Long-term validation of Aeolus L2B wind products at Punta Arenas, Chile and Leipzig, Germany

Holger Baars[1], Joshua Walchester[1,3], Elizaveta Basharova[1,3], Henriette Gebauer[1,3], Martin Radenz[1], Johannes Bühl[1], Boris Barja[2], Ulla Wandinger[1], and Patric Seifert[1]

[1]Leibniz Institute for Tropospheric Research (TROPOS), Leipzig, Germany
[2]Atmospheric Research Laboratory, University of Magallanes, Punta Arenas, Chile
[3]University of Leipzig, Leipzig, Germany

Correspondence: Holger Baars, baars@tropos.de

Abstract. Ground-based observations of horizontal winds have been performed at Leipzig (51.35° N, 12.43° E), Germany, and at Punta Arenas (53.15° S, 70.91° W), Chile, in the framework of the German initiative EVAA (Experimental Validation and Assimilation of Aeolus observations) with respect to the validation of the Mie and Rayleigh wind products of Aeolus (L2B data). In Leipzig, at the Leibniz Institute for Tropospheric Research (TROPOS), radiosondes have been launched for the Aeolus overpasses on each Friday (ascending orbit) since mid of May 2019. In Punta Arenas, scanning Doppler cloud radar observations have been performed in the framework of the DACAPO-PESO campaign (dacapo.tropos.de) for more than 3 years from the end of 2018 until the end of 2021 and could be used to validate Aeolus measurements on its ascending and descending orbit. We present two case studies and long-term statistics of the horizontal winds derived with the ground-based reference instruments compared to Aeolus Horizontal Line-of-Sight (HLOS) winds. The wind products of Aeolus considered are the Mie cloudy and Rayleigh clear product. It was found that the deviation of the Aeolus HLOS winds from the ground reference is usually of Gaussian shape, which allowed the use of the median bias and the scaled median absolute deviation (MAD) for the determination of the systematic and random errors of Aeolus wind products, respectively. The case study from August 2020 with impressive atmospheric conditions at Punta Arenas shows that Aeolus is able to capture strong wind speeds up to more than 100 $\mathrm{ms^{-1}}$.

The long-term validation has been performed in Punta Arenas covering the period from December 2018 to November 2021 and in Leipzig from May 2019 until September 2022. This analysis showed that the systematic error of the Aeolus wind products could be significantly lowered during the mission lifetime with the changes introduced into the processing chain (different versions are called baselines). While in the early mission phase, systematic errors of more than 2 $\mathrm{ms^{-1}}$ (absolute values) were observed for both wind types (Mie and Rayleigh), these biases could be reduced with the algorithm improvements, such as the introduction of the correction for temperature fluctuations at the main telescope of Aeolus (M1 temperature correction) with Baseline 09. Hence, since Baseline 10, a significant improvement of the Aeolus data was found leading to a low systematic error (close to 0 $\mathrm{ms^{-1}}$) and nearly similar values for the mid-latitudinal sites on both hemispheres. The random errors for both wind products were first decreasing with increasing baseline but later increasing again due to performance losses

of the Aeolus emitter. ~~However,~~ Nevertheless, no significant increase in the systematic error ~~is only slightly affected by this issue, so that~~ of the Aeolus wind products was found. Thus, one can conclude that the uncertainty introduced by the reduced atmospheric return signal received by Aeolus is mostly affecting the random error.

Even when considering ~~these issues~~all the challenges during the mission, we can confirm the general validity of Aeolus observations during its lifetime. ~~This proves the general concept of~~ Therefore, this space explorer mission could demonstrate that it is possible to perform active wind observations from space with the applied technique.

1 Introduction

In 2018, the Aeolus satellite of the European Space Agency (ESA) was launched with the goal to improve weather forecast through global measurements of wind profiles (Stoffelen et al., 2005; Reitebuch, 2012). To obtain ~~vertical~~ vertically resolved wind measurements around the globe, the High-Spectral-Resolution (HSR) Doppler lidar ALADIN (Atmospheric Laser Doppler Instrument) was installed on board. It has been the first time ~~,~~ that a lidar with Doppler capabilities as well as with ~~high spectral resolution~~ high-spectral-resolution capabilities has been operated in space. Given this unique and novel space-borne technique, it is possible to actively measure vertical profiles of the line-of-sight (LOS) wind in clear sky by using air molecules as tracer (Rayleigh methodology) and in cloudy atmospheric regions ~~,~~ by using cloud particles as tracer ~~(Mie methodology - de Kloe et al., 2023; Tan et al., 2008; Baars et al., 2020b)~~ (Mie methodology, de Kloe et al., 2023; Tan et al., 2008; Baars et al., 2020b). The profiles of ~~line-of-sight~~ LOS wind velocity (35° off nadir) are then projected to a plane parallel to the Earth's surface to obtain the horizontal line-of-sight (HLOS) wind, i.e., one wind component of the horizontal wind vector (near west-east direction). Besides that, the ~~technology~~ HSR lidar technology onboard Aeolus can be used to obtain profiles of aerosol and cloud optical properties as ~~spin-off product (e.g., Baars et al., 2021; Flament et al., 2021; Siomos et al., 2021)~~ spin-off product (e.g., Flament et al., 2021; Baars et al., 2021; Siomos et al., 2021; Abril-Gago et al., 2022a; Gkikas et al., 2023).

The main goal of the mission is the assimilation of the wind products into numerical weather prediction (NWP) models to demonstrate its benefit for weather forecast (Stoffelen et al., 2006; ESA, 2008, 2023). This has meanwhile be done at several meteorological centers (Rennie et al., 2021; Rani et al., 2022; Martin et al., 2022) and a clear positive impact on forecast skills has been reported (ECMWF, 2019a, b).

Given the novelty, extensive validation efforts (Calibration/Validation - Cal/Val) have been needed to verify the observations, detect unforeseen challenges (instrument and processing wise) and develop respective correction or calibration updates in order to make such a data assimilation within near-real time (less than 3 hours) possible at all. For this reason, an intense feedback from Cal/Val teams was desired and obtained to work together with the Aeolus DISC (Data, Innovation, and Science Cluster) and ESA itself on the improvement and stability of instrument and products.

Since the launch in 2018, several challenges were identified by DISC and ESA according to the feedback from the Cal/Val teams (e.g., Krisch et al., 2020). Some important issues are listed in the following:

- ~~Initially lower~~ Lower laser energy with a more rapid decline than expected (Simonelli et al., 2019; Reitebuch et al., 2020; Lux et al., 2020b),

- Switch to second laser with different beam characteristics which were also changing over time (Straume et al., 2020),

- Occurrence of increased background noise for some pixels (hot pixels) on the ACCD (Accumulation Charge-Coupled Devices) of ALADIN (Weiler et al., 2021a),

- Changes in the wind accuracy according to differences in temperature at the main telescope mirror of ALADIN (Weiler et al., 2021b).

DISC and ESA have worked hard on these features to improve the stability of the instrument and its products ~~. This has been which is~~ a prerequisite for the direct assimilation. As the above-mentioned issues influence the use, e.g., the assimilation, of the Aeolus data, continuous and long-lasting validation becomes very important. Most of the operational validation of Aeolus products was performed with NWP models (using of course also assimilated measurements, Chen et al., 2021; Hagelin et al., 2021; Martin et al., 2021; Liu et al., 2022; Zuo et al., 2022) while a direct validation with dedicated measurements ~~have~~ has been rare or ~~rather covering~~ covered only a short period and usually only a certain geographic region (Baars et al., 2020a; Witschas et al., 2020; Lux et al., 2020a; Baars et al., 2020b; Chen et al., 2021; Martin et al., 2021; Belova et al., 2021; Iwai et al., 2021; Zuo et al., 2022; Wu et al., 2022; Geiß et al., 2022; Lux et al., 2022b; Witschas et al., 2022).

The Leibniz Institute for Tropospheric Research (TROPOS) performed direct, long-term Aeolus-dedicated measurements of the wind vector at two distinct locations in the framework of the cooperation project EVAA (Experimental Validation and Assimilation of Aeolus observations, Baars et al., 2020a) between the Ludwig-Maximilians-Universität of Munich, the German Aerospace Center (DLR), the German ~~Weather~~ Meteorological Service (DWD) and TROPOS. In Leipzig, Germany (~~51~~51.35°~~N, 12~~ N, 12.43° E), dedicated radiosondes ~~were~~ have been launched for the weekly overpass on Aeolus' ascending orbit since May 2019. We analysed this radiosonde data covering the period from June 2019 ~~, while in~~ until September 2022, thus the mission period for which the second laser (FM-B) was in operation. In Punta Arenas, Chile (~~53.1~~53.15°~~S, 70.9~~ S, 70.91° W), continuous remote-sensing observations of LACROS (Radenz et al., 2021) served as one of the very rare Southern Hemispheric Aeolus validation ~~measurements~~ sites (Ratynski et al., 2022; Zuo et al., 2022). Namely, scanning Doppler cloud radar data, which have been collected in the framework of the DACAPO-PESO campaign (Radenz et al., 2021) ~~since~~ from November 2018 ~~,~~ until November 2021 were used for validation activities ~~besides occasional radiosonde launches~~ covering the ascending and descending orbit. In addition, occasional radiosondes were launched at the local airport, but due to the low amount, we have used this data only for case study validation of both, Aeolus and the ground-based cloud radar observations, but not for the statistical validation approach..

In the work presented here, we ~~want to assess~~ assessed the performance of Aeolus and its wind products ~~for the entire period of its nominal lifetime , i.e. 3 years,~~ (both Rayleigh clear and Mie cloudy) for a large part of its lifetime by using our long-term reference measurements. We also ~~intend to evaluate~~ evaluated the potential improvements of the wind products by the introduction of new algorithm versions (baselines) into the operational retrieval chain. We thus ~~aim~~ aimed for analysing the overall
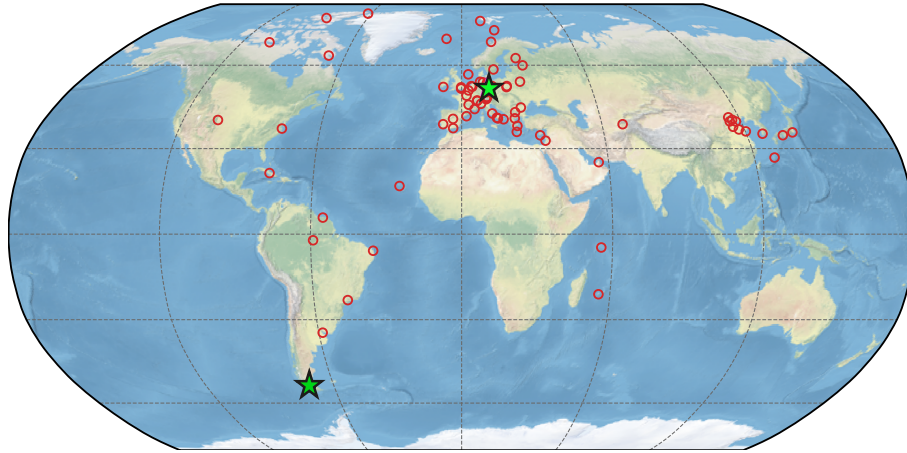
3

Figure 1. Map of Aeolus Cal/Val stations. The ones considered in this work are highlighted as ~~green yellow~~ green stars.

performance of Aeolus and the improvements introduced by new processor versions and calibration strategies based on two single sites located in the ~~Northern and Southern Hemispheric~~ northern and southern mid-latitudes.

The paper is structured as follows: In Section 2, the experimental setup including the campaign locations and instrumenta-

95 tion is described. In Section 3, we explain the methodology applied to derive horizontal wind from the ground-based ~~remote sensing~~ remote-sensing instrument and our validation strategy with respect to Aeolus. Selected case studies are discussed in ~~Sec.~~Section 4 to explain the methodology and show already some potentials and limitations of Aeolus. Finally, long-term statistics ~~taking also~~ are presented and discussed with respect to the different algorithm versions ~~of Aeolus into account, are presented and discussed concerning the Aeolus performance in Sec. 5~~ and the overall performance of Aeolus during its lifetime

100 in Section 5 and 6. Last but not least, the drawn conclusions are presented.

2 Experimental setup

2.1 Campaign locations

Measurements at two mid latitudinal locations have been used for the validation activities described here: Leipzig, Germany and Punta Arenas, Chile. Their locations are shown in Fig. 1 together with other ground-based stations which contribute to the

105 validation of Aeolus.

2.1.1 Punta Arenas, Chile

The remote-sensing supersite LACROS (Leipzig Aerosol and Cloud Remote Observations System) has been operated at Punta Arenas, Chile (~~53.14~~53.15°S, ~~70.88~~ S, 70.91° W) from November 2018 to November 2021 for the DACAPO-PESO campaign ~~(Radenz et al., 2021)~~

4

110 (Dynamics, Aerosol, Cloud And Precipitation Observations in the Pristine Environment of the Southern Ocean, Radenz et al., 2021)

. Thereby, one of the first multi-year ground-based remote sensing data ~~set in the Southern~~ sets in the southern mid-latitudes was obtained. The LACROS instrumentation comprises a PollyXT ~~Raman-polarization-lidar~~ Raman-polarization lidar (Engelmann et al., 2016; Baars et al., 2016), a CHM15kx ceilometer, a MIRA-35 scanning cloud Doppler radar (Görsdorf et al., 2015), a HATPRO microwave radiometer, and a Streamline Doppler lidar. Additionally, radiosondes (Lockheed Martin LMS6)

115 could be launched at the airport of Punta Arenas for dedicated objectives.

Punta Arenas is an ideal location for the validation of Aeolus in terms of wind conditions~~:~~. A strong circumpolar flow is a characteristic feature of the Southern Ocean with the southern tip of South America being the only barrier in the latitude band from 47° S to 63°~~S. Low pressure~~ S. Low-pressure systems embedded in this flow usually pass through the Drake passage south of Punta Arenas causing prevailing wind directions between south-west and north-west. A comprehensive description of

120 the meteorological conditions is provided in Radenz et al. (2021).

Aeolus overpasses considered for the validation were on ~~Wednesday~~Wednesdays, ca. 23:26 UTC on the ascending orbit, and on Thursdays, at around 09:56 UTC on the descending orbit. For the presented study, the scanning Doppler cloud radar has been ~~primarily used~~ used for the long-term validation and is thus explained in more detail in Section 2.2. The Doppler lidar performed scans for the horizontal wind as well, but due to the very low amount of particles in Punta Arenas, the performance

125 during the scans was not optimum for the Aeolus validation. Mostly, wind retrievals were restricted to the local boundary layer. But due to the relatively long distance to the Aeolus ground track (often more than 50 km) and the complex orography, the data were not useful for the Aeolus validation.

2.1.2 Leipzig, Germany

At the ACTRIS (Aerosol, Clouds and Trace Gases Research Infrastructure) site of Leipzig, Germany (51.35° N, 12.43° E),

130 Aeolus Cal/Val activities were focused on dedicated radiosonde launches (see Sec. 2.2.2). These launches took place for the ascending orbit of Aeolus on Friday evening at around 16:50 UTC.

Leipzig is located in central Europe being in the intermediate state between maritime and continental climate. Prevailing winds are usually westerlies, but due to wave activities winds from all ~~direction~~ directions can be observed. Leipzig is located in the low-land area. No orographical obstacles are around the city, making it a perfect location for the validation of Aeolus.

135 ## 2.2 Instrumentation

2.2.1 Scanning Doppler cloud radar

~~Continous~~ Continuous measurements were conducted with a 35 GHz Doppler cloud radar of type Metek MIRA35 (Görsdorf et al., 2015). Once per hour, the stare mode (vertical profiling) was interrupted for a Range-Height-Indicator (RHI) and Plan Position Indicator (PPI, also called VAD - Variable Azimuth Display) scan ~~.~~ from minute 29 to 36 of each hour. The PPI scan

140 started around minute 35 and lasted for 60 seconds and covered one full 360° rotation made in 6° steps. Only the PPI scans are considered for the horizontal wind retrieval~~.~~, which were performed at an elevation angle of $\varepsilon = 85°$. A pulse repetition

frequency of 5000 Hz gives a maximum unambiguous radial velocity of 10.56 ~~m/s~~ ms^{-1} , while the range resolution of 31.17 m is determined by the pulse length of 208 ns. Frequent cloud occurrence over Punta Arenas makes this instrument a perfect tool for retrievals of horizontal wind profiles, particularly during austral winter (Seifert et al., 2020; Radenz et al., 2021). The methodology for retrieving wind information from scanning Doppler ~~remote sensing~~ remote-sensing instruments is described in more detail in Sec. 3.1.

2.2.2 Radiosonde

Radiosondes of type Vaisala RS41 (Jauhiainen et al., 2014; Jensen et al., 2016) were launched at Leipzig each Friday for the regular Aeolus evening overpass (on its ascending orbit) since May 2019. The launch time of the radiosondes has been at 16 UTC, thus ca. 50 minutes before the Aeolus overpass, to have a good coverage of the atmospheric conditions up to about 25 km . Usually, the complete ascent up to the burst height is about 2 hours and thus perfectly centered around the overpass time to have the best temporal coverage as possible. We therefore assume that the horizontal drift of the radiosonde does not introduce a systematic bias to our statistical validation analysis. The RS41 delivers profiles of temperature, humidity, pressure, wind speed and direction. The uncertainty for the wind products is estimated to be between 0.4 and 1 ~~m/s~~ ms^{-1} for the wind velocity and 1° for the wind direction based on calculations of the Global Climate Observing System Reference Upper-Air Network (GRUAN, Dirksen et al., 2014). Even though these estimations are based on Vaisala radiosonde type RS92, there is no significant difference in the uncertainty between both radiosonde types as they are based on the same technique to derive wind velocity and direction (Jensen et al., 2016). A bigger gap in coverage occurred during winter 2020/2021 with only sporadic radiosondes (the reason was local access restriction due to COVID-19), but in total more than 125 launches could be completed. These radiosonde profiles were not assimilated so that they can serve as an independent reference for Aeolus products.

