Mattis et al. (2004) (solid) and Marseille et al. (2003) (dashed).

function of altitude from a multiyear 2000–2003 dataset of a Raman lidar operated at 532 nm and 355 nm laser wavelength over Leipzig, Germany. From their Table 2 and Fig. 3 the black solid curve in Fig. 3 was extracted based on an average value of 1.4 in the PBL, decreasing linearly to a value of 0.8 between 5 and 8 km and then increasing to a value of 1 in the upper part of the troposphere. They note that the Ångström exponent shows a large variability of 0.5 and 0.8 standard deviation in the PBL and free-troposphere respectively, depending on observed aerosol type and relative humidity. In a comparative study of lidar aerosol measurements at different wavelengths Vaughan et al. (1998) parameterized the Ångström exponent through $\mathring{A}(\beta_0(z)) = -0.104 \times \ln(\beta_0(z)) - 0.62$ (see also Marseille et al. (2003)) with β_0 the aerosol backscatter at 10.6 μm as obtained from flight campaigns over the North and South Atlantic in 1989, (Vaughan et al., 1995). The corresponding dashed curve in Fig. 3 is opposite to the Leipzig curve that may be explained by different compositions of aerosols for these distinct regions. The retrieval algorithm discussed in the previous sections does not distinguish between different types of aerosols. We selected the Ångström exponent profile from the more recent dataset from Mattis et al. (2004) for the particle wavelength conversion.

From Eq. (10) the particle extinction at 355 nm, α_p^{355} , is obtained from the retrieved particle backscatter, β_p^{532} , and lidar ratio, S_p^{532} at 532 nm. The particle backscatter, β_p^{355} , at 355 nm then follows from a priori knowledge of the lidar ratio, S_p^{355} at 355 nm from

$$\alpha_p^{355}(z) = \beta_p^{532} S_p^{532} \left(\frac{532}{355} \right)^{\mathring{A}(z)} ; \beta_p^{355}(z) = \frac{\alpha_p^{355}(z)}{S_p^{355}} \quad (11)$$

The ratio of the lidar ratio at 532 nm and 355 nm is a function of particle size and shows a complex behaviour around unity, but may vary from -50% to +70%. In addition the relative humidity (RH) may change the size of hygroscopic particles substantially by up to 40% for $\text{RH} > 70\%$ (Flamant

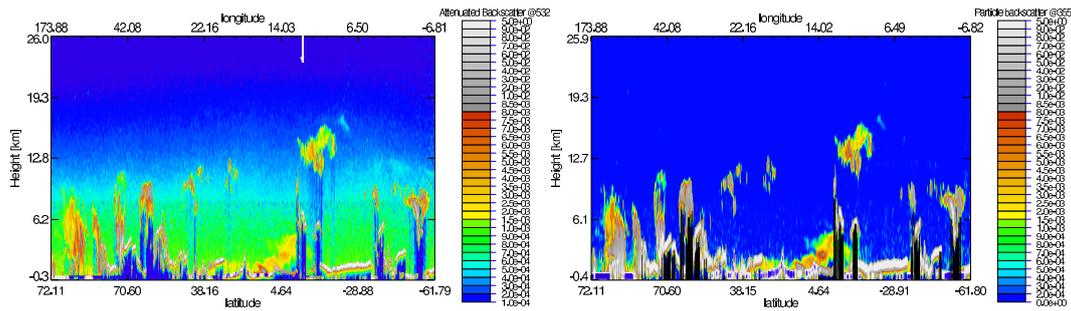


Fig. 4. Left, CALIPSO 532 nm total attenuated backscatter ($\text{km}^{-1} \text{sr}^{-1}$) measured on 1 January 2007 at 00:22–00:49 UTC. The white arrow at -4.32 degrees latitude, 12.12 degrees longitude points to the location of the raw data displayed in Fig. 1. Right, retrieved particle backscatter ($\text{km}^{-1} \text{sr}^{-1}$) from the inversion algorithm and converted to 355 nm wavelength. Black spots indicate regions where no particles could be retrieved by the algorithm.

et al., 2008-II). Omar et al. (2009) shows typical variations from unity of -56% (clean marine) to 133% (polluted continental) for 532 nm and 1064 nm wavelengths. Ackermann (1998) shows that the lidar ratio at 532 nm and 355 nm is similar, on average, for continental and maritime aerosol and only shows a large discrepancy for desert dust with a doubling of the value at 355 nm wavelength. Also Mattis et al. (2004) finds on average similar values for the lidar ratio at both wavelengths. Based on these results and noting that the lidar inversion algorithm does not discriminate between different types of aerosols the lidar ratio at 355 nm was taken identical to the value obtained at 532 nm.

2.1.4 Results

Four months of CALIPSO level-1 532 nm total attenuated backscatter (release version 2.01) were processed with the lidar inversion algorithm described in the previous sections. These months include January, April, August and October 2007, thus covering all seasons. CALIPSO data observed in daytime conditions were not used because of the additional solar background signal from objects within the sensing volume and illuminated by the sun like water clouds and earth surfaces with a large albedo. Only data observed at nighttime, with substantially larger SNR, were used to minimize the noise contamination in the retrieved particle backscatter profiles. In addition the CALIPSO nighttime atmospheric conditions are in better agreement with those observed by Aeolus at local dawn and dusk, see section 4. The total dataset included about 1800 (half) orbits and the inversion algorithm was operated automatically over the whole dataset without manual intervention needed.

Figure 4 shows a typical example of the raw CALIPSO attenuated backscatter and the retrieved particle backscatter at 355 nm for a half CALIPSO orbit. The first part of the orbit, on the left hand side of the figure, is in ascending node then crosses the North Pole at 82 degrees and then descending until 61 degrees in the Southern Hemisphere. The scene shows optically thin cirrus clouds at 73°N (descending node)

over Western Russia at 10 km altitude, between 48°N and 36°N over Eastern Europe and the Mediterranean at 10 km altitude, between 29°N and 22°N over the African continent at 12 km altitude, between 2°S and 18°S in the Tropics along the west coast of Southern Africa between 11 and 17 km altitude with an enhanced optical depth near 10°S , a tenuous cirrus cloud at 22°S at 17 km altitude and between 42°S and 52°S around 9 km altitude. In the polar region in the Northern part of Russia large cloud systems are found extending from the PBL to 9 kilometers altitude. Dense clouds are found between 5 and 10 km in the mid-latitude regions. Stratus clouds with cloud tops below 2 km are found over the South Atlantic between 12°S and 34°S . Aerosols are found up to 2 km altitude throughout the scene with more pronounced events over the African continent between 38°N and 18°N and a strong aerosol dust layer between 14°N until 8°S that reaches a maximum altitude of 5 km.

The retrieved particle backscatter in the right panel of Fig. 4 reveals the prominent features as observed in the left panel of Fig. 4. The signal below dense water clouds is generally too small and no particle backscatter is obtained from the algorithm in these areas. Spurious isolated features above the PBL are probably related to signal noise rather than atmospheric particles.

Figure 5 shows the clouds as detected by the algorithm and their estimated effective lidar ratio. The threshold profile, Eqs. (7,8), enables the detection of most of the cirrus clouds including part of the tenuous cloud at 22°S , 17 km altitude. The strong aerosol layer near the Equator is not identified as a cloud layer because the threshold value in the lower part of the atmosphere below 8 km was selected to detect only the strongest signals from dense clouds. The orange bar between the surface and 8 km corresponds to the imposed default 35 (sr) lidar ratio value for aerosol. At some locations this value has been reduced by the algorithm to avoid divergence, in particular below or above dense clouds. Because profiles are processed individually the lidar ratio values inside the clouds may differ. The small lidar ratio values for

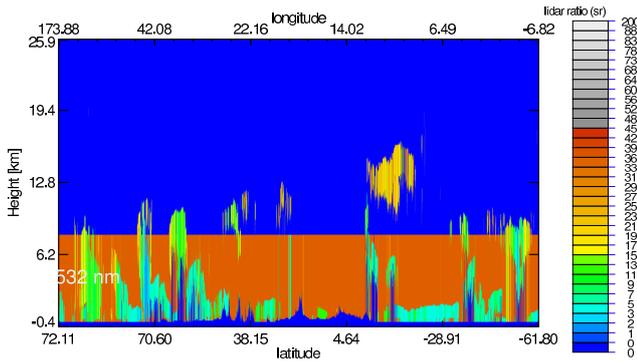


Fig. 5. Cloud and aerosol effective particle lidar ratio (sr) as obtained from the retrieval algorithm for the scene in the left panel of Fig. 4. Dark blue spots denote areas that are free of particles (upper troposphere, stratosphere and below the earth surface). The orange bar between the surface and 8 km corresponds to the imposed default 35 (sr) lidar ratio value for aerosols.

dense water clouds may not be realistic but is merely the result of forcing the solution within the boundaries imposed by the divergence criteria to avoid excessive large backscatter solutions in the lower part of the clouds.

Figure 6 shows statistics of the fraction of detected clouds and the retrieved lidar ratio values as a function of altitude and climate zone. This is the combined result from all seasons. The climate zones are defined as follows: polar regions 70° – 90° N/S, midlatitude regions 40° – 70° N/S, subtropical region 20° – 40° N/S and the tropics 20° N– 20° S. The small cloud fraction above the tropopause for the various climate zones is explained by spurious noise artefacts in the signal rather than real clouds, except for PSCs over the South Pole area in August. Most of the PSCs over the North Pole area are too tenuous to be captured by the cloud layer detection algorithm. Here, it is noted that the application of the particle layer detection scheme was limited to the lowest 20 km of the atmosphere and thus not well tuned to process PSCs at higher altitudes. These are treated as fully transparent aerosol by the algorithm. The mean effective lidar ratio is generally close to 17 (sr) at high altitudes, i.e., close to the product of the multiple scattering factor of 0.7 and the default 24 (sr) lidar ratio for ice clouds. The mean value decreases at lower altitudes compatible with the decreasing default lidar ratio values for mixed ice/water and water clouds.