For Aeolus overpasses ~~in~~ at Punta Arenas, dedicated radiosondes were launched irregularly. The radiosonde type deployed in Punta Arenas was Lockheed Martin LMS6 and delivered ~~next to~~ also profiles of temperature, humidity ~~and pressureprofiles also profiles of~~ , pressure, wind speed and direction. In total, 41 radiosondes were launched during the 3 years campaign. Due to the irregularity of the launches, the radiosonde data was used only for case study analysis and thus also mainly for the validation of the cloud-radar-derived winds.

3 Methodology

 The methodology how Aeolus retrieves the HLOS is described in many other publications, e.g. in detail in Krisch et al. (2022); Rennie et al. (2021) but also in Baars et al. (2020b); Weiler et al. (2021b); Chou et al. (2022); Witschas et al. (2022); Bley et al. (2022) and references therein. However, the methodology how to retrieve horizontal wind vector from ground-based Doppler instruments, like the Doppler cloud radar at Punta Arenas, is not straight forward as several methodologies exists. Here, we give a short overview, which
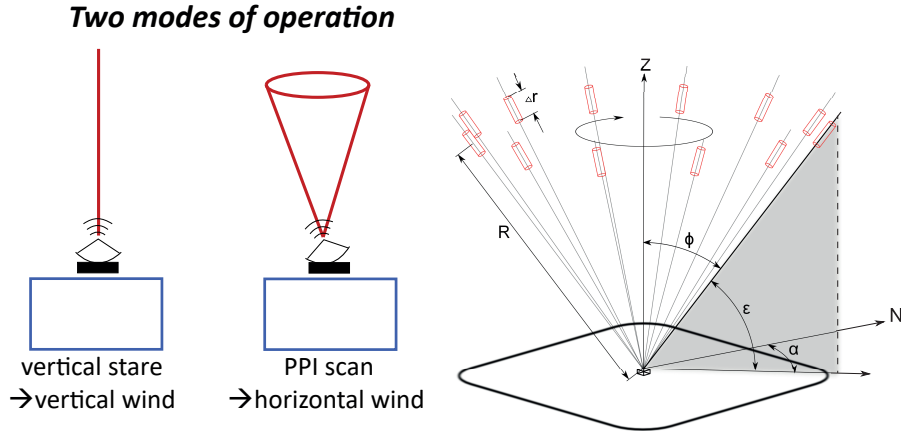
Figure 2. Left: Sketch of the different operating modes of the Doppler cloud radar. Stare mode is used for retrieval of vertical wind speed, while PPI/VAD scans are used to retrieve profiles of horizontal wind speed and direction. Right: Scanning geometry and nomenclature for the PPI/VAD scans as used in this work. The sketch was presented in Päschke et al. (2015) under Creative Commons Attribution 3.0 License and is shown here with permission of the authors.

methods are used and, thus, how the HLOS is retrieved from the cloud radar (Section 3.1) and about the general validation strategy for Aeolus (Section 3.2).

3.1 Retrieving horizontal wind profiles from radar

While the ~~remote sensing~~ remote-sensing instruments of TROPOS in Punta Arenas are usually operating ~~at~~ in stare mode (vertical profiling), regular PPI scans have been performed with the Doppler cloud radar to obtain the horizontal wind vector. For these scans, ~~measurements are taken by rotating in~~ the radar is rotated around azimuth α ~~through a full circle~~ with a fixed elevation angle ~~ϵ~~ ε (which is set to $85°$). A sketch showing the ~~scan patterns~~ different measurement modes is provided in Fig. 2, left.

The measured line-of-sight (LOS) Doppler velocity v_{LOS} at the range R and azimuth angle α is retrieved as the mean of the measured Doppler velocities for a given range band Δr. The Root Mean Square (RMS) of this distribution is used to calculate the uncertainty of v_{LOS}. In the following, we use the notation ~~which~~ that is given in Päschke et al. (2015)~~and~~, which is shown in Fig. 2, right, and neglect that all variables are a function of the range R to allow better reading. The final result is of course depending on R and thus gives a vertical profile.

In line with the standard approach of deriving horizontal wind from PPI scans ~~(Browning and Wexler, 1968)~~ (Browning and Wexler, 1968; Päschke et al., 2015), the mean horizontal wind speed $v_{\mathrm{advection}}$ can be approximated by fitting the measured v_{LOS} with a trigonometric function of the azimuthal coordinate of the scan corrected for positioning errors ($\alpha_{\mathrm{corrected}}$):

$$v_{\mathrm{projection}}(\alpha_{\mathrm{corrected}}) = v_{\mathrm{advection}} \cdot \cos(\alpha_{\mathrm{corrected}} - \alpha_{\mathrm{wind}}) + B + \sigma. \tag{1}$$

7

This formula gives the horizontal wind direction α_{wind} and the horizontal wind speed $v_{\text{advection}}$, with $v_{\text{projection}}$ being the horizontal component projected from v_{LOS}:

$$v_{\text{projection}} = \frac{v_{\text{LOS}}}{\cos(\epsilon)} \frac{v_{\text{LOS}}}{\cos(\varepsilon)}. \tag{2}$$

The term σ in Eq. ~~1~~ (1) reflects the remaining variation and is aimed to be minimized. For the fit procedure, $v_{\text{advection}}$, α_{wind}, and B are dependent variables, chosen to minimise σ ~~, reflecting~~ which is the remaining residual. The extra term B ~~reflects~~ stands for the contribution of two factors to the measured Doppler velocity: The divergence in the wind field and the vertical component of the average wind velocity. Both effects are neglected within the following analysis, as it is also done for the Aeolus HLOS retrieval.

Three different fit methodologies are used to derive the horizontal wind vector. The first one is equivalent to the methodology described in Päschke et al. (2015):

1. A least square regression is ~~used~~ applied to fit $v_{\text{projection}}$ and $\alpha_{\text{corrected}}$ considering also their uncertainties. ~~The fit procedure results in a covariance matrix which is used to calculate the uncertainty.~~ This method is subject to ~~a flaw of Doppler folding by the nature of the measurement process~~ Doppler folding (Ray and Ziegler, 1977). This means, v_{LOS} that exceeds the Nyquist velocity will appear as smaller measured velocity in the opposite direction. Usually, this effect will result in a poor fit quality with a high residual, as the measured $v_{\text{LOS}}(\alpha_{\text{corrected}})$ and thus $v_{\text{projection}}(\alpha_{\text{corrected}})$ will not approximate a trigonometric function, and can thus be discarded.

The other two fit methods that are ~~used~~ applied to retrieve wind velocities from the raw Doppler velocity data are based on the method by Tabary et al. (2001), which uses the approximation of an azimuthal derivative of the velocity distribution. This method is performed in two different ways~~.~~:

2. The ~~usual way, as recommended in Tabary et al. (2001), is to approximate the gradient~~ horizontal wind vector profile is retrieved from the gradient $\partial v_{\text{LOS}}/\partial \alpha_{\text{corrected}}$ which is approximated by overlapping piece-wise linear fits centred around each point of the initial distribution ~~. This approach allows to estimate $\partial v_{\text{LOS}}/\partial \alpha_{\text{corrected}}$ which is used to estimate the horizontal wind velocity and direction. This~~ as recommended by Tabary et al. (2001). This procedure is the second method and is usually consistent with the first method, but may lead to higher standard deviations due to removal of data points and the extra stages of calculation. Conversely, when Doppler folding occurs, this method is able to fit to a transformed version of the data with much higher accuracy.

3. The third method is applied because processing large numbers of linear fits as for the second method can sometimes be numerical unstable. This backup method ~~is using~~ applies the direct differences between consecutive values divided by the azimuthal distance. This approach is consistent with the former ones but leads to correspondingly higher errors because it excludes the averaging that occurs with the linear fit procedure. ~~But on~~ On the other hand, if the previous method fails to converge, this (third) Doppler-folding-safe methodology can be ~~used~~ applied to derive the horizontal wind vector.

All three methods are performed for each range R ~~. In the final data set~~to calculate the horizontal wind vector. In a final step, a best estimate is ~~then computed~~computed, which selects the method with the lowest error ~~.~~out of the three methods. In the data set, the retrieval results from all three methods plus the best estimate is stored. This best estimate ~~it~~is then used for the comparison with the Aeolus winds~~.~~, however not considering cloud-radar-derived HLOS winds with an error higher than $10\,\mathrm{ms}^{-1}$.

3.2 Aeolus validation strategy

For the validation of Aeolus, we focus on the L2B wind products ~~, we use~~obtained by the Rayleigh methodology in clear air, called Rayleigh clear winds, and with the Mie methodology in clouds, called Mie cloudy winds. A more thorough description of the different products can be found, e.g., in the product description document (de Kloe et al., 2023) or in Baars et al. (2020b). If not otherwise stated, from now on we use the term "Rayleigh" for the Rayleigh clear wind products and the term "Mie" for the Mie cloudy wind products not considering the theoretically available Rayleigh cloudy and Mie clear winds. Rayleigh winds are delivered at 87 km horizontal resolution, while the Mie wind resolution has been mainly at 15 km (setting flexible, see Table 1).

We analyse all Aeolus-derived Horizontal Line-of-Sight (HLOS) wind speeds (i.e., at different altitudes) ~~which were assigned to coordinates which~~from the Rayleigh and Mie products according to their mean coordinates (in the center of the horizontal width) that are within a radius of 100 ~~km~~km[1] around the measurement site. The used radius is a good compromise between the number of available comparison data and the representativeness between the two different measurements as shown in other studies (Geiß et al., 2019; Cossu et al., 2022). Accordingly, two overpasses per week ~~fulfilling these conditions were~~for each station fulfil these conditions and have been suitable for validation:

– Punta Arenas: Wednesdays ~~around~~at 23:26 UTC and Thursday ~~around~~at 09:56 UTC,

– Leipzig: Fridays ~~around~~at 16:50 UTC and Sundays ~~around~~at 05:29 UTC.

The ~~validation is done for the so-called Aeolus L2B "Rayleigh clear" wind~~performed orbit shift has not changed the considered overpasses for the stations, but the mean distance from the ground site to the Aeolus wind curtains has changed. Before the orbit shift, this mean distance was 17 km and 42 km on Fridays and Sundays at Leipzig, and 27 km and ~~"Mie cloudy" wind products. A detailed description of these wind types is given in, e.g., de Kloe et al. (2023) and Baars et al. (2020b). Rayleigh clear winds are delivered at 87 km horizontal resolution while the Mie wind resolution has been mainly at 15 km (setting is flexible).~~75 km on Wednesday and Thursdays at Punta Arenas, respectively[2]. After the orbit shift, it changed to 33 km and 93 km at Leipzig, and 75 km and 85 km at Punta Arenas, respectively. As the orbit of Aeolus is slightly varying and the distances given hear are only mean values, some of the overpasses at Punta Arenas were outside the 100 km criterion so that we adjusted the radius to 120 km. For Leipzig this was not needed, as only Friday overpasses were considered.

For a better understanding of the general procedure, an example of the Aeolus Rayleigh wind profiles over Punta Arenas is

---

[1]120~~km~~ km in Punta Arenas after orbit shift in June 2021.

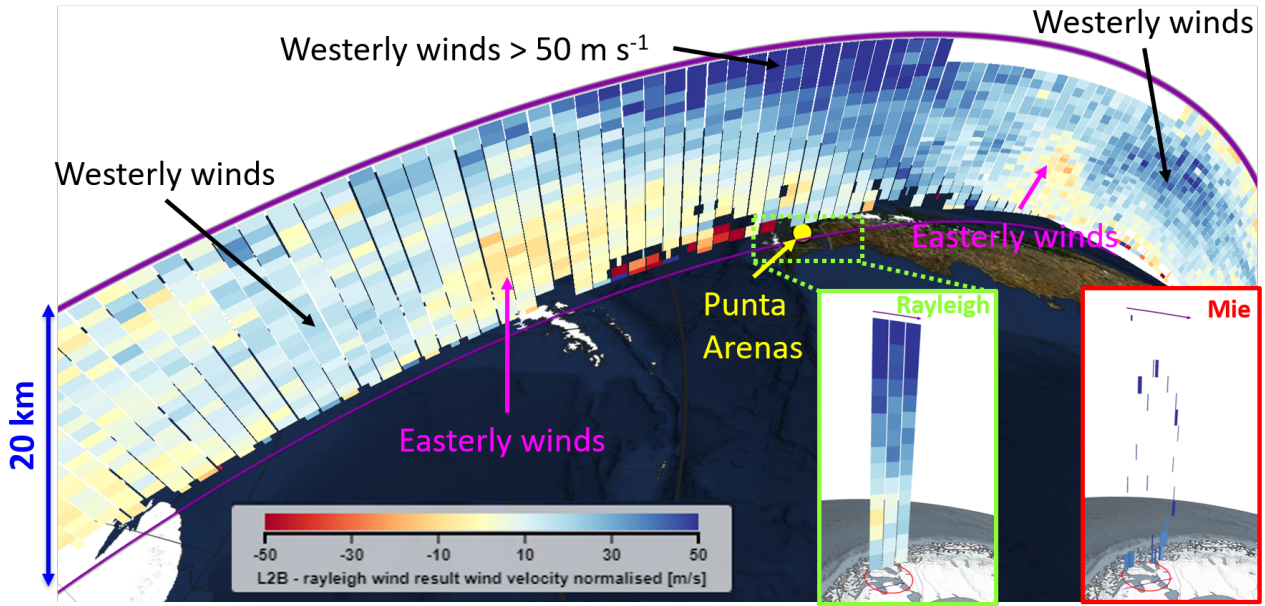[2]calculated by ESA based on Orbit Scenario Files

Figure 3. Example for L2B wind curtain around Punta Arenas, Chile, on 30 September 2020 visualized with VIRES (Santillan et al., 2019). The Rayleigh ~~clear~~ wind product is shown for the whole curtain. The Mie wind product (red box) and the Rayleigh wind ~~product~~ products (green box) 100 km around Punta Arenas ~~is~~ are shown in the lower right.

shown in Fig. 3. On 30 September 2020, strong westerly winds (blueish colors) occurred over Punta Arenas at altitudes above 5~~km~~ km. Closer to the South Pole, easterly winds (reddish colors) were prevailing. A patchy wind speed pattern was observed close to Punta Arenas near the ground, caused by cloud contamination of the Rayleigh winds. Given the example in Fig. 3, one sees that depending on the actual track of Aeolus, 1–3 wind profiles fulfil ~~this~~ the criterion of being within 100~~km~~ km radius of the observational site (see green box in Fig. 3). Considering 15~~km vertical~~ km horizontal resolution for the Mie product since 5 March 2019 (before the resolution was 87~~km~~ km), one can have up to 13–20 "Mie winds" for one altitude range ~~around the~~ within 100~~km~~ km of the ground-based location (see red box in Fig. 3). For the validation of Aeolus products, the ~~temporal closest observation of the~~ temporally closest ground-based ~~measurements~~ cloud radar measurement has been used ~~with allowing a maximum-difference~~ allowing for a maximum-time-difference threshold of 1 hour. This threshold shall ensure similar atmospheric conditions for the validation. For the radiosonde data, such a temporal constraint is not needed as the radiosondes have been launched directly for the Aeolus overpass.

Furthermore, we converted the wind speed v_{gb} ~~and direction φ_{gb}~~ v_{ref} and direction φ_{ref} obtained with the ~~ground-based (gb, i.e. cloud radar) instruments (see Sec. 3.1)~~ reference instruments (subscript ref, i.e., cloud radar and radiosonde) to Aeolus-like HLOS winds $v_{gb_{HLOS}}$ $v_{ref_{HLOS}}$ with the equation described in Baars et al. (2020b):

$$v_{gb_{HLOS}\,ref_{HLOS}} = v_{gb\,ref} \cdot \cos(\varphi_{Aeolus} - \varphi_{gb\,ref}). \tag{3}$$

10

Table 1. Overview of the different algorithm ~~version~~ versions (called baseline) for the processing of the Aeolus data together with some important additional information.

Baseline	Period	Start date for operational processing	Additional info
B02	Sep 2018 – May 2019	8 Sep 2018	Mie wind ~~res. from 87 km to 15 km~~horizontal resolution from 87 km to 15 km in March 2019
B03	May 2019 – June 2019	16 May 2019	
B04	June 2019	14 June 2019	Hot Pixel correction (Weiler et al., 2021a)
B05	June 2019 – September 2019	28 June 2019	Switch to ~~Laser~~ laser FM-B
B06	June 2019 – October 2019	5 September 2019	Adapted response calibration for FM-B
B07	October 2019 – April 2020	31 Oct 2019	
B08	April 2020	2 April 2020	Unit change of HLOS error from ms^{-1} to cm s^{-1}
B09	April 2020 – July 2020	20 April 2020	M1 temp. correction (Weiler et al., 2021b), public data release
~~B10*~~ B10	June 2019 – Oct 2020	2019 reprocessed since 9 July 2020 operational	reprocessed FM-B data set included in Baseline 10 covering June to December 2019
~~B11*~~ B11	June 2019 – May 2021	2019 reprocessed since 8 October 2020 operational	reprocessed FM-B data set included in B11 covering the period June 2019 to 10 October 2020
B12	May 2021 – Dec 2021	26 May 2021	Orbit shift in June 2021
B13	Dec 2021 – Mar 2022	6 December 2021	
B14	Mar 2022 – Sep 2022	29 March 2022	

φ_{Aeolus} is the azimuth angle of Aeolus, which is obtained from the Level 2B data and differs depending on global position. The uncertainties of the ground-based observations were propagated forward. The derived ground-based profiles of HLOS wind were then vertically averaged to the Aeolus-range bin thickness (500 to 2000 m, mostly higher resolution near ground and coarse resolution at high altitudes: Bley et al., 2022; ESA, 2020) to allow a one-to-one comparison. ~~By doing so~~This means, we do not aim to discuss here the small-scale wind variability within the relatively large Aeolus range bins but rather concentrate on performance of the space-borne instrument.

During the lifetime of Aeolus, several algorithm versions of the processing chain (so-called baselines) were released and applied~~. Some~~, some of them in operational mode, some of them to reprocess parts of historical Aeolus data. Thus, for certain dates in the Aeolus data set, several versions exist (processed with different baselines), while for other periods only one baseline was applied. An overview of the different baselines of Aeolus covering the observational period of our ground-based reference measurements (i.e., up to ~~summer~~ autumn 2022) is given in Table 1. Two major steps for boosting the performance of Aeolus were made. With Baseline 04, the so-called hot pixel correction (Weiler et al., 2021a) was introduced. Before that, single pixels on the ACCD of Aeolus had a higher dark current and thus biased the retrieved winds. A second important step

11

was the introduction of a correction with respect to changes in the telescope temperature of Aeolus (M1 temperature correction, Weiler et al., 2021b). This correction was implemented with Baseline 09 and should have brought a significant improvement of the performance of the Aeolus winds.

The switch from laser FM-A to laser FM-B was performed from 12 June 2019 until 28 June 2019 and led to Baseline 05. However, a new response calibration needed to be applied, which was obtained in August 2019 and led to Baseline 06. The FM-B data before that date was then reprocessed. In June 2021, the orbit of Aeolus was shifted to favor the ground-based observations in Cabo Verde during the Joint Aeolus Tropical Atlantic Campaign (JATAC, Fehr et al., 2021). Therefore, mean horizontal distances of the Aeolus wind products to the ground-based reference stations changed and as a result we increased the maximum radius for Punta Arenas to 120 km to still be able to obtain two overpasses per week (as already discussed above).

Accounting for changes in units for the uncertainties within the Aeolus products between different baselines, all values for Aeolus horizontal wind speed and errors were transformed into ms^{-1}. Beside this unit correction, all baselines were treated equally. Furthermore, next to the provided validity flag within the Aeolus wind products, additional quality measures, i.e., error thresholds for Mie and Rayleigh winds (5 ms^{-1} and 8 ms^{-1}, respectively), have been applied. This means that wind products flagged valid but with an error higher than these thresholds were discarded. This approach is consistent with DISC/ESA recommendations (e.g., Witschas et al., 2020; Rennie and Isaksen, 2020) and studies by other Cal/Val teams (e.g., Guo et al., 2021; Chen et al., 2021; Iwai et al., 2021; Abril-Gago et al., 2022b).

For the statistical analysis presented in Section 5 and 6, the Rayleigh and Mie wind products were treated separately. To obtain statistical metrics, a straight-line fit between the ground-reference winds and the Aeolus winds using a orthogonal distance regression (ODR) to include the effects of errors has been computed. We also created histograms of the deviations (reference wind minus Aeolus wind) in the range of -15 to +15 ms^{-1} (with higher velocities being assigned to the outside bins) to check for a Gaussian distribution shape. A walk-through example for the statistical analysis is given in Section 5, after the general validation strategy based on case studies is discussed next.

4 Case studies

To illustrate the validation strategy and discuss the potentials and drawbacks, we present two interesting case studies performed for Punta Arenas in the following.

4.1 Punta Arenas - 6 February 2020

A schematic overview on how winds are retrieved from the ground-based observations and then compared to Aeolus wind products is shown in Fig. 4 for the case of 6 February 2020, representative for the Southern Hemispheric summer. The atmospheric conditions above Punta Arenas on this day are presented in Fig. 4, top, left, by means of the Cloudnet Target Catego-
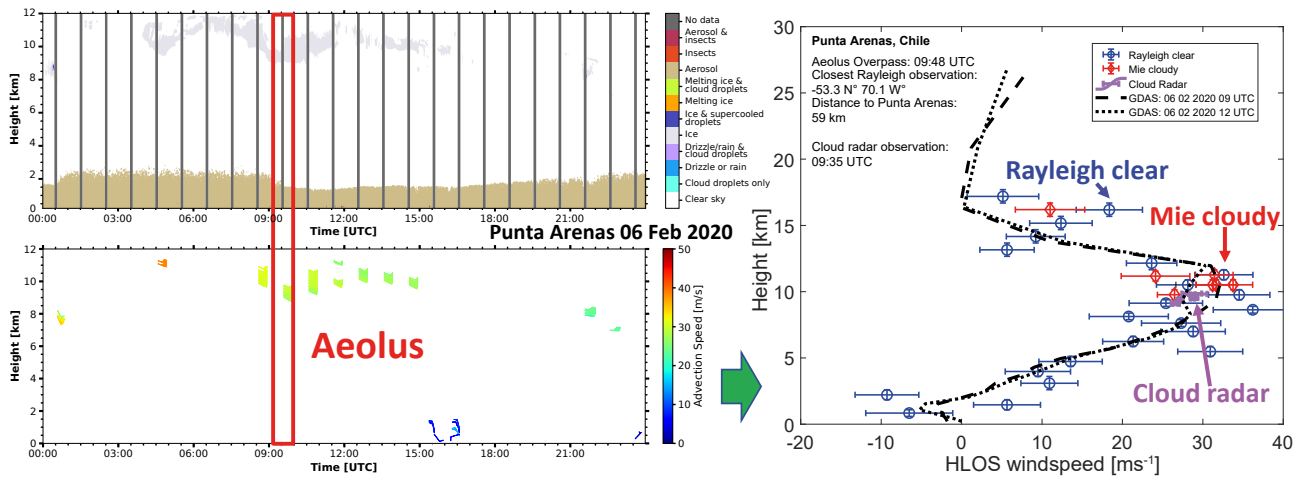
Figure 4. Schematic overview on the methodology used in this ~~paper~~ study for the example of 6 February 2020 in Punta Arenas~~: How are winds retrieved from the Doppler cloud radar and finally compared to Aeolus~~. ~~Left~~Top, left: Cloudnet Target ~~categorization~~ Categorization obtained from the combination of ~~vertical~~ vertically measuring (in stare mode) cloud radar, ~~Ceilometer~~ ceilometer and microwave radiometer. ~~Center~~Bottom, left: ~~Resulting winds~~ Wind speed and direction indicated by arrows as retrieved with the Doppler cloud radar scans at minute 35 of each full hour in regions of cloud occurrence. Right: Comparison of the HLOS winds from Doppler cloud radar to Aeolus products (Mie ~~cloudy~~ and Rayleigh~~clear~~, Baseline 11) for the closest overpass. GDAS model winds are shown ~~as well~~ for comparison as well.

rization ~~(Illingworth et al., 2007; Tukiainen et al., 2020)~~ (Illingworth et al., 2007; Tukiainen et al., 2020; Radenz et al., 2022) derived from the ~~vertical~~ vertically starring active remote sensing instrumentation (~~Cloud radar and Ceilometer~~cloud radar and ceilometer) and the passive microwave radiometer. During this day, a nearly cloud-free aerosol layer from ground up to 1.5~~km~~ km altitude was observed with enough particles to be identified by Cloudnet Target Categorization (dark yellow). Partly, the cloud radar observed a return signal within this aerosol layer, which is attributed to insects (~~red colors~~not shown). Between 2 and 8~~km, clear sky~~ km, clear-sky conditions (white color) were found, while ice clouds (dark grey) occurred sporadically above 8~~km altitude.~~ km altitude. ~~Vertical profiles of the horizontal wind vector from~~ Horizontal wind vector observations retrieved with the Doppler cloud radar from the hourly scans at minute 35 of each hour were therefore available ~~when~~ in atmospheric regions where clouds were existing~~,~~ see Fig. 4, ~~center.~~lower, left.

The Aeolus overpass on this day was at 09:48 UTC. Thus, the temporally closest wind profile ~~of~~from the Doppler Cloud radar (at 09:35 UTC) plus the HLOS profiles extracted from GDAS1[3] data for 9 UTC and 12 UTC were used to compare with the Aeolus products (Mie ~~cloudy and Rayleighclear~~and Rayleigh) as can be seen in Fig. 4, right. The closest distance between the Aeolus ground track and the ground site was 59 km on this day.

In this example comparison, the advantages and drawbacks of the used reference instrument becomes clear. The cloud radar is only able to retrieve winds in regions were clouds are existent (on this day between 8 and 12~~km~~ km) and no information

---

[3]Global Data Assimilation System (GDAS), ARL Archive: GDAS1 data set, available at: https://www.ready.noaa.gov/gdas1.php

can be obtained in clear-sky regions. In regions of clouds, however, the winds can be obtained with high vertical resolution and high quality. The GDAS-derived vertical profile of HLOS is available for all atmospheric states (independent of cloud occurrence, clear sky etc.) but in coarser resolution. As GDAS data are a result of data assimilation, it is no direct validation measure and therefore shown only for consistency checks and not for Aeolus validation itself.

According to Fig. 4, right, the cloud-radar-derived HLOS wind taken at 09:35 UTC provides values of 27 to 31 ms^{-1} in the cloudy region around 10 km. GDAS HLOS wind speeds in this altitude region change from about 33 ms^{-1} to 28 ms^{-1} from 9 to 12 UTC, respectively. Due to the good agreement of the both data sources with respect to the HLOS, the validity of the cloud-radar derived winds can be assumed. Aeolus-derived wind profiles within a radius of 100 km were available in clear air (Rayleigh winds) and at top of clouds (Mie winds). The Mie winds available at 10 to 12 km indicate, thus, the presence of clouds at similar altitudes at the Aeolus track (more than 59 km away from the ground site). If the cloud deck would be persistent and optically thick over the whole horizontal Aeolus track, no Aeolus winds would be available below the altitude of the clouds due to the strong light attenuation within the cloud. As this is not the case because Rayleigh winds are available down to the surface, a broken cloud deck and/or optically thin clouds between 10 and 12 km with clear sky below and above can be considered in the Aeolus observation - which is in excellent agreement with the atmospheric scene observed over Punta Arenas by LACROS (Fig. 4, top, left).

On this specific day in austral summer 2020, a good agreement between the Mie winds and the cloud-radar-derived winds were obtained at an altitude of around 10 km. Also, the delivered Rayleigh winds in this altitude region agree well with the radar and also with GDAS. The coexistence of Rayleigh and Mie winds in one altitude range is possible because of the broken cloud deck Rayleigh and Mie winds can coexist (Reitebuch et al., 2018; Rennie et al., 2020; de Kloe et al., 2023) and all single Aeolus products within the defined horizontal radius of 100 km are considered. In contrast, at 17 km height, GDAS and Aeolus disagree for the only one Mie observation there. The reason for that is yet unclear but might be related to atmospheric heterogeneity, uncertainties of GDAS (which are not provided) or a misclassification of Aeolus as 17 km is usually well above the local tropopause and thus no clouds are expected (and not seen in the ground observations). However, smoke occurrence from Australian bush fires at this altitude range cannot be ruled out (Ohneiser et al., 2020), but the ground based PollyXT lidar observations do not show a significant enhanced backscatter at this region later the day, when the clouds at around 10 km disappeared, which would explain a misclassification by Aeolus. Rayleigh winds show also deviations around 17 km towards higher HLOS wind speed in accordance to the Mie observations.

Below the cloud deck at 10 km, Rayleigh winds are partly matching the model data (GDAS), but with a tendency of higher Aeolus wind speeds down to around 5 km altitude. Sporadically, lower Rayleigh HLOS wind speeds were also observed. Deviations within the lowest 3 km might be caused by horizontal inhomogeneity within

the 100~~km~~ km radius around the ground-based station. For the statistics presented below in ~~Sec. 5~~Section 5 ff., we use wind derived with the Doppler cloud radar ~~derived winds~~ and compare them to the equivalent Aeolus HLOS winds. For the example case presented here, this means that a comparison to Mie ~~cloudy~~ winds is possible for the height range around about 11~~km~~ km, as this is the only region for which ~~cloud-radar derived~~ cloud-radar-derived winds and Aeolus Mie ~~cloudy winds do~~ winds coexist. Rayleigh ~~clear~~ wind comparisons can be done at the same height range (reference measurements from the cloud radar). The regions between the ground and 9~~km~~ km altitude and above 11~~km~~ km cannot be covered for the comparison due to the missing ground-based measurement data. Radiosondes could cover this gap, but they were launched at Punta Arenas only irregularly for the Aeolus validation so that a meaningful long-term validation is not possible. We did not aim at a ~~comparison~~ validation with model data, as this is done regularly at ECMWF (Rennie and Isaksen, 2020; Rennie et al., 2021) and by other validation activities (e.g., Martin et al., 2021; Liu et al., 2022; Chen et al., 2021; Hagelin et al., 2021; Rani et al., 2022). Instead, we concentrate on the direct measurements made ~~at groundand space~~from ground.

4.2 Punta Arenas - 18 August 2021

The second case study from Punta Arenas presents an observation from the Southern Hemispheric winter. At this day, beside the Doppler radar, also a local radiosonde launch was available.

~~Atmospheric~~ The atmospheric conditions on this day were remarkable ~~. While considerable normal~~ as presented in Figure 5. While typical HLOS wind speeds between 5 and 20~~m/s~~ ms^{-1} were observed in the troposphere, a steady increase in wind speed was observed above the tropopause (at ca. 9.5 km) leading to a maximum ~~measured wind speed~~ wind speed measured by the radiosonde of more than ~~360 km/h. This Aeolus-observed wind profile was also present in the GDAS data~~100 ms^{-1}. The Aeolus-derived wind speed profile (overpass at 23:27 UTC, with the closest distance of 47 km to Punta Arenas) is in agreement with the GDAS (21 and 00 UTC) and the radiosonde data (launched 23 UTC), and thus gives confidence of its reality. ~~The reason for such high wind speeds was a shift of the polar vortex toward South America. As a consequence, a high wind speed band was directly located above Punta Arenas. Its horizontal extent is quite broadly distributed~~ The horizontal extent of the strong winds around Punta Arenas ~~as can been~~ can be seen in Fig. 5~~. There,~~, bottom, right, which shows the Aeolus Rayleigh ~~clear~~ HLOS wind speed~~curtain is presented as separate plot in the bottom, right, demonstrating the extent of the high-speed wind band.~~. It demonstrates the potential of Aeolus to detect such features.

The comparison of Aeolus derived HLOS winds with the radiosonde, cloud radar (23:35 UTC) and GDAS is shown in Fig. 5 and reveals that the Rayleigh ~~clear~~ product above 10~~km fits perfectly to~~ km follows the observed radiosonde ~~winds~~ wind profile and GDAS products. A considerable deviation was observed only at the top most Aeolus range bin (around 24~~km~~ km). There, the wind speed measured by Aeolus was considerably lower compared to the radiosonde wind. The reason is not ~~clear~~yet clear, but might be simply due to the strong drift of the radiosonde ~~due to~~because of the high wind speed. Below the tropopause ~~,~~ ~~the agreement of the Aeolus Mie cloudy~~ at ca. 9.5 km , the Aeolus Mie winds and Rayleigh ~~clear winds~~ winds agreed mostly with the radiosonde~~was excellent as well~~. The GDAS data, however, ~~shows~~ show a significant variation at 8~~km~~ km between the two profiles at 21 UTC and 00 UTC. This behaviour implies that fast changes in HLOS, i.e., in wind speed and direction, have taken place in this atmospheric region. It furthermore gives confidence, that the 1-hour time window for validation is a
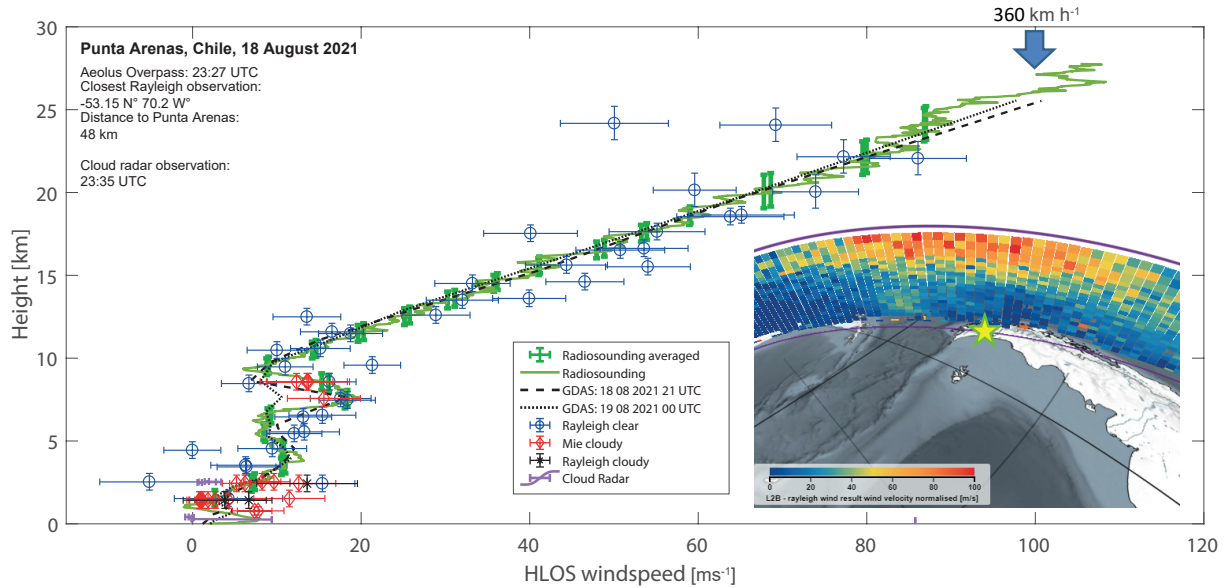
Figure 5. Comparison of ~~Radiosonde~~radiosonde, Doppler cloud radar~~wind~~, and GDAS HLOS winds to Aeolus products on 18 August 2021 ~~, thus representing~~ in southern hemispheric winter conditions. The Aeolus Rayleigh ~~clear~~ wind curtain for the ~~analysed~~ analyzed overpass as visualized with VIRES (Santillan et al., 2019) is shown in the lower right.

reasonable time period for the wind validation.