We further elaborate on the tropical upper troposphere because the presence of cirrus clouds in combination with relatively large wind-shears as discussed in section 3 makes this region challenging for Aeolus for the retrieval of representative winds. Table 1 shows the main characteristics of the cirrus clouds as obtained from the retrieval algorithm for the winter (January 2007) period. Tropical cirrus clouds are defined as clouds located between 20° N and 20° S with a minimum cloud bottom of 10.5 km. A total of 207268 cirrus cloud layers covering a total of 4155360 125 m bins were

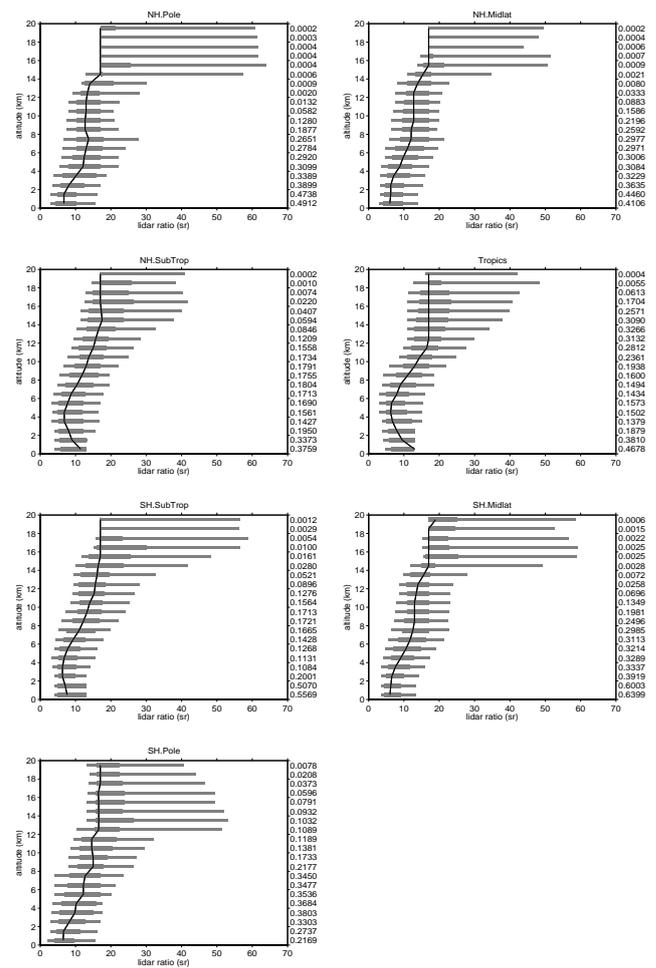


Fig. 6. Cloud statistics derived from 4 months of CALIPSO 532 nm total attenuated backscatter data. The bar lines denote effective cloud lidar ratio (sr). The black solid line denotes the median value, the thick gray bar the 25–75% percentile range and the small bars the 10–25% and 75–90% percentile range. The numbers on the right-hand side denote the fraction of cloud occurrence.

detected, where we note that multiple cloud layers may be found in a single profile. Tropical cirrus cloud temperature is generally between -40 and -80 degrees Celsius with a mean value of about -60 degrees. The median value of 18 (sr) for the lidar ratio in tropical cirrus cloud is in agreement with the value of 19 (sr) reported in Young (1995). The retrieval algorithm yields the parameters at 532 nm wavelength and converts to 355 nm as explained in section 2.1.3. In general, backscatter and extinction is larger at 355 nm than at 532 nm. The larger scattering ratio at 532 nm is explained by the substantially smaller molecular backscatter at this wavelength. The median cloud thickness is smaller than the difference of the median cloud top and cloud bottom, because multiple cloud layer can be found on top of each other. Tropical cirrus cloud is optically thin in general with a two-way transmission of more than 0.6 at 355 nm in 75% of the cases. This

tropical cirrus clouds		P10	P25	P50	P75	P90
log10 β_p at 532 nm	($\text{m}^{-1} \text{sr}^{-1}$)	-8.58	-6.54	-5.81	-5.26	-4.91
log10 α_p at 532 nm	(m^{-1})	-7.29	-5.26	-4.53	-4.06	-3.75
ρ at 532 nm		1.01	1.88	7.07	22.53	49.21
one-way transmission at 532 nm		0.65	0.84	0.94	0.97	0.99
log10 β_p at 355 nm	($\text{m}^{-1} \text{sr}^{-1}$)	-8.40	-6.36	-5.63	-5.09	-4.73
log10 α_p at 355 nm	(m^{-1})	-7.11	-5.08	-4.35	-3.89	-3.57
ρ at 355 nm		1.00	1.32	2.73	6.99	14.72
one-way transmission at 355 nm		0.53	0.78	0.91	0.96	0.99
effective lidar ratio at 532/355 nm	(sr)	13.4	16.8	18.0	29.0	50.0
cloud top	(km)	13.8	15.1	16.6	17.3	17.7
cloud bottom	(km)	11.1	11.8	13.2	14.7	15.8
cloud thickness	(km)	1.25	1.75	2.50	3.75	4.87

Table 1. Tropical cirrus cloud statistics for the winter (January 2007) period. Statistics are given by percentiles p_i with i denoting the percentage of data with values smaller than the table value. β_p and α_p denote ice cloud particle backscatter and extinction respectively, ρ is the scattering ratio.

mean that Aeolus is well capable to penetrate the tropical lower troposphere and yield good quality winds even below tropical cirrus clouds, i.e., a 40% signal loss yields an increase of the HLOS wind error standard deviation by a factor of $1/\sqrt{0.6} = 1.3$ below the cloud. The error of corresponding (Rayleigh channel) winds is still within the mission requirement.

Similar tropical cirrus cloud statistics have been determined for the other seasons. Overall the results are compatible with a slightly larger median ice particle (log10) backscatter of -5.47 , a smaller median lidar ratio of 17 (sr) and transmission of 0.90 in the summer. Median cloud top altitudes are lower for all other seasons with values of 15.9, 15.3 and 15.6 km in spring, summer and autumn respectively. Also the median cloud bottom altitudes are lower with values of 12.8, 12.3 and 12.4 in spring, summer and autumn respectively. The median cirrus cloud layer thickness is similar for all seasons.

Figure 7 shows the combined statistics of the retrieved aerosol (thus excluding clouds) backscatter at 355 nm for all seasons in 2007 as a function of climate zone. This includes more than 1 million profiles for all climate zones except for the polar regions with 0.64 (NH-Pole) and 0.85 (SH-Pole) million profiles. The red and black lines correspond to the CALIPSO aerosol percentiles with the red line denoting the median. For reference aerosol backscatter from two other datasets are included in the figure. The green curve (with the lowest particle backscatter values) is the reference model atmosphere (RMA) median profile as derived by (Vaughan et al., 1995) and used as a reference for many Aeolus related studies e.g. (Marseille et al., 2003), (Stoffelen et al., 2005). The backscatter database of Vaughan is based on flight campaign data with a lidar operated at $10.6\mu\text{m}$ in 1989 over the Atlantic near Iceland, the UK, Azores and Ascension Island. A scaling law was used to convert from $10.6\mu\text{m}$ to 355 nm laser wavelength (Marseille et al., 2003). This 1989 period

has been characterized as an extremely clean period. The statistics correspond to the background aerosol density, excluding clouds although the higher deciles show some cirrus. The purple curve with the highest particle backscatter is the retrieved median aerosol backscatter profile from LITE (Winker et al., 1996). The LITE period of September 1994 has been characterized as dirty, with an increased aerosol density in the upper troposphere and stratosphere due to the Pinatubo volcanic eruption in 1991. We note that the Vaughan and LITE reference profiles in Fig. 7 are identical in all climate zone panels, i.e., no effort was done to generate statistics for different climate zones for these datasets since they serve mainly as a reference for the CALIPSO results. The reference profiles were only added to those climate zones that were covered by the measurement campaigns.

The lower CALIPSO aerosol percentiles (10% and 25%) are generally outside the bounds of the figures, meaning that their values are either extremely small or zero for low aerosol regimes where the signal is dominated by noise as explained in section 2.1.2. Because of the large dataset the median aerosol backscatter profiles show an overall realistic exponential decrease of aerosol density with altitude. The decreasing numbers on the right hand side of the figure when getting closer to the surface are explained by missing data below dense clouds and/or orography. The median CALIPSO profile is generally between the "clean" RMA and "dirty" LITE median profiles throughout the atmosphere.

Figure 8 shows the median aerosol backscatter for the different seasons. No backscatter profiles from nighttime CALIPSO orbits are obtained for the Northern Hemisphere Polar area in August and the Southern Hemisphere Polar area in January, because these areas are lit by the sun in these periods. The CALIPSO aerosol product does not show a large seasonal dependence except for the Southern Hemisphere polar area in August with an increased aerosol loading above 10 km due to the presence of PSCs. CALIPSO aerosol is well

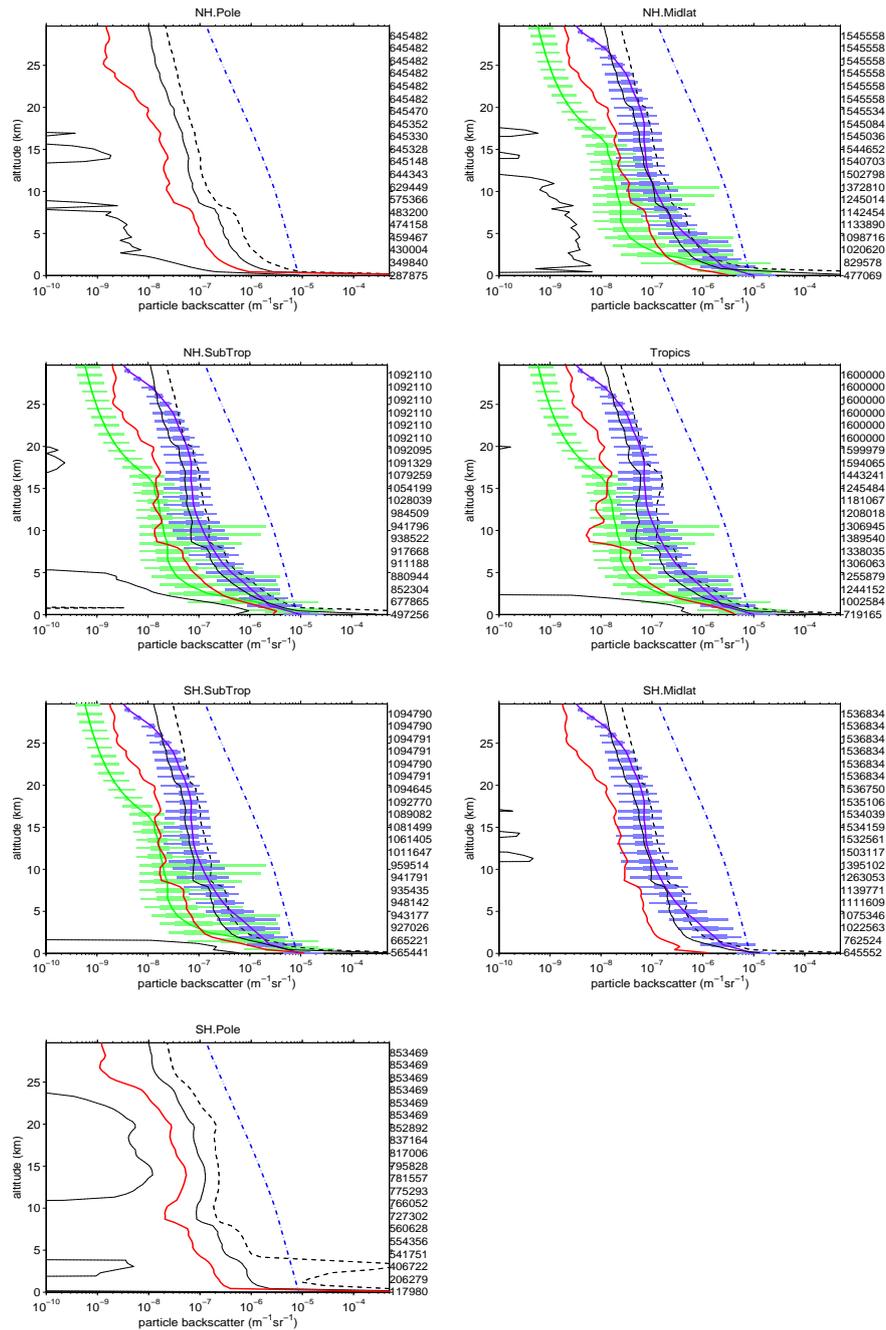


Fig. 7. Backscatter statistics at 355 nm wavelength as a function of climate zones for the whole month of January 2007. The red and black curves correspond to one month of retrieved backscatter data from CALIPSO; the red line denotes the median, the black solid lines denote the 25% and 75% percentiles, the black dashed line denotes the 90% percentile. The green solid line denotes the reference model atmosphere (RMA) median profile, see the text for details. The purple solid line corresponds to the median backscatter obtained from LITE. The bars correspond to the 10–25%, 25–75% and 75–90% percentiles. The dash-dotted blue line shows molecular backscatter for reference. The numbers on the right hand side of the figure show the number of data used for the CALIPSO aerosol statistics.

between the reference "clean" and "dirty" aerosol conditions in the free troposphere and stratosphere. In the tropics a drop in the aerosol density is observed between 8 and 15 km that can be explained by an overestimate of cirrus cloud transmission. The algorithm then finds negative aerosol backscatter

solutions below the cirrus cloud. The aerosol content in the PBL is generally smaller than for the reference curves that can be explained by the conservative algorithm divergence check criterion that prevents large backscatter retrievals in aerosol-rich areas.

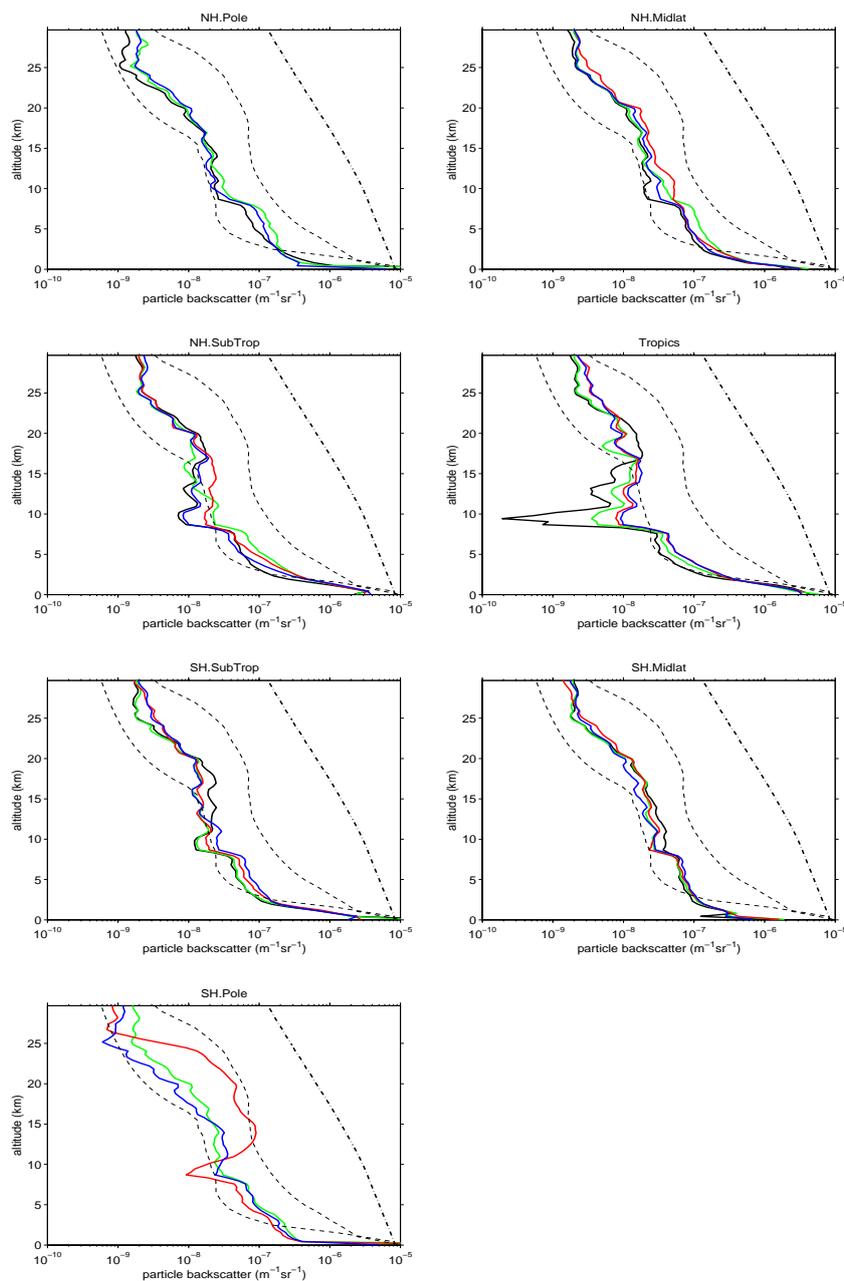


Fig. 8. CALIPSO median aerosol backscatter ($\text{m}^{-1} \text{sr}^{-1}$) at 355 nm wavelength for January (black), April (green), August (red) and October (blue) 2007. The dashed lines represent the reference “clean” and “dirty” median aerosol backscatter profiles and the dash-dotted line is molecular backscatter similar as in Fig. 7. The reference curves are fixed for all climate zones.

The median aerosol scattering ratio is about 1.01 throughout the troposphere increasing to 1.1 in the PBL. Because PSCs above 20 km are treated as fully transparent aerosols the aerosol scattering ratio over the South-Pole region in August is substantial with a median value between 1.05 and 1.1 and a scattering ratio exceeding a value of 2 in 10% of the scenes (not shown). The full transparency assumption yields a slight underestimate of the retrieved ice particle backscatter of PSCs.