Usually, two ~~full~~ Aeolus Rayleigh wind observations (of 87~~km~~ km horizontal length in case of Rayleigh~~clear~~) per Aeolus height bin ~~lay~~ exist in the validation radius of 100~~km~~ km around the ground station when using the centre coordinates of the
400 wind products and are thus considered for the validation. Having a look at the HLOS observations at around 4.7~~km~~ km, one sees that one wind product of Aeolus fits very well to the reference wind profiles, while the other one shows considerable deviations (around 10~~m/s~~ ms^{-1} lower HLOS). This finding implies regional variations in the wind pattern. As a consequence, the observed outliers in the Aeolus Rayleigh ~~clear~~ winds at 3 and 4.7~~km~~ km can be attributed to horizontal (and thus also temporal) heterogeneity in the wind field. This behaviour shows the difficulty in comparing Aeolus HLOS winds to the ground-
405 based observations, because a perfect co-location in space and time can usually never be achieved. However, we ~~are confident~~ consider that these meteorological variations do not lead to additional biases in the statistics presented in ~~Sec. 5ff.~~ Section 5 ff., but are properly covered by the statistical methodologies in terms of random error.

5   Example for statistical validation: Baseline 11 at Leipzig and Punta Arenas

To obtain statistical measures for the performance of Aeolus and its algorithms, we ~~analysed~~ analyzed the Aeolus HLOS data
410 by Doppler cloud radar and radiosonde (from now on called reference instruments) as described above for the locations of Leipzig, Germany, and Punta Arenas, Chile. To illustrate that approach, the validation of Aeolus Baseline 11 products around
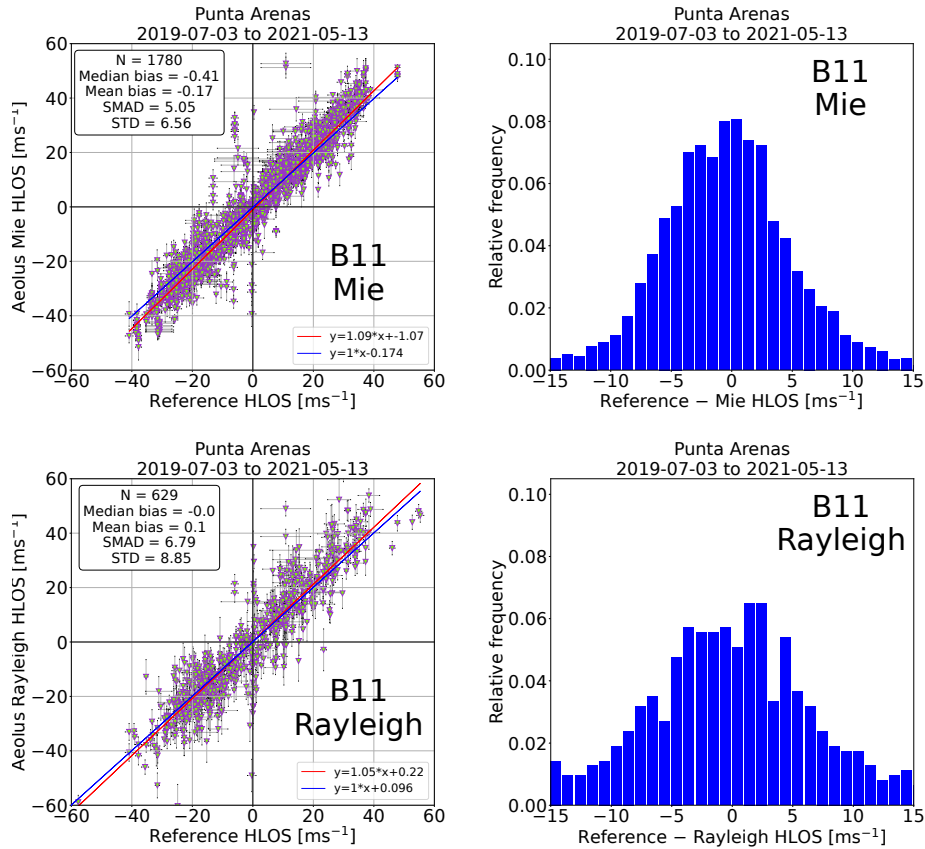
Figure 6. Long-term wind statistic for Baseline 11 for the location of Punta Arenas. Top: Aeolus Mie ~~Wind~~ wind vs. ~~Ground~~ ground Doppler ~~Observations~~ observations, ~~Bottom~~ bottom: Rayleigh ~~Wind~~ wind vs. ~~Ground~~ ground Doppler ~~Observations~~ observations, ~~Left~~ left: 1:1 statistic with respective measures (median and mean bias, ~~Right~~ scaled median absolute deviation (SMAD), and standard deviation (STD) provided in ms^{-1} , and N indicating the number of samples), right: ~~Frequency~~ frequency distribution of differences between the two data sources.

Punta Arenas are shown in Fig. 6. The ~~1:1 statistic~~ direct comparison (left column) and the frequency distribution of the deviation from the Doppler cloud radar ~~instrument~~ values (reference minus Aeolus, right column) are shown for the Mie (top row) and Rayleigh winds (bottom row). A ~~general~~ generally good agreement can be seen between Aeolus and the cloud radar

415 being most of the time close to the one-to-one line and thus justifying the use of the orthogonal distance regression (ODR) for fitting the data. ~~The uncertainties of the obtained slopes are thus low (<0.015 m/s for both, Rayleigh clear and Mie cloudy).~~

More data points could be evaluated for the Mie winds (in total ~~1817~~1780) than for the Rayleigh winds (~~642~~629), which is a logical consequence of the higher resolution of the Mie winds and the fact that the cloud radar ~~only derives winds~~ derives winds only in regions with cloud occurrence. Thus, the validation is more meaningful with respect to Mie winds.

420 For ~~that windtype~~the Mie wind, we obtained a slope of 1.1 ~~m/s~~ with the ODR. When forcing the slope to be unity, the resulting offset is equal to the mean bias as expected for a Gaussian distribution. A median bias of ~~-0.43 m/s~~ − 0.41 ms^{-1} and a mean

17

Table 2. Overview of the metrics obtained for the validation of Baseline 11 at Punta Arenas (left) and Leipzig (right).

(a) Aeolus vs. cloud radar at Punta Arenas

	Mie	Rayleigh
Number of points	1780	629
Slope	1.09	1.05
Median bias (ms^{-1})	−0.41	0
Mean bias (ms^{-1})	−0.17	0.1
Scaled MAD (ms^{-1})	5.05	6.79
Standard deviation (ms^{-1})	6.56	8.85

(b) Aeolus vs. radiosonde at Leipzig

	Mie	Rayleigh
Number of points	1751	2361
Slope	0.93	0.97
Median bias (ms^{-1})	−0.35	−0.46
Mean bias (ms^{-1})	0.03	−0.22
Scaled MAD (ms^{-1})	4.59	5.77
Standard deviation (ms^{-1})	5.34	6.75

bias of ~~-0.387 m/s~~ $-\,0.17$ ms^{-1} was derived (i.e., Aeolus measures less than the ground-based reference) together with a standard deviation of ~~7.2 m/s~~ 6.6 ms^{-1} and a scaled median absolute deviation (MAD) of 5.1 ~~m/s~~ ms^{-1} . For the Rayleigh wind validation, we obtained respective values of 1.05 ~~m/s~~ (slope), ~~-0.1 m/s~~ 0.1 ms^{-1} (mean bias), ~~-0.003 m/s~~ 0 ms^{-1} (median

425 bias), ~~9.1 m/s~~ 8.9 ms^{-1} (standard deviation) and ~~7 m/s~~ 6.8 ms^{-1} (scaled MAD)~~-~~, see statistics box in the left column of Fig. 6, bottom.

In the following, we use the scaled MAD as an indicator for the random error, in analogy to the median bias for the systematic error ~~also often~~ often just called bias. The median values are less sensible to outliers than the mean values, but are a valid measure for the uncertainties as long as the frequency distribution is of Gaussian shape. The philosophy of the use of the scaled

430 MAD for Aeolus comparisons is in detail explained in Martin et al. (2021), Lux et al. (2022a), and Weiler et al. (2021a).

We also performed a Z-score analysis as described in Lux et al. (2022b) and found that, for example, with a Z value of 3 for Baseline 10, 1.5 % of the Aeolus values are identified as outliers. However, in this publication we do not want to discuss the outliers of Aeolus wind products, but rather the performance of the publicly available wind data as a whole. Thus, we do not exclude outliers based on the Z-score analysis, but make a validation of the complete Aeolus data set ~~, given~~ by applying the

435 recommended error thresholds of 8 ~~m/s~~ ms^{-1} and 5 ~~m/s~~ ms^{-1} for Rayleigh and Mie winds, respectively.

The frequency distribution of the differences (Fig. 6, right) shows a ~~near Gaussian~~ near-Gaussian form giving confidence that the ~~statically~~ statistically measures described above ~~could~~ can be applied. The obtained systematic ~~error of -0.43 m/s and -0.09 m/s,~~ errors of $-\,0.41$ ms^{-1} and 0 ms^{-1} and the random errors of 5 ~~m/s~~ ms^{-1} and 7 ~~m/s~~ ms^{-1} for Mie and Rayleigh products, respectively, for Baseline 11 validation at Punta Arenas are in line with results from other validation activities for this

440 Baseline (Zuo et al., 2022; Geiß et al., 2022). An overview of the main key numbers from this statistic is given in ~~Tab.~~ Table 2.

We performed the same ~~statistics~~ statistical analysis with the radiosonde data from Leipzig for Baseline ~~11 - the~~ 11. The results are presented in Fig. 7. ~~Giving the fact ,~~ Given the fact that the radiosonde delivers wind data in both, clear and cloudy ~~sky~~ skies, it becomes clear that this reference instrument is well suited for the Aeolus Rayleigh and Mie wind validation. As radiosondes are not limited to certain atmospheric targets, a coverage up to 25 ~~km~~ km height could usually be achieved, allowing to validate

445 all HLOS winds during an Aeolus overpass. Therefore, results are not confined to the cloud-laden troposphere like in Punta
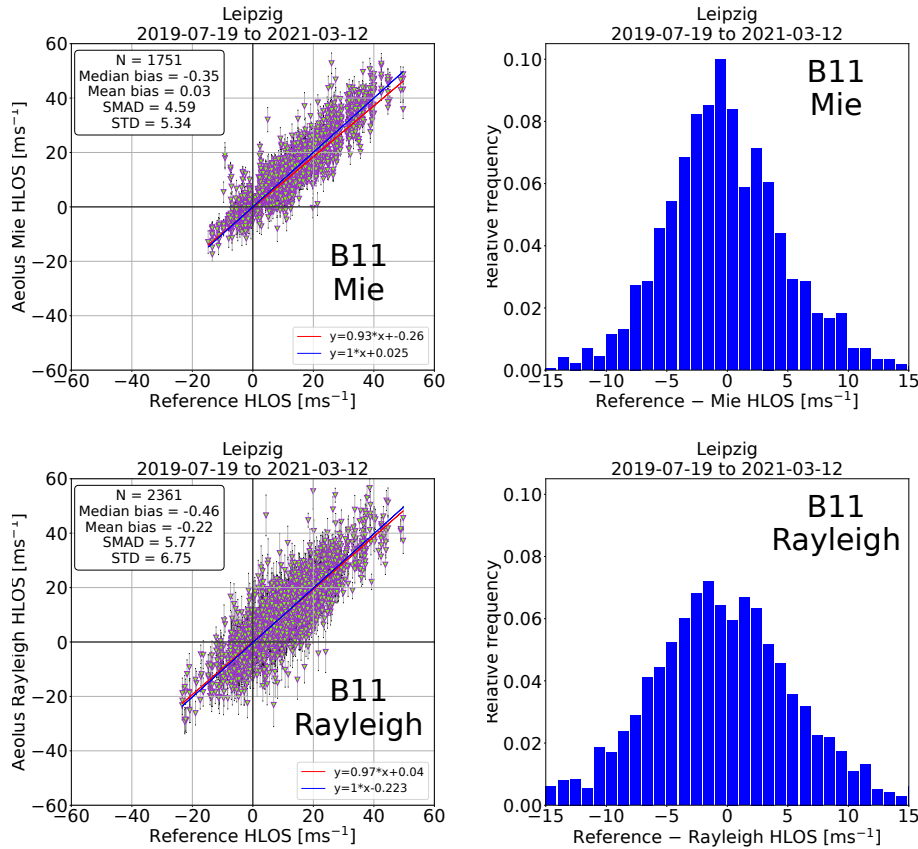
Figure 7. Long-term wind statistic for the Baseline 11 wind validation based on ~~radiosondes~~ radiosonde launches at Leipzig. Top: Aeolus Mie ~~Wind vs. Ground Doppler Observations~~wind, ~~Bottom~~bottom: Rayleigh ~~Wind vs. Ground Doppler Observations~~wind, ~~Left~~left: 1:1 statistic with respective measures (median and mean bias, ~~Right~~scaled median absolute deviation (SMAD), and standard deviation (STD) provided in ms^{-1} , and N indicating the number of samples), right: ~~Frequency~~ frequency distribution of differences between the two data sources.

Arenas. Thus, the statistical analysis is more ~~rich~~ representative in terms of data points for the Rayleigh and Mie wind ~~data points~~ as can be seen in Fig. 7, left. In total, more than 1500 and 2000 data points could be used for validating the Aeolus Mie and Rayleigh winds, respectively. The frequency distributions of the difference between the reference and Aeolus HLOS winds are of Gaussian shape for ~~, both ,~~ both Mie and Rayleigh winds and thus ~~gives~~ give again evidence for the validity of the applied

450   statistical validation approach. The direct comparison (Fig. 7, left column) shows a ~~general~~ generally good agreement with only sporadic outliers (e.g., ~~70 m/s measured by~~ $\approx 50\ \mathrm{ms}^{-1}$ in the Aeolus Rayleigh wind product while the radiosonde delivered ~~15 m/s~~less than $10\ \mathrm{ms}^{-1}$). The majority of data points are, however, near the ~~imaginary~~ 1:1 line and thus in good agreement. For the Mie winds, we obtained similar values ~~like~~ as for Punta Arenas in the Southern Hemisphere, with a systematic error of ~~-0.4 m/s~~$- 0.4\ \mathrm{ms}^{-1}$ and a random error of 4.6~~m/s~~ms^{-1} . For the Rayleigh ~~wind products~~winds, we obtained a median bias

455   of ~~-0.5 m/s~~$- 0.5\ \mathrm{ms}^{-1}$ and a random error similar to the one ~~in~~ for Punta Arenas with ~~5.7 m/s. Giving~~5.8 ms^{-1} . Given the

fact that more data points are available, the retrieved Rayleigh systematic error for Leipzig is more meaningful, even though one has to consider that latitudinal and longitudinal dependencies of the systematic error have been discovered (Martin et al., 2021; Weiler et al., 2021b) and thus single locations like Leipzig and Punta Arenas are not completely representative for the overall global performance of Aeolus. ~~This~~ The so-called harmonic oscillating bias was partly resolved ~~due to~~ with the M1 temperature correction (Weiler et al., 2021b). However, a leftover effect in this harmonic bias oscillation cannot be ruled out. Especially, if one considers that for Leipzig we could only evaluate the ascending orbit, while for Punta Arenas we evaluated both orbit types.

We also performed a radiosonde-based validation for Punta Arenas, but too less data points from the very few radiosonde launches matching the evaluation criteria have been available so that the results are not statistically significant. We thus do not further discuss this specific validation methodology for Punta Arenas with respect to the long-term analysis. A final overview of the obtained metrics for the validation of Baseline 11 is shown in ~~Tab.~~ Table 2. The same methodology has been applied to the other baselines and will be discussed below.

6 Aeolus ~~Validation~~ validation

We performed the validation analysis for ~~all available baselines, and thus for the time periods listen in Tab. 1 for~~ Punta Arenas (Doppler ~~instruments and radiosonde~~ cloud radar) and Leipzig (~~Radiosonde) . Overview of the different baselines (algorithm versions) which were used to process Aeolus data. As stated above in Sec. 3.2, partly reprocessing led to parallel existing products from different algorithm versions for certain time periods . A graphical overview is given in Fig. 8. Depending on the discussion topic further on, we either use~~ radiosonde) for all available baselines, and thus for the time periods listed in Table 1. In Section 6.1, we analyze different product versions covering the same time period to assess the changes in product quality with changing baseline ~~(Sec. 6.1) or we use the latest algorithm version (baseline) to discuss the performance of the instrument (Sec. 6.2). For example, the periods marked as red, green,~~ . Periods marked with red, light blue, magenta and orange rectangles in Fig. 8 are well appropriate ~~and have been used~~ for a baseline ~~inter-comparison. For the "red"~~ intercomparison. For the red-marked period from June 2019 – September 2019, products from four different algorithm versions are available covering already the FM-B era. The ~~orange marked time period covers~~ dark-yellow-dashed-marked time period comprises Baseline 06, 10, and 11 from July 2019 to October 2019 and the ~~magenta-marked~~ magenta-dashed-marked period represents the comparison period for B07 to B11 ~~that~~ , which covers the time from November 2019 to April 2020. The ~~"green" period with 2 different algorithms~~ light-blue-marked period reaches from June until December 2019 and from May 2020 to October ~~2020.~~ 2020 and covers two different algorithm versions. In Section 6.2 we use the latest algorithm version (baseline) available for the analyzed time period to discuss the performance of the instrument during its lifetime.