Limitations of the lidar inversion algorithm have been discussed. These include imperfect knowledge of the aerosol lidar ratio that may locally under- or overestimate the retrieved aerosol backscatter. The strong database requirement of high resolution has the negative side-effect of substantial noise in the signal to be processed. The aerosol retrieval is thus contaminated with noise. The estimation of isolated particle layers may also be contaminated through signal noise. Noise filtering techniques prior to the backscatter retrieval could

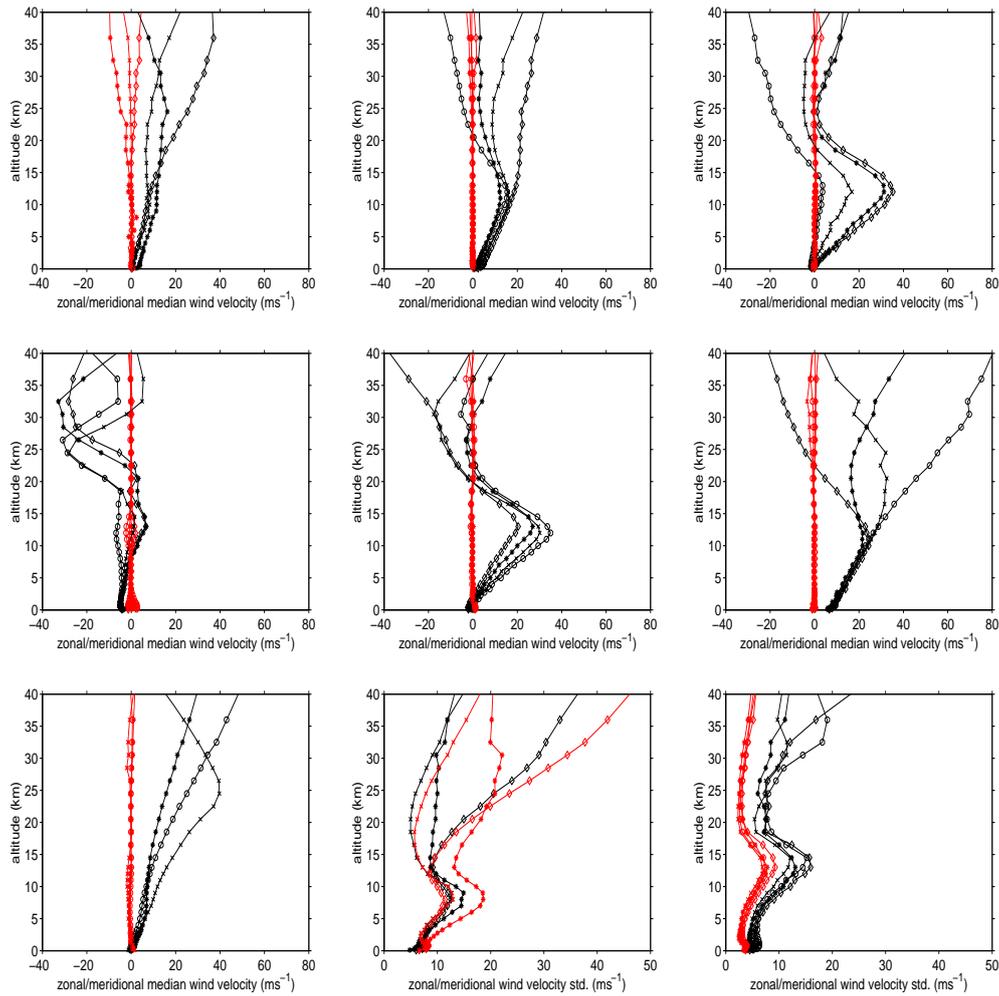


Fig. 9. ECMWF model wind statistics along CALIPSO orbits. Median zonal (black) and meridional (red) wind profiles (ms^{-1}) for the following climate zones: NH pole, NH mid-latitudes, NH Subtropics (top row from left to right), Tropics, SH subtropics, SH mid-latitudes (middle row) and SH pole (bottom row, left panel). In addition, the standard deviation from the mean wind for the NH polar region (bottom, middle) and Tropics (bottom, right). The symbols denote the various months: January (diamond), April (star), August (circle) and October (crosses).

alleviate these problems but most probably at the expense of losing real aerosol structures. It was thus decided to overestimate the atmospheric aerosol variability rather than to underestimate it. No aerosol backscatter can be determined below dense clouds. Despite these limitations the retrieved aerosol and cloud statistics are compatible with available datasets and the resulting dataset forms an extensive and high-quality dataset for Aeolus related studies. The retrieval algorithm may be improved by making use of the emerging (growing) availability of CALIPSO level-2 products. In particular the cloud-aerosol mask product and characterization of aerosols and corresponding aerosol lidar ratio can be used as a-priori knowledge in the retrieval algorithm.

3 ECMWF model parameters

The result from the previous section is a dataset of atmosphere optical properties along the CALIPSO nighttime orbit at 3.5 km horizontal and 125 m vertical resolution. The database is completed with meteorological parameters along the CALIPSO orbit. Hereto model field parameters from the European Centre for Medium-Range Weather Forecasts (ECMWF) analyses are horizontally interpolated to locations along the CALIPSO orbit and vertically to 125 m resolution, consistent with the database resolution. The ECMWF model horizontal resolution in 2007 is about 25 km with 91 vertical levels (T799L91). The model vertical levels extend from the surface up to 0.01 hPa, i.e., about 85 km. The vertical resolution is not constant but reduces with altitude from several tens of meters in the PBL to more than 2 km above 40 km.

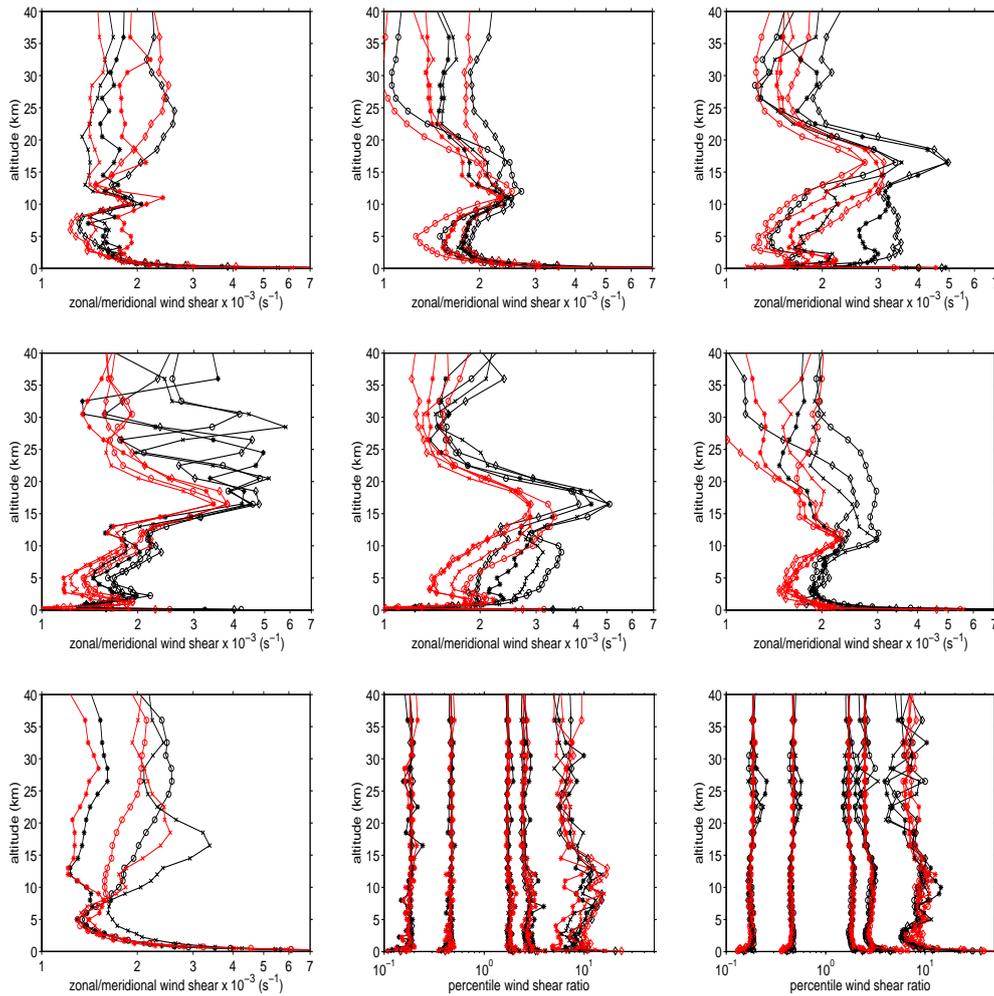


Fig. 10. Same as Fig. 9 but now for the median zonal (black) and meridional (red) absolute wind-shear (s^{-1}). The x -axis has a logarithmic scale. The middle (NH polar region) and right (tropics) panel in the bottom row show the ratio of percentile profiles and the median profile for (from left to right) the 10%, 25%, 75%, 90% and 100% (absolute maximum) percentile profiles.

ECMWF model analyses are archived daily at 6 hour resolution starting at 00:00 UTC. We perform no temporal interpolation in between analyses, i.e., the atmosphere is assumed frozen for 6h around the analysis time. Analyses closest in time to the CALIPSO observation times are used for the spatial interpolation to the CALIPSO orbit. As a result the mean absolute time delay between ECMWF model parameters and the CALIPSO optical parameters is 90 min with a maximum delay of 3h. Meteorological parameters extracted from the ECMWF model analysis fields and used in the database include: pressure, temperature, the 3 (zonal, meridional and vertical) wind components and specific humidity, all as a function of height above mean sea level. This includes conversion from model to pressure levels, that is inherent in ECMWF processing software, and conversion from pressure to height levels based on hydrostatic balance, the equation of state and virtual temperature and taking into account orography.

An improved understanding and modelling of the atmospheric dynamics is the main objective of Aeolus. Figures 9 and 10 provide an overview of the model dynamics along the CALIPSO orbits for January, April, August and October 2007 as a function of the climate zones defined in section 2.1.4. Figure 9 shows statistics of the zonal and meridional wind components. The August statistics are missing in the NH polar region because of mission night-time CALIPSO data for this region in this time of year. Similarly the January statistics are missing in SH polar region. The median meridional wind is close to zero throughout the atmosphere for all seasons and climate zones. Zonal wind maxima are mostly found near the tropopause in particular in the mid-latitude and subtropical regions and at higher altitudes in the stratosphere. The zonal wind in the tropics is generally small in the troposphere with maximum absolute values found in the stratosphere. The strongest winds are found in the SH midlatitude region in August with zonal

wind speeds exceeding 80 ms^{-1} at 40 km altitude. For two typical climates zones the standard deviation relative to the mean wind is displayed in the middle and right panels in the bottom row of Fig. 9. Both the NH polar and tropical region show a local maximum variability in both the zonal and meridional wind components near the tropopause for all seasons. This is also clear from the absolute vertical wind-shear statistics of the horizontal wind components in Fig. 10, where absolute vertical wind-shear (s^{-1}) is defined as the absolute value of the vertical gradient of the horizontal wind component, i.e., $|\partial u(z)/\partial z|$ with $u(z)$ the zonal or meridional wind component at altitude z . Model maximum median wind-shear values in the polar and mid-latitude regions are between 2×10^{-3} and 3×10^{-3} (s^{-1}) and up to 5×10^{-3} (s^{-1}) in the subtropics and tropics. Of particular interest for Aeolus is the occurrence of extreme wind-shear events. The middle and right panel in the bottom row of Fig. 10 show the 10%, 25%, 75%, 90% and 100% percentile profiles relative to the median profile. Interestingly, the ratio of the percentile and median profiles is fairly constant throughout the atmosphere for the 10, 25, 75 and 90 percentiles with values of 0.19, 0.45, 2.0, 2.7 respectively. The maximum wind-shear shows more variability with ratio values ranging from 6 to 20. From these results it is concluded that wind-shear values exceeding 0.05 (s^{-1}), or 50 ms^{-1} over typical 1-km Aeolus range bins, are occasionally found in the tropical upper troposphere that is characterized by the presence of cirrus clouds, see Fig. 6 and Table 1. Atmospheric scenes with large wind-shears in combination with varying aerosol and/or cloud densities are most challenging for Aeolus to retrieve representative winds.

4 Database representativeness for Aeolus dawn-dusk orbit

The CALIPSO overpass at nighttime is at about 1:30 a.m. local time, while Aeolus will travel in a dawn-dusk orbit with local overpass times at 6 a.m. and 6 p.m. The question is whether the atmosphere as observed by CALIPSO and adopted in the database is representative for the atmosphere that will be observed by Aeolus. Hereto, the diurnal cycle of the atmospheric dynamics is considered in this section with an emphasis on the 0–6 a.m. period. Two global regions have been selected, both bounded at 70° N and 70° S . In meridional direction, the first (Pacific) region is bounded by 90° W and 105° W , covering parts of North and Central America and the South Pacific, the other (Atlantic) region is bounded by 15° W and 0° W , covering parts of Western-Europe, West-Africa and the South Atlantic. The motivation for these regions is the availability of ECMWF model fields analyses at 00:00, 06:00, 12:00 and 18:00 UTC that correspond to 6 p.m., 0 a.m., 6 a.m. and 12 a.m. local time for region 1 and to 0 a.m., 6 a.m., 12 a.m. and 6 p.m. local time for region 2.

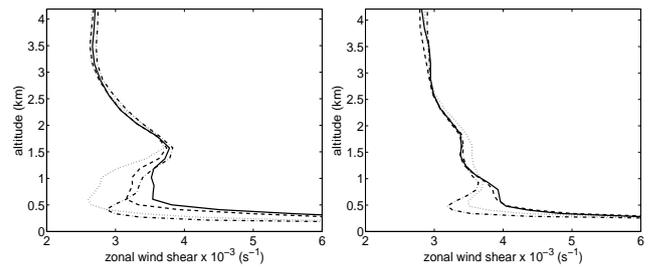


Fig. 11. ECMWF model mean absolute zonal wind-shear (s^{-1}) in the Pacific region 1 (left) and Atlantic region 2 (right) at 0 a.m. (dashed), 6 a.m. (solid), 12 a.m. (dash-dot) and 6 p.m. (dot) local time. The mean is taken over the 3-month 2007 winter period.

Figure 11 shows that the dynamical variability in the lower part of the atmosphere substantially differs between day and night. During the evening and night, the continent cools thus reducing the convective turbulence. An inversion layer may develop that decouples the PBL from the higher troposphere. The strong geostrophic wind (low-level nocturnal jet) then gives a large wind-shear at the top of the PBL. During the morning, the sun heats the surface, triggering convection. The resulting turbulence mixes the PBL and tropospheric air, thus removing the inversion and associated nocturnal jet. This process is well visualized in Fig. 11 showing that the mean zonal-wind shear in the PBL has its maximum at 6 a.m. local time with slightly smaller values at 0 a.m. and the smallest values at noon and 6 p.m. local time in the lowest 1.5 km. This result indicates that the boundary layer further stabilizes after 0 a.m. The reduced wind-shear in the PBL at noon and 6 p.m. indicates that the overnight stabilized PBL disappears due to daytime turbulence over land. The strong wind gradients near the surface with values up to 15×10^{-3} (s^{-1}) are explained by friction and orography over the continents. Similar results were found for the summer periods for both regions with on average less wind shear in the PBL than in winter (not shown) that may be explained by an earlier sunrise in the morning, stronger convection during daytime and later sunset.

A closer inspection of the dynamics at 0 a.m. and 6 a.m. in the PBL in both regions showed large similarity over the oceans with noted differences mainly persisting over the continents (not shown). Figure 11 also shows that the diurnal variations are mainly limited to the lowest 1.5 km. At higher altitudes the mean flow is less sensitive to day and night and local (land/sea) convective processes that evolve and decay during the day and is mainly driven by the large-scale atmospheric motions. From these results we conclude that the variability of the atmospheric dynamics along the CALIPSO orbit, and adopted in the database, well represents the atmosphere as will be observed by Aeolus in the free troposphere and in the PBL over the oceans. The continental PBL is generally further stabilizing after the CALIPSO overpass at 1:30 a.m. local time and thus the nocturnal jet at the top of

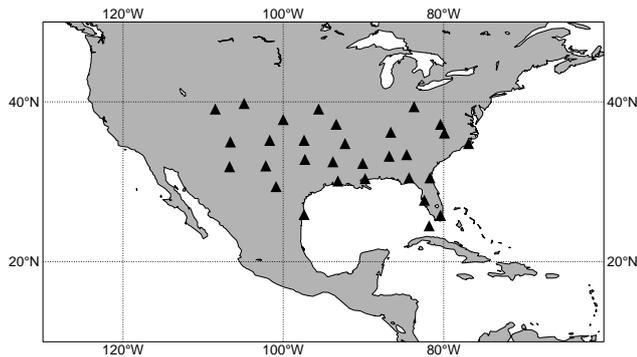


Fig. 12. Locations of high-resolution radiosondes in the US subtropical region between 20° N and 40° N and between 75° W and 110° W.

the PBL is slightly underestimated in the database as compared to the situation at dawn. At dusk (6 p.m. local time), the wind shear is generally substantially smaller than at dawn and similar to the shear at noon on average. The most challenging dynamical scenes for Aeolus in the lower part of the atmosphere are therefore found at dawn while at dusk the remaining clouds from daytime convective processes pose a challenge for Aeolus.

Here, we note that the tuning of the ECMWF model turbulence scheme is to give maximum performance of medium-range (5–10 days) weather forecasts, rather than short-term (up to two days) forecasts. The model overestimates the PBL mixing in stable conditions and as a consequence the nocturnal boundary layer low-level jet is underestimated (Baas, 2009). We further elaborate on ECMWF model shear representation in the next section.

4.1 ECMWF model validation against high resolution radiosondes

The simulated atmosphere of global NWP models is a smooth representation of the real atmosphere. Despite the gradual refinement of the model grid over the last couple of decades, the ability to resolve atmospheric processes at the smallest scales has improved to a lesser extent. In other words, the effective model resolution is substantially coarser than the model grid size (Frehlich et al., 2008), (Skamarock, 2004). As a consequence the vertical gradient of the horizontal wind components, i.e., the model wind shear, is underestimated with respect to the real atmosphere as will be observed by Aeolus. This has been further investigated by comparing wind observations from a network of high-resolution radiosonde data in the US, see Fig. 12, with 12-hour ECMWF model wind forecasts. These data sources are independent in the sense that the radiosonde observations have not been assimilated in the analysis that is used as forecast initial state. The radiosonde data are averaged over 12 s yielding a vertical resolution of about 60 m. (Houchi et al., 2010). Figure 13

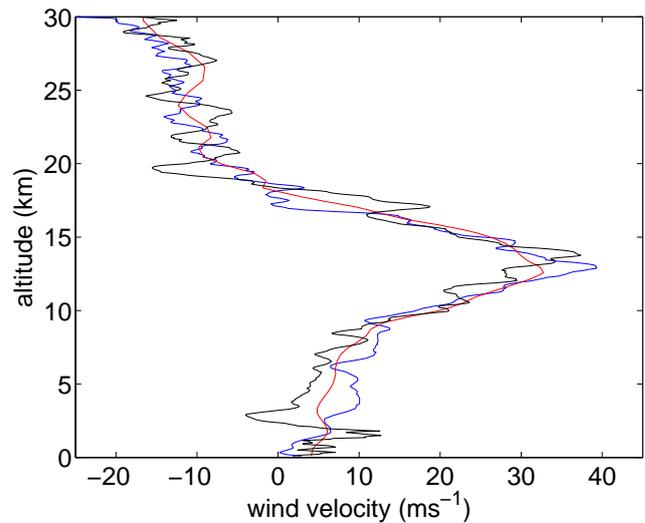


Fig. 13. High-resolution radiosonde zonal-wind profile (blue) versus ECMWF model background from a 12-h forecast (red). The black profile is the adapted model profile, see the text for details.