6.1 Comparison of ~~Baselines~~ baselines

Due to ~~reprocessing efforts by~~ the reprocessing efforts of the Aeolus team, there are certain periods in which Aeolus data ~~is~~ are available for different baselines as shown in Fig. 8. This allows ~~to validate the~~ the validation of the improvements between
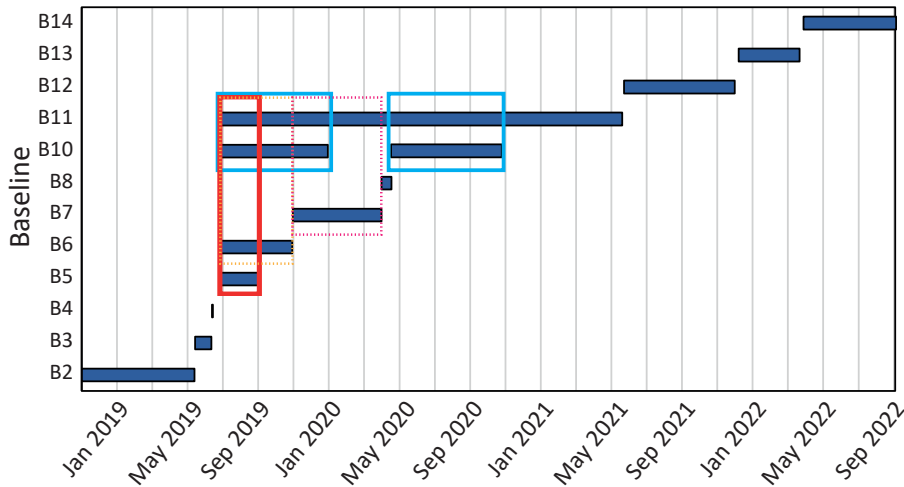
Figure 8. Overview of the different baselines (algorithm versions) that were used to process Aeolus data. The rectangles indicate the periods which have been used for the baseline comparison. The red rectangle refers to Section 6.1.1, the dark yellow dashed one to Section 6.1.2, the magenta dashed one to Section 6.1.3, and the light blue ones to Section 6.1.4.

the different baselines using the same reference data. However, a quantitative measure is not ~~straight forward~~ straightforward as due to quality control (QC) procedures etc., not the same amount of Aeolus wind data ~~is~~ are available. Nevertheless, ~~it~~ such a comparison gives a first inside into the improvement made by introducing new baselines.

6.1.1 B05, B06, B10, B11 comparison

To start with the analysis of the different baselines, we focus on the period from July to September 2019 for which data from four different baselines ~~is~~ are available: B05, B06, B10, and B11 ~~–~~ see red rectangle in Fig. 8. The switch to laser FM-B was already performed at this time. We ~~analysed~~ analyzed this period with the reference data from Punta Arenas and Leipzig. The results for Punta Arenas are shown in Fig. 9 for the Mie (top) and Rayleigh (bottom) wind products.

As this period ~~lasts~~ lasted two months, 16 overpasses in the Southern Hemispheric winter could be covered. ~~Without looking at the exact values for the comparison, it gets already obvious that the amount of data which could be compared has significantly lowered from Baseline B05 and B06 to Baseline B10 and B11 - especially evident for the Rayleigh data - indicating improved quality flags and error calculations. Also, the~~ The greatest differences can be seen for the Rayleigh winds. ~~At~~ For Baseline 05~~and 06~~, many outliers (data not close to the 1:1 line) have been observed, mainly at negative HLOS speeds, which led to a bias of ~~12 and 10 m/s~~ − 8 ms^{-1} for B05. The random error is as high as 20 ms^{-1} for the Rayleigh winds for this baseline (B05). For the Mie winds, a systematic and random error of − 3 ms^{-1} and ~~B06~~7 ms^{-1} was found, respectively. ~~Random errors are as high as 27 and 15 m/s~~With Baseline 06, the overall performance was improved on the cost of less valid data. This indicates already introduced improvement by the new response calibration which was needed since the switch to laser
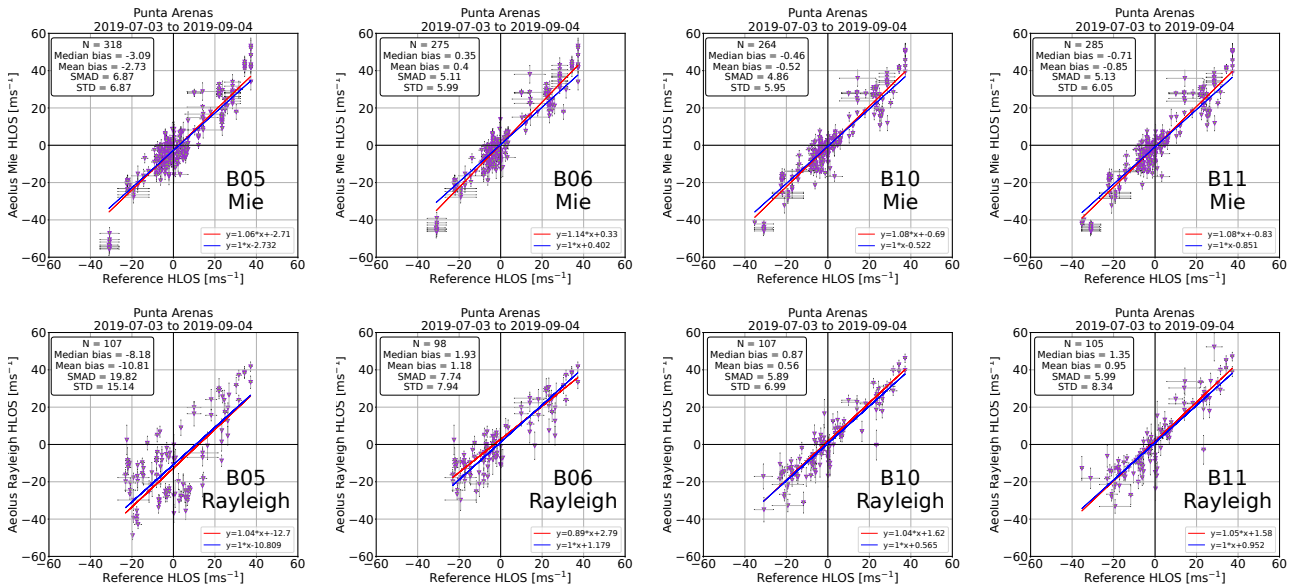
21

Figure 9. Aeolus performance as obtained by comparison to ground-based Doppler cloud radar ~~in~~ at Punta Arenas for Baselines 05, 06, 10, and 11 in the period from July to September 2019. Plots in analogy to Figures 6 and 7, left.

505 FM-B, but could be just obtained in August 2019 and thus two months after the switch. The fact that less valid Aeolus data was available might be caused by improved or more strict quality flags and error calculations (more wind observations flagged as invalid). Finally, less outliers are found for Rayleigh winds as seen in Fig. 9 (B06, Rayleigh). At this baseline (B06), the systematic and random error was as low as 0.4 (1.9) ms^{-1} and 5.1 (7.7) ms^{-1} for Mie (Rayleigh) winds, respectively. Of course, these numbers have to be assessed with care due to the low number of compared overpasses. Nevertheless, when

510 looking at the Rayleigh wind metrics of Baseline 10 ~~and 11~~, a significant improvement is found ~~. While the number of values to be compared has almost halved, most of the outliers seen in B05 and B06 are removed giving confidence for the improvement of the algorithm including quality control with the new baselines. Especially the~~ while also a slightly higher number of data are available. The introduction of the M1 temperature correction with Baseline 09 seems to have significantly improved the Rayleigh winds. Biases of ~~1 and 1.3 m/s~~ 0.9 and 1.4 ms^{-1} have been detected with random errors of about ~~7 m/s~~ 6 ms^{-1}

515 for both B10 and B11, respectively. Also here, it holds that these numbers have to be taken with care due to the relatively low amount of data, but it definitely shows that algorithms have significantly improved and systematic and random errors are at a good level to allow the use of the Rayleigh wind products.

For the Mie winds, the improvement in performance is ~~also seen, even though it is not as significant as for~~ less evident compared to the Rayleigh products, caused by the fact that Mie winds were already much more reliable for ~~B05 and~~ B06 (bias of ~~-2.3 and~~

520 ~~0.3 m/s, and~~ 0.4 ms^{-1} and random error of ~~8.7 and 5.2 m/s, respectively).~~ 5.1 ms^{-1}). The number of available measurements also stayed nearly constant. The reason for having already reliable Mie winds for B06 is that according to Weiler et al. (2021b)
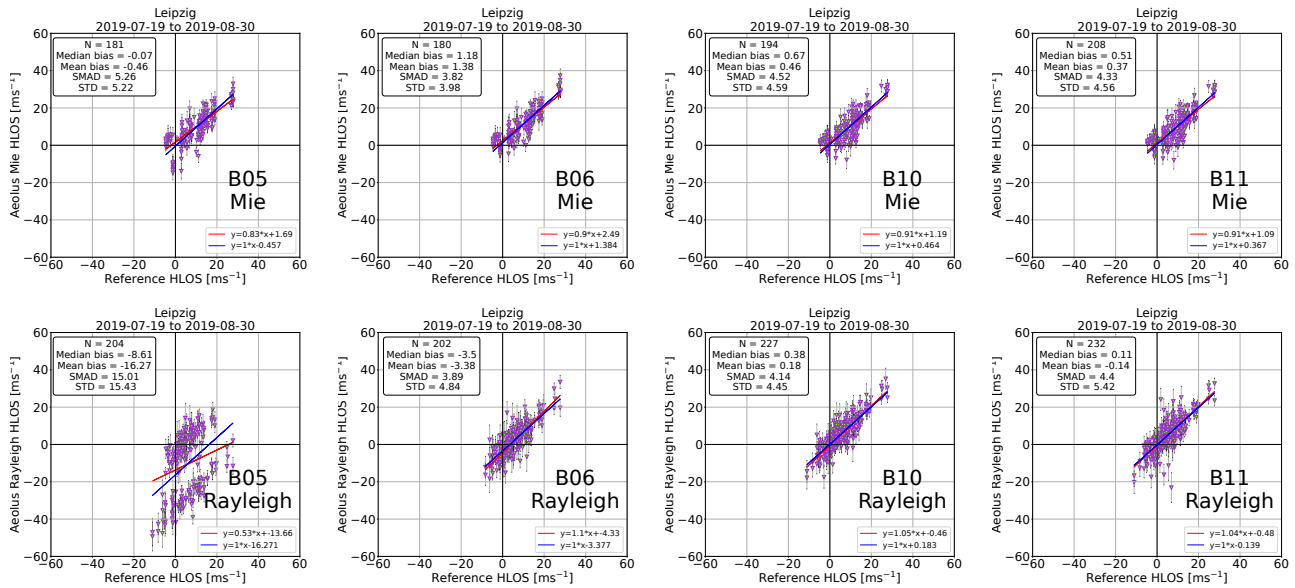
22

Figure 10. Aeolus performance as obtained by comparison to radiosonde launches ~~in~~ at Leipzig for Baselines 05, 06, 10, and 11 in the period from July to ~~September~~ August 2019. Plots in analogy to Figures 6 and 7, left.

, the Mie winds are 10 times less affected by the M1 temperature variations than the Rayleigh products due to the technical nature of the different detection schemes (e.g., also Reitebuch, 2012).

With the introduction of B10 and B11, the median bias ~~is improved to absolute values below~~ stays below absolute values of

525 1 ~~m/s~~ ms^{-1} with random errors of about ~~5m/s~~ 5 ms^{-1} for the Mie products. For both wind types, the difference between B10 and B11 itself is less significant, most probably caused by the low amount of data ~~which~~ that could be used for the comparison. An intense discussion on the B10 to B11 comparison is done later in Sec. 6.1.4 for a longer time period.

For Leipzig, a similar but not equal behaviour was observed as shown in Fig. 10. The ~~number of data points has decreased as well but not as strongly as for Punta Arenas~~ Mie systematic errors were for all compared baselines with absolute values

530 below 1.2 ms^{-1} . The magnitude of the ~~systematic and~~ random error for the Mie winds was improved, but not as significantly as over Punta Arenas ~~and also the statistical error was lower with increasing baseline for the Mie winds, but one has to state that already Baseline 05 was much more reliable compared to the B05 observations at Punta Arenas.~~ In contrast, for the Rayleigh winds the bias was high at Baseline 05 (> ~~8m/s~~ 8 ms^{-1} in absolute values) and could be significantly reduced until B11 (< 0.2 ~~m/s~~ ms^{-1}). The major step forward concerning the random error for the Rayleigh winds was found since ~~B10~~ B06

535 due to the new response calibration function, leading to a decrease from ~~16 m/s~~ 15 ms^{-1} to 4 ~~m/s.~~ ms^{-1} . Interestingly, for the B05 Rayleigh winds (Fig. 10, lower, left), a divided distribution was found with much stronger negative HLOS winds of Aeolus compared to the radiosonde reference. This behaviour has completely vanished since Baseline 06.
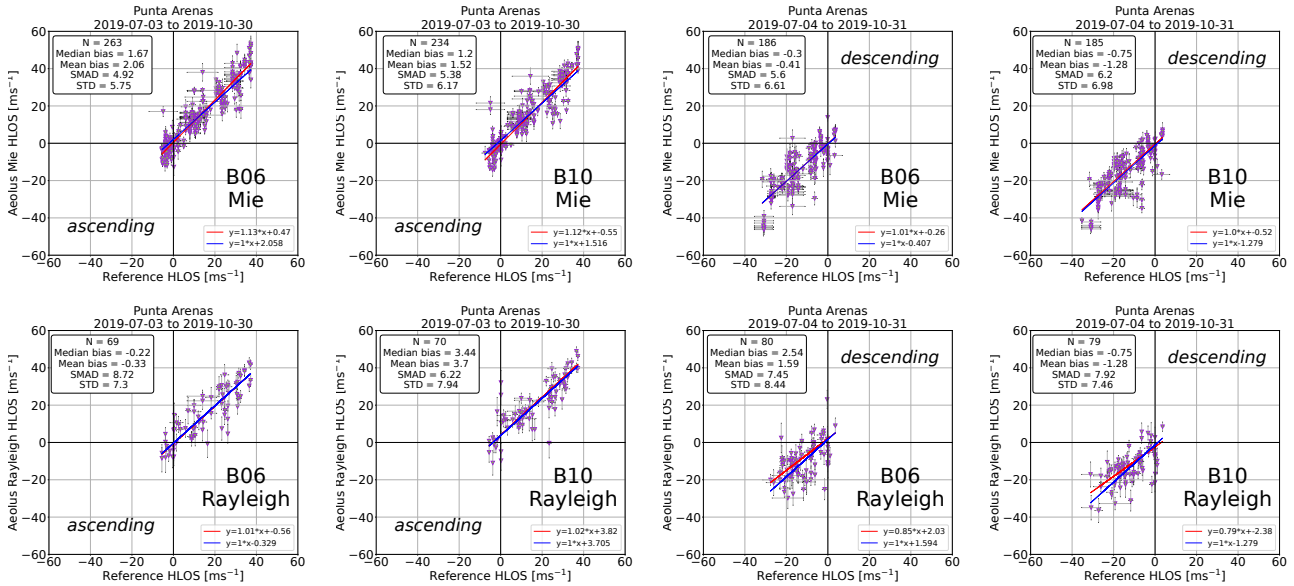
Figure 11. Comparison of Baseline 06 to Baseline 10 ~~in~~ at Punta Arenas ~~but~~ separated for ascending and descending orbit type for the period of July to October 2019. Plots in analogy to Figures 6 and 7, left.

6.1.2 B06 to B10 comparison

~~The~~ As shown in Fig 11, the improvement for B10 Rayleigh winds compared to B06 products becomes even more evident
540   if one compares the longer period available for B06 and B10 only ~~, as shown in Fig 11~~(see a dark yellow dashed rectangle in Fig. 1). Here, the longer time period from July to end of October 2019 could be considered covering 34 overpasses (17 for each orbit type – ascending and descending). Furthermore, the comparison ~~was divided into the~~ focuses on the data quality depending on the orbit types, i.e., if Aeolus measured on an ascending or a descending orbit, to estimate the effect of the M1 temperature correction introduced since B09 (Weiler et al., 2021b). Therefore, we also did not include the Leipzig data
545   (on ascending orbit only) ~~in this comparison. At a first glance there are to be no obvious difference~~ here. There are obvious differences in performance of B06 between ascending and descending orbit type for the Rayleigh ~~winds, while the~~ and Mie winds. E.g., a systematic error for the Rayleigh wind products of $-0.22\,\mathrm{ms^{-1}}$ on the ascending orbit vs. $2.54\,\mathrm{ms^{-1}}$ on the descending orbit. The majority of the outliers is seen for Mie winds ~~in~~ on the descending orbit. ~~Furthermore, it is also seen that these~~ These outliers in the Mie winds remain partly in Baseline 10, so that one can conclude that there must be other reasons for
550   the discrepancy in the Mie wind measurement than the temperature deviation at the Aeolus telescope (Weiler et al., 2021b). For example, it might be atmospheric inhomogeneity which led to the result for which Aeolus measured about ~~45 m/s~~ $-45\,\mathrm{ms^{-1}}$, but the reference instrument only about ~~30 m/s~~ $-30\,\mathrm{ms^{-1}}$.
For ~~the Rayleigh winds, also the amount of data available at~~ B10~~decreases by a factor of two compared to B06 but leading to much improved systematic and random errors. At B10~~, a systematic (random) error of 3.4 (6.2) and ~~0.7 (8.1)m/s~~ -0.8
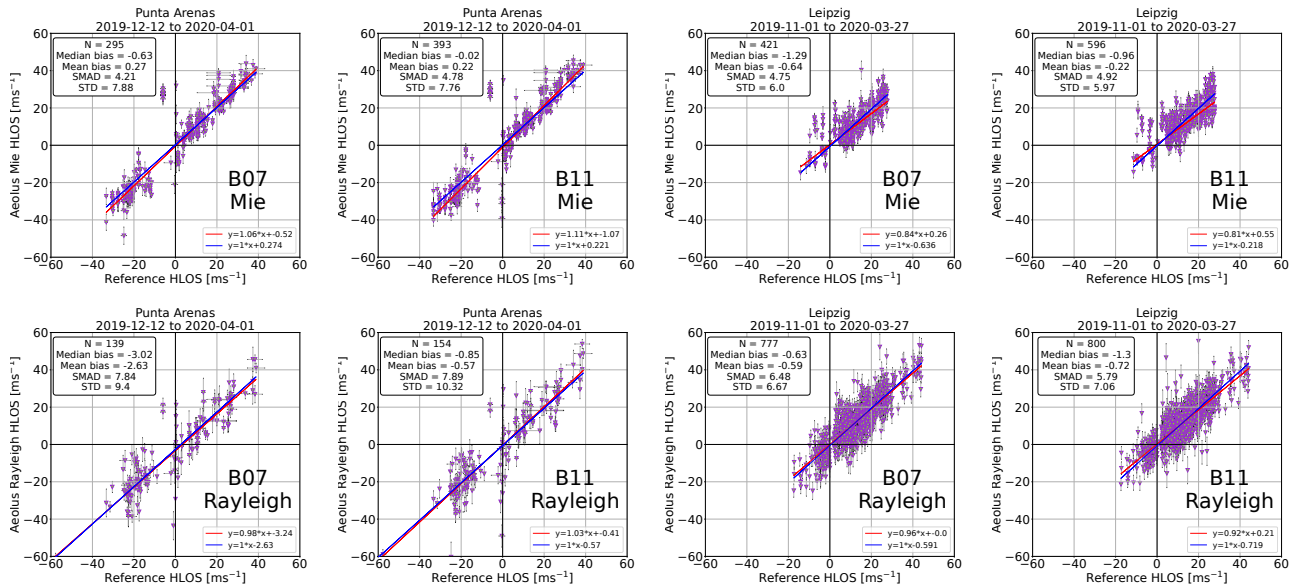
Figure 12. ~~Comparison of the~~ Aeolus performance ~~compared to~~ at the reference ~~observations at~~ stations of Punta Arenas (left) and Leipzig (right) for Baselines 07 and 11 in the period from November 2019 to April 2020. Plots in analogy to Figures 6 and 7, left.

(7.9) ms^{-1} was observed for the ascending and descending Rayleigh winds, respectively, indicating ~~however~~ still a different behaviour between ~~ascending and descending orbits~~ the different orbits types. The same is valid for the Mie winds, with systematic errors of about ~~1.5 m/s but in different direction~~ 1.2 ms^{-1} (ascending) and -0.75 ms^{-1} (descending), while the random error is similar with around 6 ms^{-1} .

6.1.3 B07 to B11 comparison

We also analysed the difference between B07 and B11 ~~which is given for a longer time series than the comparisons before to analyse the improvement made with the implementation of the M1 telescope temperature correction on a more statistically significant basis.~~ (see dashed magenta rectangle in Fig. 1). The used data set covers Punta Arenas and Leipzig data from November 2019 to April 2020, thus ~~Southern~~ southern hemispheric summer and ~~Northern~~ northern hemispheric winter conditions. It is therefore well suited to statically analyse the influence of implementation of the M1 telescope temperature correction on the data quality.

~~Giving~~

According to the results presented in Fig. 12, it becomes evident that between B07 and B11, like for B06 to B10, a significant improvement has to be attributed to the Rayleigh wind performance with much less outliers at both locations. Exemplary stated for Punta Arenas only (Fig. 12, left, bottom), a lower systematic (~~-5.2 vs. -0.9 m/s) and random error (14.6 vs.~~ -3 vs. -0.9 ms^{-1}) with almost equal random errors (about 7.8~~m/s) on the costs of the observations available (157 at B11 compared
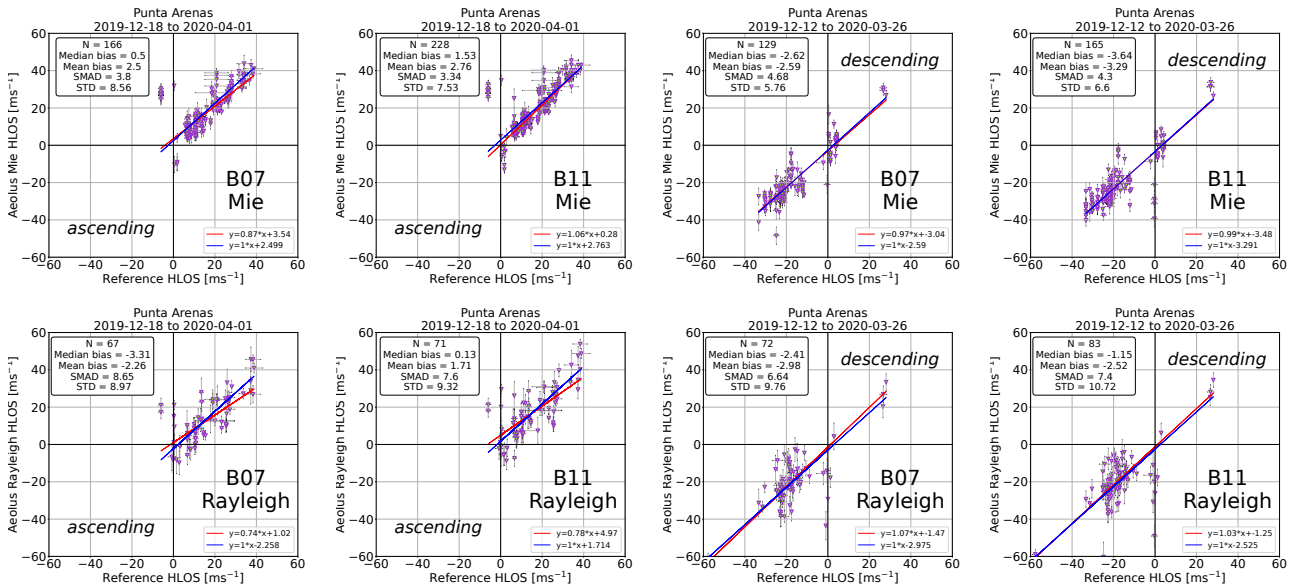
Figure 13. Comparison of Baseline 07 to Baseline 11 ~~in~~ at Punta Arenas ~~but~~ separated for ascending and descending orbit type for the period from December 2019 to April 2020. Plots in analogy to Figures 6 and 7, left.

~~to 289 at B07) were found~~ ms^{-1}) were found for the Rayleigh winds. For the Mie wind performance at Punta Arenas (Fig. 12, left, top), the systematic error improved from ~~-0.7 m/s~~ − 0.6 ms^{-1} to near 0, while the random error stayed equal with ~~4.7 m/s~~ less than 5 ms^{-1} with an even increasing number of ~~comparable observations (342 to 403~~ available observations (295 to 393). The number of ~~observations~~ available observations of the Mie winds has also increased at Leipzig (Fig. 12, right, top),

575 but with no significant changes in the errors.~~.~~, for both Mie and Rayleigh (Fig. 12, right) wind products.
As the separation between the ascending and descending orbit is one key element for identifying the harmonic bias effect, we separated the statistics according to that for Punta Arenas observations only and ~~discard~~ discarded the Leipzig data set (available for ascending orbit type only). The results are shown in Fig. 13. Without going into too much detail, we could not identify a significant improvement in the performance with respect to the Mie ~~cloudy~~ wind product in this specific data set

580 covering 5 months of observations. Based on ~~more than~~ around 150 data points, the Mie systematic error was for B07 +0.5~~m/s~~ ms^{-1} on the ascending orbit while it was almost ~~-3 m/s~~ − 3 ms^{-1} on the descending orbit and thus remarkably different. With Baseline 11, however, the differences in the bias have even increased: +1.5~~m/s~~ ms^{-1} for the ascending orbit and ~~-3.3 m/s~~ − 3.6 ms^{-1} for the descending orbit, both for the Mie wind product. Thus, based on this limited data set, no improvement was found for the Mie winds. The Rayleigh winds, however, had an equal bias of ~~almost -5 m/s at~~ around − 3 ms^{-1} for B07,

585 which significantly decreased to ~~1.5 m/s and -1.1 m/s~~ 0.1 ms^{-1} and − 1.2 ms^{-1} (ascending and descending~~respectively) but on the costs of only half the observations at~~, respectively) for B11.
These comparisons ~~are not thought as~~ discussed above are an indicator for the ~~Aeolus performance at all, but to show the changes with changing baselines. Having a significant~~ improvements made between different baselines - especially the

26

~~importance of the implementation of the M1 temperature correction for the Rayleigh wind product. They are not meant as a general statement on the Aeolus performance as the analysed time period is short. A significantly~~ large data set is key to determine statistically significant measures for a single validation site, like Leipzig or Punta Arenas. Therefore, the ~~differences in the overall product performance on a longer time series is made~~ product performance is analysed between B10 and B11 on a longer time series in the following.

6.1.4 B10 ~~vs.~~ to B11 comparison

As stated above, one major improvement step was reached by the introduction of the M1 telescope temperature correction with Baseline 09. Thus, it ~~would be~~ is also of interest to compare the algorithm versions beyond this Baseline. This is possible for B10 and B11, for which a significant amount of data ~~is~~ are available in parallel as seen in Fig. 8 - light blue rectangles. The most important differences between B10 and B11 ~~is~~ are the implementation of the Sat-LOS velocity correction, the reporting of the Rayleigh spot location and width values, and different ~~Signal-to-Noise Ratio~~ signal-to-noise ratio (SNR) thresholds for classification of Mie and Rayleigh (ESA, private communication~~/confluence~~).

However, ~~giving~~ according to Fig. 14, left, no significant difference can be found at Punta Arenas between B10 and B11 for both ~~,~~ Mie and Rayleigh wind products, despite the fact that about ~~10~~5 % more Mie winds are available, which is most probably due to the new SNR thresholds for the wind type classification. ~~The~~ In fact, the performance of the Rayleigh ~~winds is indeed slightly worse , but giving the uncertainty not statistically significant~~ and Mie winds is slightly worse (overall small increase in systematic and random error). We can only speculate about the reasons for that, but it might be due to the new wind type classification or the newly implemented Satellite-LOS velocity correction. Similar findings are made for Leipzig, Fig. 14, right, for which the radiosondes could cover a much ~~broader~~ larger height range compared to the cloud radar observations in Punta Arenas~~but covering only~~, but only for the ascending orbit. Here, the absolute bias has slightly decreased from ~~-0.44 to -0.29 m/s~~ -0.44 to -0.29 ms^{-1} for the Mie winds with similar random error, but more measurements ~~at~~ for B11 as also observed for Punta Arenas. For Rayleigh winds, no difference at all is seen, giving confidence that for this wind type Baseline 10 was already working well over the atmospheric range from 0 to 25~~km~~ km - at least on Aeolus' ascending orbit over central Europe. If one separates the orbit types for the statistical analysis, which is possible for Punta Arenas (Fig. 15), it is interesting to note that still a significant difference in the bias occurs between the two orbit types for both baselines. With respect to the comparison of ~~B10 to B11, for~~ the Mie winds on the ascending orbit between B10 to B11, the bias decreased, while on the descending orbit it increased (in terms of magnitude) from ~~-1.8 to -2.2 m/s but considering more wind values~~ -1.7 to -2.4 ms^{-1} . For Rayleigh winds, also like in Leipzig, no significant difference is seen at the geographic region of Punta Arenas between the two baselines~~but also here~~, but with significant differences between the orbit types ~~.~~ (> 2 ms^{-1} vs. ≈ -2 ms^{-1}). The random error remained equal between the baselines for both wind types.
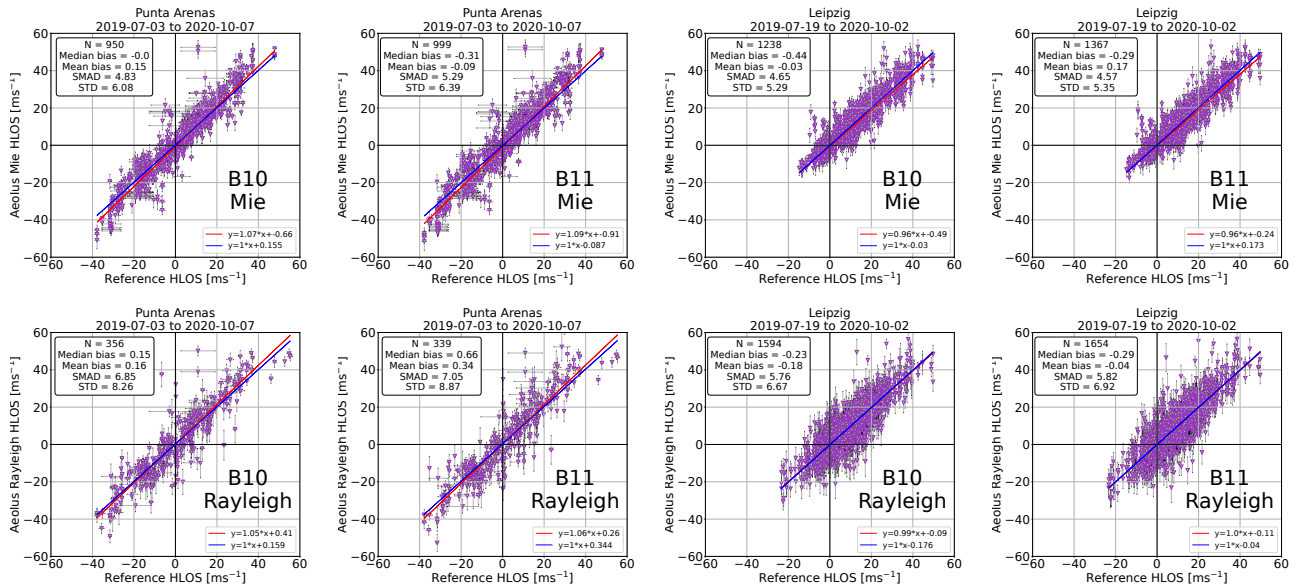
Figure 14. ~~Comparison of the~~ Aeolus performance ~~compared to~~ at the reference ~~observations in~~ stations of Punta Arenas (left) and Leipzig (right) for Baselines 10 and 11 in the period from July 2019 to October 2020. Plots in analogy to Figures 6 and 7, left.

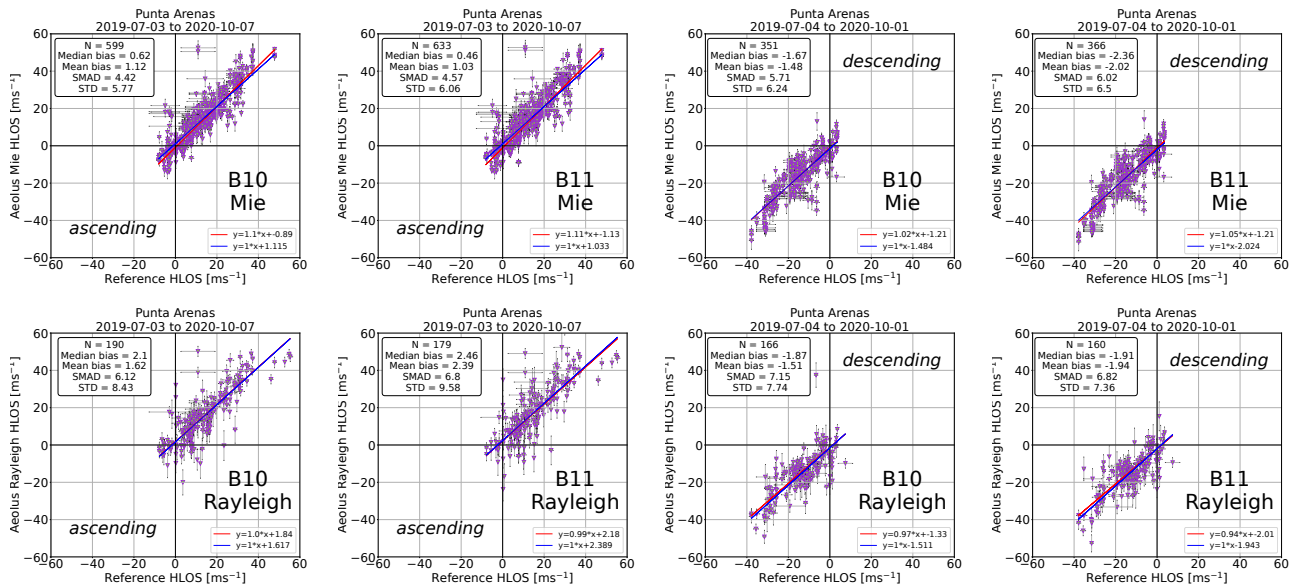**Figure 15.** Comparison of Baseline 10 to Baseline 11 ~~in~~ at Punta Arenas ~~but~~ separated for ascending and descending orbit type for the period from July 2019 to October 2020. Plots in analogy to Figures 6 and 7, left.
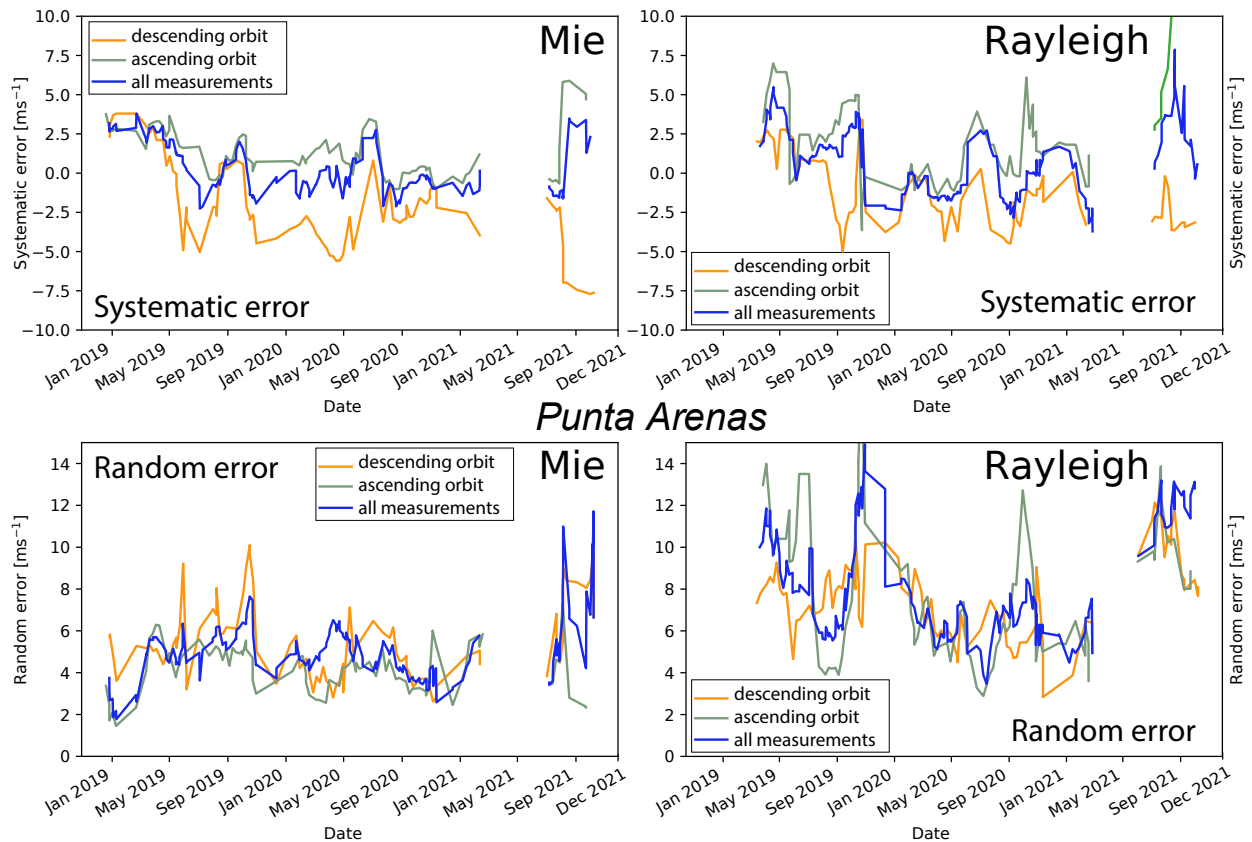
Figure 16. ~~Long term~~ Long-term evolution of the derived systematic error (top) and random error (bottom) for the Aeolus Mie (left) and Rayleigh (right) wind products obtained at the Cal/Val station of Punta Arenas for all observations (~~yellow~~blue) and separated by orbit type (ascending: ~~pink~~grey-green, descending: ~~purple~~orange). A ~~28-days~~ 50-days moving average window was applied.