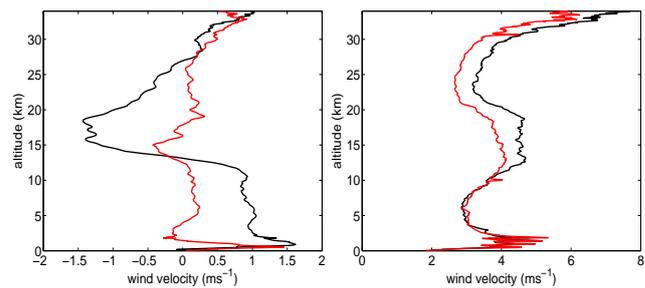


Fig. 14. Observation minus background statistics from high-resolution radiosondes and ECMWF model 12-h forecast. Bias (left) and standard deviation (right) for the zonal (black) and meridional wind component (red).

shows a typical example of a radiosonde zonal wind profile and its model counterpart. The model closely follows the radiosonde curvature but lacks the small-scale structures. To characterize the missing structures, one year (2006) of radiosonde data from the 30 locations in the US, as displayed in Fig. 12, have been used to generate observation minus background, also denoted (o-b) or background departures, statistics. Here, the background is the 12-h ECMWF model forecast verifying at the same time as the radiosonde launch and interpolated to the radiosonde launch location. Only radiosondes launched at 12:00 UTC have been used which is between 5 and 7 a.m. local time for the region under investigation. A total of 10950 radiosonde profiles were used in this study. Figure 14 shows a positive bias in the lower part of the atmosphere, indicating an underestimate of the model flow in the PBL and lower troposphere, despite the generally high model resolution of tens to hundreds of meters in the lowest few kilometers near the surface; Houchi et al. (2010)

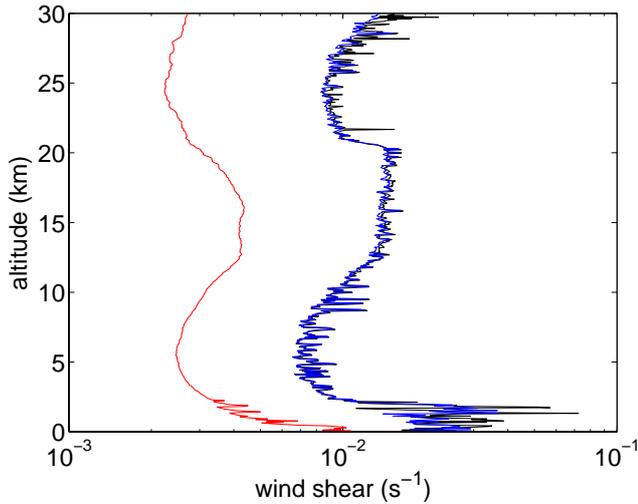


Fig. 15. Zonal wind-shear standard deviation (s^{-1}) for model winds (red curve), radiosonde winds (blue) and adapted model winds (black). The radiosonde wind and adapted model wind profiles largely overlap.

found a 1.5–2 km effective vertical resolution throughout the atmosphere. The maximum bias near 1–1.5 km is presumably linked to the model underestimate of nocturnal low-level jets as discussed in section 4. Also the standard deviation in the right panel of Fig. 14 shows a local maximum near the top of the PBL in the lower troposphere. In addition a maximum is found near the tropopause level. At higher altitudes in the stratosphere the (o-b) standard deviation increases substantially through the presence of gravity waves in the observations that are suppressed in NWP models. It is not believed that radiosonde drift from the launch location is the dominating source in (o-b) at high altitudes because the drift was generally found to be within the effective model resolution (Houchi et al., 2010).

The (o-b) dataset has been further used to define a procedure to adapt the relatively smooth database wind profiles by adding small-scale atmospheric structures. The adaptation is based on the eigenvalue decomposition of the positive definite (o-b) covariance matrix, also denoted the covariance matrix of background departures \mathbf{D}_b . Stoffelen et al. (2006) show that $\mathbf{D}_b = \mathbf{H}\mathbf{B}\mathbf{H}' + \mathbf{R}$ with \mathbf{H} the linearized observation operator that relates model fields to observations, \mathbf{B} is the background error covariance matrix, \mathbf{R} is the observation error covariance matrix and the prime denoting matrix transpose. The random measurement errors of radiosonde observations are assumed unbiased with an error standard deviation equal to one for raw data at 30 m resolution and $1/\sqrt{2}$ for 60 m resolution data (Houchi et al., 2010). The random part of \mathbf{R} then equals 1/2 times the identity matrix. \mathbf{R} also includes a so-called representativeness error. These are essentially the unresolved wind scales of the model and due to 3-D turbulence in the free troposphere. The background error

covariance matrix is related to the larger and model-resolved atmospheric scales. The objective of the NWP analysis is to analyze these scales and propagate them in the forecast. The potentially correlated part of \mathbf{R} is therefore related to lacking model small-scale atmospheric structures. Matrix \mathbf{D}_b is obtained from the (o-b) dataset of radiosondes (o) and model background (b). After subtraction of the random part of \mathbf{R} , to reduce instrument noise in the model wind adaptation, an eigenvalue decomposition of the resulting positive definite symmetric matrix yields $\mathbf{U}\mathbf{\Lambda}\mathbf{U}'$ with \mathbf{U} the unitary matrix of singular vectors and $\mathbf{\Lambda}$ the diagonal matrix of strictly positive eigenvalues. The wind adaptation vector, $\delta\mathbf{u}$, is then obtained from $\delta\mathbf{u} = \mathbf{U}\mathbf{\Lambda}^{1/2}\delta\mathbf{n}$ with $\delta\mathbf{n}$ a noise vector from normally Gaussian distributed random numbers with zero mean and standard deviation 1, i.e., $\delta\mathbf{u}$ is a correlated noise vector realization representing the combined large-scale and small-scale model errors. For the definition of the wind adaptation we are mainly interested in the latter, but a well defined methodology to separate these errors is not available. High-pass filtering or averaging techniques as proposed in (Houchi et al., 2010) can be used for this purpose but this needs further investigation. Considering the relatively low effective spatial resolution of nowadays global models we assume the small-scale model errors as the dominant error source. The black line in Fig. 13 shows an example of an adapted model wind profile using the above method.

Although the ECMWF model wind follows the mean flow very well on average, Fig. 15 shows that the ECMWF model underestimates the atmosphere wind variability. Here, the red curve shows the ECMWF model zonal wind-shear standard deviation as obtained from one year of model wind data at the radiosonde locations displayed in Fig. 12. The blue curve shows the wind-shear standard deviation from the radiosonde data which is on average a factor of 3.5 larger than the model shear with a maximum value of 10 found in the PBL between 1 and 2.5 km above the surface. The zonal wind-shear standard deviation of the adapted winds (black curve) is well in agreement with the radiosondes. Similar results apply for the meridional wind component (not shown).

These results show that the characteristics of the adapted winds are in better agreement with the real atmosphere as will be observed by Aeolus. Database model wind adaptation is needed for a realistic assessment of the occurrence of heterogeneous atmospheric scenes that are challenging for Aeolus wind profile processing. The current implementation of the model wind adaptations ignores correlation with optical structures. It is envisaged to further investigate this aspect using the combined information from temperature, humidity and wind in high-resolution radiosonde profiles. The database of high-resolution radiosondes is still under development at KNMI. Ultimately, a global coverage may provide wind adaptations as a function of global region.

5 Summary and conclusions

The definition and contents of an atmospheric database as a data source for various studies in preparation for ESA's Aeolus mission has been discussed. The database is built up of atmospheric backscatter and extinction at the 355 nm Aeolus laser wavelength as retrieved from CALIPSO night-time total attenuated backscatter data and completed with meteorological parameters from the ECMWF model. The database horizontal and vertical resolution of 3.5 km and 125 m respectively is a compromise to enable the simulation of small-scale atmospheric dynamical and optical processes inside the Aeolus sampling volumes of typically 50 km along track and 1 km in the vertical on one hand and to retrieve atmosphere optical properties with sufficient SNR from the CALIPSO lidar inversion algorithm on the other hand. The database has a full global coverage from the surface up to 40 km altitude and covers all seasons to simulate the seasonal dependence of the observed quantities. Besides the primary wind product this includes the characterization of aerosol and cloud optical properties, the so-called Aeolus secondary products (Flamant et al., 2008). The availability of a large database allows a statistical evaluation of the Aeolus mission components under investigation.

A lidar inversion algorithm has been discussed to retrieve aerosol and cloud backscatter and extinction from the CALIPSO level-1 532 nm total attenuated backscatter product at the database resolution. The objective of the algorithm is not to generate a level-2 aerosol and cloud product, but merely to retrieve the atmosphere backscatter and extinction properties at 532 nm wavelength and to convert these to 355 nm wavelength. The exact characterization of the scattering constituents is of less relevance in this context. The algorithm includes cloud detection and the calculation of the cloud transmission for isolated cloud layers. An iterative procedure is then used to calculate the cloud lidar ratio and cloud backscatter from the best fit of the retrieval function to the measured data inside the cloud layer. Aerosol backscatter and extinction in the remaining part of the atmosphere are obtained from solving the retrieval function and assuming a constant value for the aerosol lidar ratio of 35 (sr) that is typical for background aerosol. An underestimate of the aerosol lidar ratio in the algorithm results in an underestimate of the retrieved aerosol backscatter and vice versa. For the selected lidar ratio value this is the case for instance for atmospheric events of substantial smoke (biomass burning), desert dust and polluted continental aerosol. The retrieval function is prone to divergence in case of incorrectly specified values for the aerosol and/or cloud lidar ratio. The algorithm includes divergence checks to improve its robustness and was applied to 4 months of CALIPSO data without manual intervention needed.

To validate the lidar inversion algorithm, statistics of the resulting particle backscatter have been determined, including statistics of cloud coverage and cloud and aerosol opti-

cal properties. Special attention was paid to tropical cirrus clouds for which a mean lidar ratio of 18 (sr) was found in agreement with values found in the literature. The transmission of tropical cirrus clouds is generally large with a median value of 0.91 at 355 nm wavelength. This is important for Aeolus to enable the retrieval of high quality winds in the lower part of the tropical troposphere below cirrus clouds. Retrieved aerosol backscatter profiles were generally well bounded by results from flight campaigns over the Atlantic in the 1989 period, that has been characterized as a relatively "clean" period, and results from LITE in the 1994 period, that has been characterized as relatively "dirty" in particular in the stratosphere (1991 Pinatubo eruption heritage).

Meteorological parameters to complete the database were extracted from ECMWF model field analyses, horizontally interpolated to the CALIPSO orbit ground track and vertically to 125 m, consistent with the database resolution. Wind and wind-shear statistics along the CALIPSO orbit were presented that are of relevance for further Aeolus studies. The representativeness of the atmosphere as adopted in the database, i.e., corresponding to the 1:30 CALIPSO local overpass time at night, for the atmosphere as observed by the Aeolus dawn-dusk overpass was verified. It was found that the atmospheric dynamics in the boundary layer over the continents is driven by local land-sea, radiative cooling and convective processes and shows a clear diurnal cycle with the development of strong wind shear (low-level jet) at the top of the nocturnal boundary layer. It is shown that the atmosphere further stabilizes after 1:30 a.m. and the current database winds somewhat underestimate the nocturnal jet. In addition it was demonstrated that the ECMWF model underestimates the nocturnal boundary layer jet in an intercomparison of model fields with high-resolution radiosonde data. A clear diurnal cycle was not observed over the oceans and in the free troposphere and the database here well represents Aeolus measurements.

The global model winds are a smooth representation of the real atmospheric flow. It was shown that model winds substantially underestimate the atmospheric wind variability. A method was described to adapt the model winds by adding observed atmospheric wind variability. The adaptation is based on an intercomparison of model winds and radiosonde winds obtained at a high vertical resolution of 60 m. Statistics of observation minus model background clearly show that the model underestimates the low-level jet at the top of the nocturnal boundary layer. The adapted model winds are more representative for the atmospheric dynamics as will be observed by Aeolus and will result in a better assessment of the occurrence of heterogeneous atmospheric scenes that are challenging for Aeolus to process. These include heterogeneous atmospheric scenes with large optical and dynamical variability.

The atmospheric database is currently used for many studies in preparation for the launch of Aeolus. These include the development of advanced L1B and L2B processing

algorithms for the retrieval of wind products from the measured data. In addition the database is used to study zero wind calibration opportunities from ground returns. Aeolus offers the possibility of targeting geographic regions by changing the vertical range gate binning in orbit, with a maximum of on average eight times per orbit. Studies are ongoing to optimize the vertical sampling of Aeolus for the various climate zones and as a function of season. This requires the global distribution of atmospheric particles and their correlation with the atmospheric dynamics. The database is a unique data source in this respect. The recent change from burst to continuous pulsed laser mode raises new scientific questions on the optimal sampling and processing of measured data to yield maximum impact for NWP. The atmospheric database will play an important role in the process of answering these questions.

Potential improvements to the lidar inversion algorithm can be made from the emerging availability of CALIPSO level-2 products (Omar et al., 2009), (Young et al., 2009). In particular the cloud-aerosol mask product and the characterization of aerosols and their corresponding aerosol lidar ratio can be used as a-priori knowledge in the retrieval algorithm. Besides the primary wind product the database can also be used to test Aeolus level-2A algorithms for the detection of clouds and aerosol layers and the characterization of their optical properties, the so-called Aeolus secondary products (Flamant et al., 2008).

Appendix A

Backscatter retrieval function

To retrieve particle backscatter from measured attenuated backscatter, at altitude z , it was shown in section 2.1 that the so-called retrieval function, Eq. (9), needs to be solved. In its most general form the retrieval function is written as:

$$f(x) = (m+x)e^{-\gamma x} - c \quad (\text{A1})$$

where we excluded z for clarity. Here, m corresponds to molecular backscatter β_m ($\text{m}^{-1} \text{sr}^{-1}$), $\gamma = S^* \Delta z/2$ is the product of the effective lidar ratio S^* (sr) and half the bin size $\Delta z/2$ (m). Parameter c is discussed below. The solution x for $f(x) = 0$ is the particle backscatter β_p ($\text{m}^{-1} \text{sr}^{-1}$) solution at altitude z .

From Eq. (9), $c = \beta_m \beta'' / T_p^2$, with β'' the ratio of the measured attenuated backscatter and the attenuated backscatter for a particle-free (or molecules-only) atmosphere, see Eq. (6), and T_p is the total estimated one-way transmission from particles between the satellite and z . Parameter c is thus a function of (i) the known atmosphere molecular backscatter and extinction (from known atmospheric pressure and temperature), (ii) the measured attenuated backscatter and (iii) the estimated total particle optical depth between the satellite and z that in turn is a function of the total retrieved parti-

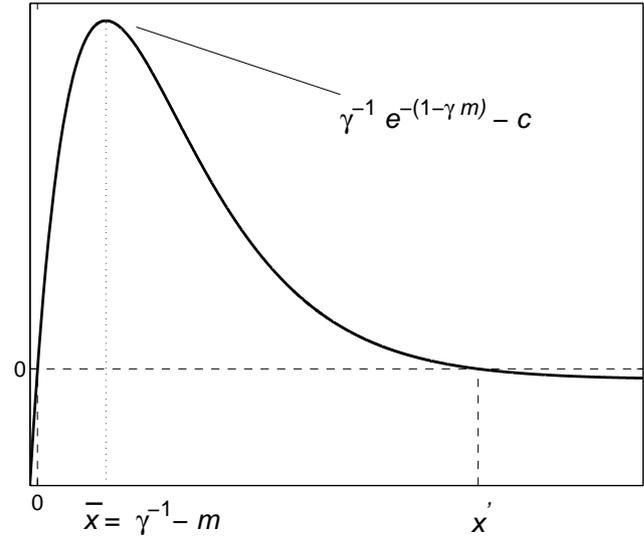


Fig. A1. Backscatter retrieval function, $f(x)$, for $c = m$, see Eq. (A1). The dotted line corresponds to the function maximum at $\bar{x} = \gamma^{-1} - m$ with function value $\gamma^{-1} \exp[-(1-\gamma m)] - c$. The dashed lines correspond to the two solutions $x = \{0, x'\}$ for $f(x) = 0$.

cle backscatter at atmospheric levels above z and parameter γ . The retrieved particle backscatter profile is thus obtained from a recursive procedure of solving Eq. (A1). The procedure is initiated at high altitudes where virtually no particles are found and the final result is sensitive to the parameter γ used in the retrieval function both directly through the argument in the exponent and indirectly through parameter c .

Parameter x is the only unknown in Eq. (A1) but an analytical solution for $f(x) = 0$ is not known and the solution is found iteratively. Figure A1 shows that $f(x)$ has a global maximum at $\bar{x} = \gamma^{-1} - m$ with function value $\gamma^{-1} \exp[-(1-\gamma m)] - c$. In the limit for x to infinite $f(x)$ goes to $-c$ and $f(0) = m - c$. The retrieval function has either no, one or two solutions depending on the value of c . For values of c smaller than m , as used in the figure, the function curve is lifted while shifted downwards for values of c larger than m . For large values of c the total function curve drops below zero and no solution for the retrieval function is found. This is called divergence of the algorithm. Divergence occurs when the value for the particle lidar ratio, S^* , in the algorithm is substantially larger than in the real atmosphere. The total particle transmission is then underestimated (smaller estimated T_p than in reality), yielding too large values for c eventually causing divergence. The algorithm checks for divergence and reduces the particle lidar ratio accordingly.

For the idealized situation of a particle-free atmosphere and no signal noise, β'' equals 1 throughout the profile. T_p equals one at the top of the atmosphere and thus c equals m . Then $x = 0$ is the obvious and physically realistic solution. There is however a second possible solution on the right

hand side of \bar{x} . For typical background aerosol with $S^* = 35$ (sr), $\Delta z = 125$ (m) and a typical molecular backscatter of 4×10^{-7} ($\text{m}^{-1} \text{sr}^{-1}$) at 532 nm wavelength at 15 km altitude, γ equals 2188 and $\bar{x} = 4.6 \times 10^{-4}$ ($\text{m}^{-1} \text{sr}^{-1}$) yielding an unrealistic scattering ratio value of about 1143. This calculation implies that in case of two possible solutions for the retrieval function the solution left from the function maximum is the physically most reliable solution. The estimated total particle transmission T_p then stays unchanged at a value of one and thus c equals m when going further down in the profile. The final result is the input zero particle backscatter profile.