6.2 Error evolution during lifetime

In the following, we assess the long-term performance of Aeolus. Thus, it is a mix of instrument performance and algorithm improvements. Figure 16 shows the temporal evolution of the systematic error (median bias, top) and random error (scaled MAD, bottom) for the Mie (left) and Rayleigh (right) products for the full 3-year data set at Punta Arenas. The temporal evolution was computed by using a ~~28-days~~ 50-day moving average window~~. The most recent available baseline was used and the~~, i.e., 7 full weeks with 14 Aeolus overpasses (7 for each orbit type). The newest baseline release was used for this analysis, i.e., for periods for which several baselines co-exist, the one with the highest number was analysed. The results are shown for all validation measurements (blue line) and are split into ascending ~~and descending orbit.~~ (grey-green) and descending orbits (orange). Note that we here also present Aeolus data which is not yet public, i.e., from the very early mission time and thus this should not be regarded as final performance indicator of Aeolus. At ~~this Southern Hemispheric~~ the southern hemispheric mid-latitude location of Punta Arenas, the systematic error of the Mie wind decreased from around

3 ms^{-1} in the beginning towards almost 0 ms^{-1} since June 2019 (for the combined observations including both orbit types, blue color). However, a difference between the wind products of the separate two orbit types (orange: descending, grey-green: ascending) becomes obvious especially for the period between October 2019 and August 2020. Sporadic outliers like in September 2020 might be due to certain weather conditions in Punta Arenas. The increase at the end of the observational period in 2021 might be attributed to the orbit shift performed for Aeolus in June 2021 and the resulting larger distances to the validation site. We also had to increase the radius from 100 to 120 km to still be able to validate both orbit types. Thus, the significant increase in magnitude of the systematic error might be attributed to the increased distance (mean distance 75 km and 85 km compared to 27 km and 75 km before the orbit shift). The random error of the Mie winds at Punta Arenas varies between 2 and 10 ms^{-1}, but with most of the higher values after the orbit shift. The increase in random error since beginning 2019 might be attributed to the reduced return signal with laser FM-A and the calibration procedures after the laser switch (e.g., Parrinello et al., 2022). Additionally, the random error on the descending orbit shows a significant increase since the orbit shift.

For the Rayleigh winds, a significant improvement in terms of bias can be seen shortly after the start of the observations. Afterwards, the systematic error of the Rayleigh wind product seem to fluctuate between -5 and $+5$ ms^{-1} during the whole analyzed period, which might be an indicator for a reduced meaningfulness of the reference observation, which are available in cloudy atmospheric regions only compared to the Rayleigh winds available in clear air regions only, for a 50-day averaging window. Nevertheless, generally the systematic error for the descending orbit is mostly negative while the one for the ascending orbit is mainly positive. Therefore, the total retrieved bias fluctuates between positive and negative values. The random error of the Rayleigh winds has significantly improved in course of the mission lifetime from more than 10 ms^{-1} in the beginning of the observations to values of around 5 ms^{-1} in the middle of the analysed period. After the orbit shift in June 2021, the magnitude of the systematic and random error increased for all orbit types. Detailed reasons are yet unclear but might be simply attributed to the larger distances between Aeolus and the ground-reference instruments after the orbit change.

A similar analysis was made for the Leipzig data set, which covers the ascending orbit only, but therefore is available until beginning of autumn 2022 (and thus includes Baseline 13 and 14, completely). The results are shown in Fig. 17. Similarly to Punta Arenas, the temporal evolution of the systematic and random error of the Mie and Rayleigh products for Leipzig have been analyzed by means of the median bias and the scaled MAD, respectively. A 50-day moving average window was applied, i.e., one smoothing window contained 7 overpasses, as the ascending orbit is analysed only for Leipzig.

The analysis for this location reveals, in accordance with the analysis for Punta Arenas before the orbit shift, that the systematic
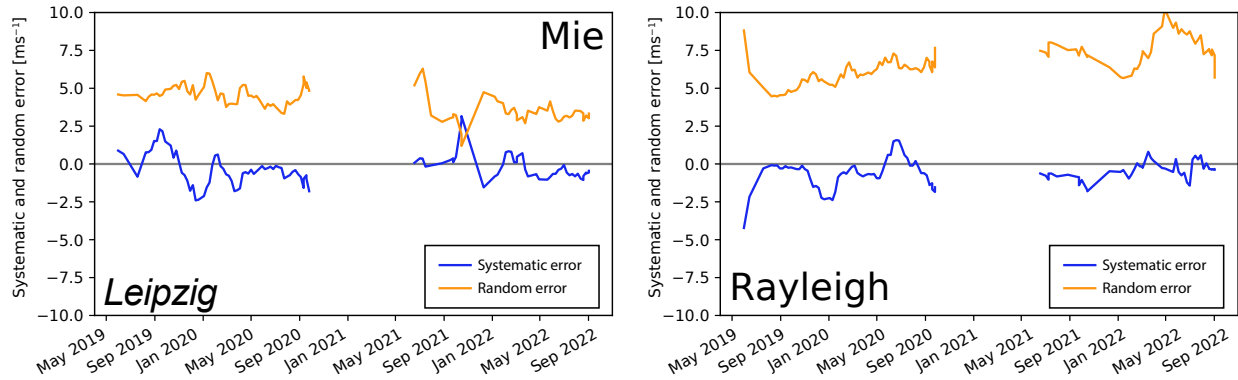
Figure 17. Temporal evolution of the derived systematic error (~~black~~ blue - median bias) and random error (~~purple~~ orange - scaled MAD) for the Aeolus Mie (left) and Rayleigh (right) wind products obtained at the Cal/Val station of Leipzig (ascending orbit). A ~~28-days~~ 50-days moving average window was applied.

665 error of the ~~Mie wind systematic error has significantly decreased with the start of the FM-B period in June 2019. Systematic error values~~ Mie wind product was close to 0~~m/s were already achieved end of 2019.~~ ms^{-1} for the entire FM-B period (from June 2019 until September 2022). The random error for the Mie products has been stable at values around 4 to 5 ~~m/s until end~~ ms^{-1} until mid of 2021. Afterwards, it decreased ~~down~~ significantly followed by an increase and stabilization at 3 to 4~~m/s~~ ms^{-1} until the end of the analysis period. Please note the observational gap which occurred during winter/spring

670 2020/2021 due to ~~COVID~~ COVID-19 restrictions. Thus, no radiosonde (reference) data ~~was~~ were available.

For the Rayleigh winds, also a positive performance trend was observed. The magnitude of the systematic error decreased significantly from values of around 4~~m/s~~ ms^{-1} to magnitude values below 2~~m/s~~ ms^{-1}. The random error also decreased until the end of 2019 due to performance improvements obtained with the switch to laser FM-B but later continuously increased as a result of the continuously decreasing return signal at Aeolus (e.g., Parrinello et al., 2022). The Rayleigh random errors at

675 Leipzig, however, stayed always below 10~~m/s~~ ms^{-1}.

6.3   Validation summary

We performed a validation analysis for both Aeolus wind products (Mie and Rayleigh winds) for the period for which our reference observations (Doppler cloud radar and radiosonde) were available. We thus considered several different baselines (see ~~Tab.~~ Table 1). The main results in terms of systematic and random error for Punta Arenas (Doppler radar) and Leipzig

680 (~~Radiosonde~~ radiosonde) are summarized in ~~Tab.~~ Table 3.

According to ~~Tab.~~ Table 3, left half, the systematic error of the Aeolus wind products could be significantly lowered with the changes introduced into the processing chain (different baselines). While in the early mission phase, systematic errors of more than 2~~m/s~~ ms^{-1} (absolute values) were observed for both wind types, these biases could be reduced with the algorithm improvements, such as new calibration procedures or the M1 temperature correction with Baseline 09. Hence, since Baseline

31

Table 3. Overview of the systematic error (median bias) and the random error (scaled MAD) for the different baselines of Aeolus derived with the reference measurements at Punta Arenas and Leipzig. All values are in $\mathrm{ms^{-1}}$.

	Systematic error				Random error			
	Punta Arenas		*Leipzig*		*Punta Arenas*		*Leipzig*	
	Rayleigh	Mie	Rayleigh	Mie	Rayleigh	Mie	Rayleigh	Mie
B02	3.1	2.73	-	-	10.97	4.68	-	-
B05	−8.18	−3.09	−8.61	−0.77	19.82	6.87	15.01	5.26
B06	1.51	0.91	−3.14	1.09	7.98	5.57	4.76	4.23
B07	−3.02	−0.63	−0.63	−1.29	7.84	4.21	6.48	4.75
B10	0.15	−0.0	−0.23	−0.44	6.85	4.83	5.76	4.65
B11	−0.0	−0.41	−0.46	−0.35	6.79	5.05	5.77	4.59
B12	0.56	−0.69	−0.86	0.42	11.2	5.4	7.33	4.26
B13	-	-	−0.37	−0.04	-	-	5.97	4.24
B14	-	-	−0.34	−0.67	-	-	8.49	3.2

685 10, a significant improvement of the Aeolus data was found leading to a low bias (close to 0 ~~m/s~~ $\mathrm{ms^{-1}}$) for the Rayleigh winds and nearly similar values for the mid-latitudinal sites on both hemispheres. ~~As the performance of the laser onboard Aeolus has been getting less ideal (Parrinello et al., 2022), the~~ The systematic error of the Mie winds was already significantly reduced with Baseline 06. The random errors for the wind products as shown in ~~Tab. 3~~ Table 3, right part, are first decreasing with increasing baseline, but later increasing again ~~. However, the~~ as a result of the performance losses of the laser onboard

690 Aeolus (Parrinello et al., 2022), but mainly only affecting the Rayleigh winds. The systematic error is only slightly affected by this issue, so ~~that~~ one can conclude that the uncertainty introduced by the reduced atmospheric return signal received by Aeolus is mostly affecting the random error - of course on the cost of having less valid wind data, but at least no significant additional bias seems to be introduced.

7 Conclusions

695 To validate the novel wind lidar mission Aeolus, we have gathered long-term validation data at two mid-latitudinal sites but at different hemispheres. More specifically, we have performed regular radiosonde launches for the weekly Aeolus overpasses at Leipzig, Germany (51.35° N, 12.43° E), since May 2019. We also operated a scanning Doppler cloud radar in Punta Arenas, Chile (~~53.2~~ 53.15° S, ~~70.9~~ 70.91° W), so that horizontal wind speed and direction could be retrieved in the vicinity of clouds. ~~Additionally, occasional radiosonde launches were performed.~~ We used all these data sources to validate the overall Aeolus

700 performance with respect to mission time, the algorithm (baseline) version applied to Aeolus data, and the orbit type (ascending, descending, both). It was found that the deviation of the Aeolus HLOS winds from the ground-reference is of Gaussian shape. As systematic error indicator we thus applied the median bias of this distribution, while the random error was attributed to the

scaled median absolute deviation in accordance to previous validation work on Aeolus and agreement within ESA and DISC (e.g., Lux et al., 2020a). It should be noted, that while the radiosonde data performed at Leipzig on Aeolus' ascending orbit is covering the whole atmosphere from ground to ca. 22-25 km height, the Punta Arenas reference measurements with cloud radar are restricted to cloudy regions in the troposphere, i.e., the results from Punta Arenas are not representing stratospheric wind observations. Nevertheless, as clouds above Punta Arenas can be generally found at all altitudes and frequently, we consider the troposphere as well covered. The main findings, i.e., the systematic and random error by baseline, have been summarized in Table 3. In general, we have found an improving performance of the HLOS wind products with respect to the baseline development. This effect was however partly masked by the effect of lower instrumental performance of Aeolus during its lifetime, especially for the random error. From the whole Aeolus lifetime, we mainly analysed the period that was conducted with the spare laser called FM-B (starting with Baseline 05). Even when considering the issues with the lower-than-expected emitted laser energy and the received atmospheric return signal (e.g., Parrinello et al., 2022), which constantly decreased despite many efforts made, we can confirm the general validity of Aeolus observations during the lifetime. The systematic error of both wind products (Rayleigh and Mie) has significantly decreased as a result of newly introduced baselines with new calibrations and corrections. While at the beginning of the mission, absolute values as high as 5 ms^{-1} were observed for the systematic error, it was continuously reduced to values close to 0 ms^{-1} before the public release of the Aeolus data in April 2020. This proves the general concept of this Earth explorer mission to perform active wind observations from space. The random error has been indeed higher than requested by the mission requirements. But compared to the loss in return signal, the performance of Aeolus has been still in a range bringing a significant benefit for the numerical weather forecast as demonstrated, e.g., at ECMWF (Rennie et al., 2021), DWD (Martin et al., 2022), and NCMRWF (Rani et al., 2022). The data set gathered at Punta Arenas, Chile and Leipzig, Germany in the course of the validation project EVAA will stay of high value for the Aeolus mission. It can, for example, further be used to validate new algorithm versions applied to historical Aeolus data or to test new methodological approaches. Such efforts will continue even after the satellite has stopped measuring and will help to foster potential follow-on activities for active wind measurements from space as it is currently planned.

Data availability. The radiosonde data from Leipzig are available at ESA Atmospheric Validation Data Centre (EVDC). Aeolus data are available via ESA Aeolus Online Dissemination System. The Cloudnet data used in this study are generated by the Aerosol, Clouds and Trace Gases Research Infrastructure (ACTRIS) and are available from the ACTRIS Data Centre using the following link: https://hdl.handle.net/21.12132/1.c06a2c60c7504072.

Author contributions. HB has conceptualized the study and led the manuscript writing. JW has developed the algorithm for retrieving horizontal wind from ground-based Doppler radar (and lidar) under the supervision of MR and JB. EB has developed the methodology for the comparison of Aeolus to the ground-reference observations under supervision of HB. JW has finally incorporated these previous works for the long-term analysis under supervision of HB. BB and PS have been responsible for the Punta Arenas operations together with MR and

735 JB, HG for the radiosonde launches in Leipzig. UW contributed her expertise on space-borne profiling and general supervision. All authors have contributed to the intense discussions and the manuscript.

Disclaimer. The presented work includes preliminary data (not fully calibrated/validated and not yet publicly released) of the Aeolus mission
740 that is part of the European Space Agency (ESA) Earth Explorer Programme. This includes wind products from processor versions before Baseline 10 and/or aerosol and cloud products from before baseline version 11, which have not yet been reprocessed. The processor development, improvement and product reprocessing preparation are performed by the Aeolus DISC (Data, Innovation and Science Cluster), which involves DLR, DoRIT, ECMWF, KNMI, CNRS, S&T, ABB, Serco and TROPOS, in close cooperation with the Aeolus PDGS (Payload Data Ground Segment). The analysis has been performed in the framework of the Aeolus Scientific Calibration and Validation Team (ACVT).

References

Abril-Gago, J., Guerrero-Rascado, J. L., Costa, M. J., Bravo-Aranda, J. A., Sicard, M., Bermejo-Pantaleón, D., Bortoli, D., Granados-Muñoz, M. J., Rodríguez-Gómez, A., Muñoz Porcar, C., Comerón, A., Ortiz-Amezcua, P., Salgueiro, V., Jiménez-Martín, M. M., and Alados-Arboledas, L.: Statistical validation of Aeolus L2A particle backscatter coefficient retrievals over ACTRIS/EARLINET stations on the Iberian Peninsula, Atmospheric Chemistry and Physics, 22, 1425–1451, https://doi.org/10.5194/acp-22-1425-2022, 2022a.