Instrument noise contaminates the solution of the retrieval function. To demonstrate the effect we again consider an atmospheric region at high altitudes that is free of particles. From Eq. (4) and setting particle backscatter and transmission terms to zero and one respectively yields for the measured attenuated backscatter $\beta'(i) = \beta_m(i)T_m^2(i-1)\tau_m^2(i) + n(i)$ with n denoting the signal noise contribution. From Eqs. (5,6) we have for $\beta_m\beta'' = 1 + n(i)/[\beta_m(i)T_m^2(i-1)\tau_m^2(i)]$ and thus for c in the retrieval function Eq. (A1), $c = m + n/T_m$ where we removed the index i for simplicity and with T_m the total two-way molecular transmission from the top of atmosphere until and including bin i . The situation $n = 0$ was discussed above. For a negative noise realisation of n , the function curve in Fig. A1 is lifted and the solution for x will be either negative or right from \bar{x} . Both solutions are physically unrealistic and the solution is set to zero in the algorithm. For a positive noise realisation of n the function curve is shifted downwards and the solution for x will be positive. The retrieval function can not discriminate between noise and particles and the solution is accepted. As a consequence the estimated atmosphere particle transmission T_p decreases and c is larger than m in the bins below, yielding increasing backscatter solutions further down in the profile. The noise contamination thus propagates through the profile yielding physically unrealistic large backscatter values in the upper part of the atmosphere. The aerosol lidar ratio is therefore set to zero in the lidar inversion algorithm for atmospheric regions that are known to be free of aerosols.

Similar arguments apply in case of an incorrect specification of the aerosol lidar ratio. As discussed in Sect. 2.1.2 an underestimate of the aerosol lidar ratio in lofted aerosol layers yields an overestimate of the calculated layer transmission, i.e., the value for c in the retrieval function is underestimated (the function curve is lifted) resulting in negative aerosol backscatter values below the aerosol layer. Vice versa, an overestimate of the aerosol lidar ratio yields an underestimate of the layer transmission and too large values for c (the function curve is shifted downwards) in underlying bins potentially causing divergence of the algorithm further down in the profile. The algorithm checks for divergence and reduces the lidar ratio accordingly. In general can be stated that an underestimate of the particle lidar ratio in the algorithm results in an underestimate of retrieved mean particle backscatter in the profile and vice versa.

Acknowledgements. This paper is one of the results of the project VAMP (Vertical Aeolus Measurement Positioning) funded by the European Space Agency (ESA). The authors thank ESA and members of the project team for stimulating discussions. Special thanks go to the NASA Langley Atmospheric Science Data Center for making available the CALIPSO data.

Edited by: C. Senff

References

- Ackermann, J.: The Extinction-to-Backscatter Ratio of Tropospheric Aerosol: A Numerical Study, *J. Atmos. Oceanic Technol.*, 15, 1043–1050, 1998.
- Baas, P.: Turbulence and Low-Level Jets in the Stable Boundary Layer, Wageningen University, The Netherlands, PhD Thesis, 152 pp., 2009.
- Blossey, P. N., Bretherton, C. S., Cetrone, J., and Khairoutdinov, M.: Cloud-Resolving Model Simulations of KWAJEX: Model Sensitivities and Comparisons with Satellite and Radar Observations, *J. Atmos. Sci.*, 64, 1488–1508, 2007.
- Bodhaine, B. A., Wood, N. B., Dutto, E. G., and Slusser, J. R.: On Rayleigh Optical Depth Calculations, *J. Atmos. Oceanic Technol.*, 16, 1854–1861, 1999.
- Dabas, A., Denneulin, M., Flamant, P., Loth, C., Garnier, A., and Dolfi Bouteyre, A.: Correcting winds measured with a Rayleigh Doppler lidar from pressure and temperature effects, *Tellus A*, 60(2), 206–215, 2008.
- Flamant, P., Cuesta, J., Denneulin, M., Dabas, A., and Huber, D.: ADM-Aeolus retrieval algorithms for aerosol and cloud products, *Tellus A*, 60(2), 273–288, 2008.
- Flamant, P., Cuesta, J., Dabas, A., Huber, D., and Olivier, M.: ADM-Aeolus cloud and aerosol spin-off products, L2A Algorithm Theoretical Basis Document (ATBD) – version 4.2. Report document AE_TN_IPSL_GS_001 available from ESA under contract number 18366/04/NL/MM CR 1, 2008.
- Frehlich, R. and Sharman, R.: The Use of Structure Functions and Spectra from Numerical Model Output to Determine Effective Model Resolution, *Mon. Weather Rev.*, 136, 1537–1553, 2008.
- Hogan R. J.: Fast Lidar and Radar Multiple-Scattering Models. Part I: Small-Angle Scattering Using the Photon Variance-Covariance Method, *J. Atmos. Sci.*, 65, 3621–3635, 2008.
- Houchi, K., Stoffelen, A., Marseille, G. J., and de Kloe, J.: Comparison of Wind and Wind-Shear Climatologies Derived from High-Resolution Radiosondes and the ECMWF Model, *J. Geophys. Res.*, 115, D22123, doi:10.1029/2009JD013196, 2010.
- Marseille, G. J. and Stoffelen, A.: Simulation of wind profiles from a space-borne Doppler wind lidar, *Q. J. Roy. Meteorol. Soc.*, 129, 3079–3098, 2003.
- Mattis, I., Ansmann, A., Müller D., Wandinger, U., and Althausen, D.: Multiyear aerosol observations with dual-wavelength Raman lidar in the framework of EARLINET, *J. Geophys. Res.*, 109, D13203, doi:10.1029/2004JD004600, 2004.
- Omar, A. H., Winker, D. M., Kittaka, C., Vaughan, M. A., Liu, Z., Hu, Y., Trepte, C. R., Rogers, R. R., Ferrare, R. A., Lee, K. P., Kuehn, R. E., and Hostetler, C. A.: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, *J. Atmos. Oceanic Technol.*, 26, 1994–2014, 2009.

- Penndorf, R.: Tables of the refractive index for standard air and the Rayleigh scattering coefficient for the spectral region between 0.2 and 20 microns and their application to atmospheric optics, *J. Opt. Soc. Amer.*, 47, 176–182, 1955.
- Pinnick, R. G., Jennings, S. G., Chyleck, P., Ham, C., and Grandy, W. T.: Backscatter and extinction in water clouds, *J. Geophys. Res.*, 88, 6787–6796, 1983.
- Skamarock, W. C.: Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra, *Mon. Weather Rev.*, 132, 3019–3032, 2004.
- Schuster, G. L., Dubovik, O., and Holben, B. N.: Angstrom exponent and bimodal aerosol size distributions, *J. Geophys. Res.*, 111, D07207, doi:10.1029/2005JD006328, 2006.
- Spinhirne, J. D., Palm, S. P., Hart, W. D., Hlavka, D. L. and Welton, E. J.: Cloud and aerosol measurements from GLAS: overview and initial results, *Geophys. Res. Lett.*, 32, L22S03, doi:10.1029/2005GL023507, 2005.
- Stoffelen, A., Pailleux, J., Källén, E., Vaughan, J. M., Isaksen, I., Flamant, P., Wergen, W., Andersson, E., Schyberg, H., Culoma, A., Meynart, M., Endemann, M., and Ingmann, P.: The Atmospheric dynamics mission for global wind measurement, *B. Am. Meteorol. Soc.*, 86, 73–87, 2005.
- Stoffelen, A., Marseille, G. J., Bouttier, F., Vasiljevic, D., de Haan, S., and Cardinali, C.: ADM-Aeolus Doppler wind lidar Observing System Simulation Experiment, *Q. J. Roy. Meteorol. Soc.*, 132, 1927–1947, 2006.
- Tan, D., Andersson, E., de Kloe, J., Marseille, G. J., Stoffelen, A., Poli, P., Denneulin, M. L., Dabas, A., Huber, D., Reitebuch, O., Flamant, P., LeRille, O., and Nett, H.: The ADM-Aeolus wind retrieval algorithms, *TellusA*, 60(2), 191–205, 2008.
- Vaughan, J. M., Brown, D. W., Nash, C., Alejandro S. B., and Koenig, G. G.: Atlantic atmospheric aerosol studies 2. Compendium of airborne backscatter measurements at 10.6 μm , *Geophys. Res. Lett.*, 100, 1043–1065, 1995.
- Vaughan, J. M., Geddes, N. J., Flamant, P. H., and Flesia, C.: Establishment of a backscatter coefficient and atmospheric database, Report available from ESA under contract number ESA-CR, 12510/97/NL/RE, 1998.
- Young, S. A.: Analysis of lidar backscatter profiles in optically thin clouds, *Appl. Opt.*, 34(30), 7019–7031, 1995.
- Young, S. A. and Vaughan, M. A.: The Retrieval of Profiles of Particulate Extinction from Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observation (CALIPSO) Data: Algorithm Description, *J. Atmos. Oceanic Technol.*, special CALIPSO edition, 1105–1119, 2009.
- Weissmann, M. and Cardinali, C.: Impact of airborne Doppler lidar observations on ECMWF forecasts, *Q. J. Roy. Meteorol. Soc.*, 133, 107–116, 2007.
- Winker, D. M.: Accounting for Multiple Scattering in Retrievals from Space Lidar, *Proc. SPIE*, 5059, 128–139, 2003.
- Winker, D. M., Couch, R. H., and McCormick, M. P.: An overview of LITE, NASA's Lidar In Space Technology Experiment, *Proc. SPIE*, 84, 164–180, 1996.
- Winker, D. M., Hunt, W. H., and McGill, M. J.: Initial performance assessment of CALIOP, *Geophys. Res. Lett.*, 34, L19803, doi:10.1029/2007GL030135, 2007.