Abril-Gago, J., Ortiz-Amezcua, P., Bermejo-Pantaleón, D., Andújar-Maqueda, J., Bravo-Aranda, J. A., Granados-Muñoz, M. J., Navas-Guzmán, F., Alados-Arboledas, L., Foyo-Moreno, I., and Guerrero-Rascado, J. L.: Validation Activities of Aeolus Wind Products in the Southeastern Iberian Peninsula, Atmospheric Chemistry and Physics Discussions, 2022, 1–32, https://doi.org/10.5194/acp-2022-732, 2022b.

Baars, H., Kanitz, T., Engelmann, R., Althausen, D., Heese, B., Komppula, M., Preißler, J., Tesche, M., Ansmann, A., Wandinger, U., Lim, J.-H., Ahn, J. Y., Stachlewska, I. S., Amiridis, V., Marinou, E., Seifert, P., Hofer, J., Skupin, A., Schneider, F., Bohlmann, S., Foth, A., Bley, S., Pfüller, A., Giannakaki, E., Lihavainen, H., Viisanen, Y., Hooda, R. K., Pereira, S. N., Bortoli, D., Wagner, F., Mattis, I., Janicka, L., Markowicz, K. M., Achtert, P., Artaxo, P., Pauliquevis, T., Souza, R. A. F., Sharma, V. P., van Zyl, P. G., Beukes, J. P., Sun, J., Rohwer, E. G., Deng, R., Mamouri, R.-E., and Zamorano, F.: An overview of the first decade of Polly[NET]: an emerging network of automated Raman-polarization lidars for continuous aerosol profiling, Atmos. Chem. Phys., 16, 5111–5137, https://doi.org/10.5194/acp-16-5111-2016, 2016.

Baars, H., Geiß, A., Wandinger, U., Herzog, A., Engelmann, R., Bühl, J., Radenz, M., Seifert, P., Althausen, D., Heese, B., Ansmann, A., Martin, A., Leinweber, R., Lehmann, V., Weissmann, M., Cress, A., Filioglou, M., Komppula, M., and Reitebuch, O.: First results from the German Cal/Val activities for Aeolus, EPJ Web Conf., 237, 01 008, https://doi.org/10.1051/epjconf/202023701008, 2020a.

Baars, H., Herzog, A., Heese, B., Ohneiser, K., Hanbuch, K., Hofer, J., Yin, Z., Engelmann, R., and Wandinger, U.: Validation of Aeolus wind products above the Atlantic Ocean, Atmospheric Measurement Techniques, 13, 6007–6024, https://doi.org/10.5194/amt-13-6007-2020, 2020b.

Baars, H., Radenz, M., Floutsi, A. A., Engelmann, R., Althausen, D., Heese, B., Ansmann, A., Flamant, T., Dabas, A., Trapon, D., Reitebuch, O., Bley, S., and Wandinger, U.: Californian wildfire smoke over Europe: A first example of the aerosol observing capabilities of Aeolus compared to ground-based lidar, Geophysical Research Letters, 48, e2020GL092 194, https://doi.org/doi:10.1029/2020GL092194, 2021.

Belova, E., Kirkwood, S., Voelger, P., Chatterjee, S., Satheesan, K., Hagelin, S., Lindskog, M., and Körnich, H.: Validation of Aeolus winds using ground-based radars in Antarctica and in northern Sweden, Atmospheric Measurement Techniques, 14, 5415–5428, https://doi.org/10.5194/amt-14-5415-2021, 2021.

Bley, S., Rennie, M., Žagar, N., Pinol Sole, M., Straume, A. G., Antifaev, J., Candido, S., Carver, R., Fehr, T., von Bismarck, J., Hünerbein, A., and Deneke, H.: Validation of the Aeolus L2B Rayleigh winds and ECMWF short-range forecasts in the upper troposphere and lower stratosphere using Loon super pressure balloon observations, Quarterly Journal of the Royal Meteorological Society, 148, 3852–3868, https://doi.org/10.1002/qj.4391, 2022.

Browning, K. A. and Wexler, R.: The Determination of Kinematic Properties of a Wind Field Using Doppler Radar, Journal of Applied Meteorology and Climatology, 7, 105 – 113, https://doi.org/10.1175/1520-0450(1968)007<0105:TDOKPO>2.0.CO;2, 1968.

Chen, S., Cao, R., Xie, Y., Zhang, Y., Tan, W., Chen, H., Guo, P., and Zhao, P.: Study of the seasonal variation in Aeolus wind product performance over China using ERA5 and radiosonde data, Atmospheric Chemistry and Physics, 21, 11 489–11 504, https://doi.org/10.5194/acp-21-11489-2021, 2021.

795 Chou, C.-C., Kushner, P. J., Laroche, S., Mariani, Z., Rodriguez, P., Melo, S., and Fletcher, C. G.: Validation of the Aeolus Level-2B wind product over Northern Canada and the Arctic, Atmospheric Measurement Techniques, 15, 4443–4461, https://doi.org/10.5194/amt-15-4443-2022, 2022.

Cossu, F., Portabella, M., Lin, W., Stoffelen, A., Marseille, G.-J., Vogelzang, J., and de Haan, S.: Characterization of Aeolus wind measurement errors, in: Aeolus 3rd Anniversary Conference, Taormina, Italy, https://az659834.vo.msecnd.net/eventsairwesteuprod/

800 production-nikal-public/7731204c60c548cb97f0214f5334a267, 2022.

de Kloe, J., Stoffelen, A., Tan, D., Andersson, E., Rennie, M., Dabas, A., Poli, P., and Huber, D.: ADM-Aeolus Level-2B/2C Processor Input/Output Data Definitions Interface Control Document, https://earth.esa.int/eogateway/documents/20142/37627/Aeolus-L2B-2C-Input-Output-DD-ICD.pdf, last access: 5 May 2023, 2023.

Dirksen, R. J., Sommer, M., Immler, F. J., Hurst, D. F., Kivi, R., and Vömel, H.: Reference quality upper-air measurements: GRUAN data

805 processing for the Vaisala RS92 radiosonde, Atmos. Meas. Tech., 7, 4463–4490, https://doi.org/10.5194/amt-7-4463-2014, 2014.

ECMWF: Aeolus data impact tests confirm potential of new wind data for NWP, https://www.ecmwf.int/en/about/media-centre/news/2019/aeolus-data-impact-tests-confirm-potential-new-wind-data-nwp, last access: 8 May 2020, 2019a.

ECMWF: Tests show positive impact of new Aeolus wind data on forecasts, https://www.ecmwf.int/en/about/media-centre/news/2019/tests-show-positive-impact-new-aeolus-wind-data-forecasts, last access: 8 May 2020, 2019b.

810 Engelmann, R., Kanitz, T., Baars, H., Heese, B., Althausen, D., Skupin, A., Wandinger, U., Komppula, M., Stachlewska, I. S., Amiridis, V., Marinou, E., Mattis, I., Linné, H., and Ansmann, A.: The automated multiwavelength Raman polarization and water-vapor lidar Polly[XT]: the neXT generation, Atmos. Meas. Tech., 9, 1767–1784, https://doi.org/10.5194/amt-9-1767-2016, 2016.

ESA: ADM-Aeolus Science Report, Tech. rep., ESA, https://www.esa.int/About_Us/ESA_Publications/ESA_SP-1311_i_ADM-Aeolus_i, last access: 5 May 2023, 2008.

815 ESA: A Guide to Aeolus Range Bin Settings, online, https://earth.esa.int/eogateway/news/a-guide-to-aeolus-range-bin-settings, last access: 31 March 2023, 2020.

ESA: Aeolus Mission Summary, https://earth.esa.int/eogateway/missions/aeolus/description, last access: 5 May 2023, 2023.

Fehr, T., Piña, A., Amiridis, V., Baars, H., von Bismarck, J., Borne, M., Cazenave, Q., Chen, S., Flamant, C., Gaetani, M., Knipperz, P., Koopman, R., Lemmerz, C., Marinou, E., Močnik, G., Parrinello, T., Reitebuch, O., Skofronick-Jackson, G., Straume, A. G., and Zenk, C.:

820 The Joint Aeolus Tropical Atlantic Campaign – First results for Aeolus calibration/validation and science in the tropics, ESA Atmospheric Science Conference (2021), Online, 22.–26. November 2021, 2021.

Flament, T., Dabas, A., Trapon, D., Lacour, A., Ehlers, F., Baars, H., and Huber, D.: Aeolus aerosol and cloud product, in: EGU General Assembly 2021 (Online), European Geosciences Union, https://doi.org/doi:10.5194/egusphere-egu21-14390, 2021.

Geiß, A., Lehmann, V., Leinweber, R., Reitebuch, O., and Weissmann, M.: Methodology and Case Studies for the Validation of Aeolus

825 Observations by means of Radar Wind Profilers, in: ESA Living Planet Symposium 2019, Milan, Italy, 2019.

Geiß, A., Martin, A., Lehmann, V., Leinweber, R., Reitebuch, O., and Weissmann, M.: Long-term validation of the Aeolus L2B wind product with the German radar wind profiler network, in: Aeolus 3rd Anniversary Conference, Taormina, Italy, 2022.

Gkikas, A., Gialitaki, A., Binietoglou, I., Marinou, E., Tsichla, M., Siomos, N., Paschou, P., Kampouri, A., Voudouri, K. A., Proestakis, E., Mylonaki, M., Papanikolaou, C.-A., Michailidis, K., Baars, H., Straume, A. G., Balis, D., Papayannis, A., Parrinello, T., and Amiridis,

830 V.: First assessment of Aeolus Standard Correct Algorithm particle backscatter coefficient retrievals in the eastern Mediterranean, Atmospheric Measurement Techniques, 16, 1017–1042, https://doi.org/10.5194/amt-16-1017-2023, 2023.

Guo, J., Liu, B., Gong, W., Shi, L., Zhang, Y., Ma, Y., Zhang, J., Chen, T., Bai, K., Stoffelen, A., de Leeuw, G., and Xu, X.: Technical note: First comparison of wind observations from ESA's satellite mission Aeolus and ground-based radar wind profiler network of China, Atmospheric Chemistry and Physics, 21, 2945–2958, https://doi.org/10.5194/acp-21-2945-2021, 2021.

835 Görsdorf, U., Lehmann, V., Bauer-Pfundstein, M., Peters, G., Vavriv, D., Vinogradov, V., and Volkov, V.: A 35-GHz Polarimetric Doppler Radar for Long-Term Observations of Cloud Parameters—Description of System and Data Processing, Journal of Atmospheric and Oceanic Technology, 32, 675 – 690, https://doi.org/10.1175/JTECH-D-14-00066.1, 2015.

Hagelin, S., Azad, R., Lindskog, M., Schyberg, H., and Körnich, H.: Evaluating the use of Aeolus satellite observations in the regional numerical weather prediction (NWP) model Harmonie–Arome, Atmospheric Measurement Techniques, 14, 5925–5938,
840 https://doi.org/10.5194/amt-14-5925-2021, 2021.

Illingworth, A., Hogan, R., O'connor, E., Bouniol, D., Brooks, M., Delanoë, J., Donovan, D., Eastment, J., Gaussiat, N., Goddard, J., et al.: Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, Bulletin of the American Meteorological Society, 88, 883–898, 2007.

Iwai, H., Aoki, M., Oshiro, M., and Ishii, S.: Validation of Aeolus Level 2B wind products using wind profilers, ground-based Doppler
845 wind lidars, and radiosondes in Japan, Atmospheric Measurement Techniques, 14, 7255–7275, https://doi.org/10.5194/amt-14-7255-2021, 2021.

Jauhiainen, H., Survo, P., Lehtinen, R., and Lentonen, J.: Radiosonde RS41 and RS92 key differences and comparison test results in different locations and climates, in: TECO-2014, WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observations, Saint Petersberg, Russia, https://library.wmo.int/pmb_ged/iom_116_en/Session%203/P3_16_Juhiainen_Radiosonde_
850 RS41_RS92_Key_Differences_Comparison_TestResults.pdf, last access: 5 May 2023, 2014.

Jensen, M. P., Holdridge, D. J., Survo, P., Lehtinen, R., Baxter, S., Toto, T., and Johnson, K. L.: Comparison of Vaisala radiosondes RS41 and RS92 at the ARM Southern Great Plains site, Atmos. Meas. Tech., 9, 3115–3129, https://doi.org/10.5194/amt-9-3115-2016, 2016.

Krisch, I., Lemmerz, C., Lux, O., Marksteiner, U., Reitebuch, O. Weiler, F., Witschas, B., Bracci, F., Meringer, M., Schmidt, K. Huber, D., Nikolaus, I., Vaughan, M., Dabas, A., Flament, T., Trapon, D., Isaksen, L., Rennie, M. Abdalla, S., Donovan, D., de Kloe, J., Marseille,
855 G.-J., Stoffelen, A., Wernham, D., Kanitz, T., Straume, A.-G., von Bismarck, J., Bley, S., Fischer, P., and Parrinello, T.: Data quality of Aeolus wind measurements, in: EGU General Assembly 2020, Online, 4-8 May 2020, https://doi.org/10.5194/egusphere-egu2020-18186, 2020.

Krisch, I., Hindley, N. P., Reitebuch, O., and Wright, C. J.: On the derivation of zonal and meridional wind components from Aeolus horizontal line-of-sight wind, Atmospheric Measurement Techniques, 15, 3465–3479, https://doi.org/10.5194/amt-15-3465-2022, 2022.

860 Liu, H., Garrett, K., Ide, K., Hoffman, R. N., and Lukens, K. E.: A statistically optimal analysis of systematic differences between Aeolus horizontal line-of-sight winds and NOAA's Global Forecast System, Atmospheric Measurement Techniques, 15, 3925–3940, https://doi.org/10.5194/amt-15-3925-2022, 2022.

Lux, O., Lemmerz, C., Weiler, F., Marksteiner, U., Witschas, B., Rahm, S., Geiß, A., and Reitebuch, O.: Intercomparison of wind observations from the European Space Agency's Aeolus satellite mission and the ALADIN Airborne Demonstrator, Atmos. Meas. Tech., 13, 2075–
865 2097, https://doi.org/10.5194/amt-13-2075-2020, 2020a.

Lux, O., Wernham, D., Bravetti, P., McGoldrick, P., Lecrenier, O., Riede, W., D'Ottavi, A., Sanctis, V. D., Schillinger, M., Lochard, J., Marshall, J., Lemmerz, C., Weiler, F., Mondin, L., Ciapponi, A., Kanitz, T., Elfving, A., Parrinello, T., and Reitebuch, O.: High-power and frequency-stable ultraviolet laser performance in space for the wind lidar on Aeolus, Opt. Lett., 45, 1443–1446, https://doi.org/10.1364/OL.387728, 2020b.

870   Lux, O., Lemmerz, C., Weiler, F., Marksteiner, U., Witschas, B., Rahm, S., Geiß, A., Schäfler, A., and Reitebuch, O.: Retrieval improvements for the ALADIN Airborne Demonstrator in support of the Aeolus wind product validation, Atmospheric Measurement Techniques, 15, 1303–1331, https://doi.org/10.5194/amt-15-1303-2022, 2022a.

Lux, O., Witschas, B., Geiß, A., Lemmerz, C., Weiler, F., Marksteiner, U., Rahm, S., Schäfler, A., and Reitebuch, O.: Quality control and error assessment of the Aeolus L2B wind results from the Joint Aeolus Tropical Atlantic Campaign, Atmospheric Measurement Techniques,

875   15, 6467–6488, https://doi.org/10.5194/amt-15-6467-2022, 2022b.

Martin, A., Weissmann, M., Reitebuch, O., Rennie, M., Geiß, A., and Cress, A.: Validation of Aeolus winds using radiosonde observations and numerical weather prediction model equivalents, Atmospheric Measurement Techniques, 14, 2167–2183, https://doi.org/10.5194/amt-14-2167-2021, 2021.

Martin, A., Weissmann, M., and Cress, A.: Investigation of dynamical scenarios leading to particularly high impact of Aeolus on NWP

880   forecasts, EGUsphere, 2022, 1–23, https://doi.org/10.5194/egusphere-2022-1150, 2022.

Ohneiser, K., Ansmann, A., Baars, H., Seifert, P., Barja, B., Jimenez, C., Radenz, M., Teisseire, A., Floutsi, A., Haarig, M., Foth, A., Chudnovsky, A., Engelmann, R., Zamorano, F., Bühl, J., and Wandinger, U.: Smoke of extreme Australian bushfires observed in the stratosphere over Punta Arenas, Chile, in January 2020: optical thickness, lidar ratios, and depolarization ratios at 355 and 532 nm, Atmospheric Chemistry and Physics, 20, 8003–8015, https://doi.org/10.5194/acp-20-8003-2020, 2020.

885   Parrinello, T., Straume, A. G., Von Bismark, J., Tran, V., Romanazzo, M., Wernham, D., Krisna, T. C., Sathe, A., Colangeli, G., Garsva, S., Fehr, T., Krisch, I. Reitebuch, O., and Rennie, M.: Aeolus: ESA's wind mission3+ years in space. Status and future challenges, in: ESA Living Planet Symposium 2022, Bonn, Germany, https://earth.esa.int/living-planet-symposium-2022-presentations/25.05.Wednesday/Nairobi_1-2/0830-1010/01_Parrinello_Session_B2_11_Version_3.pdf, last access: 5 May 2023, 2022.

Päschke, E., Leinweber, R., and Lehmann, V.: An assessment of the performance of a 1.5 μ m Doppler lidar for operational vertical wind

890   profiling based on a 1-year trial, Atmospheric Measurement Techniques, 8, 2251–2266, https://doi.org/10.5194/amt-8-2251-2015, 2015.

Radenz, M., Bühl, J., Seifert, P., Baars, H., Engelmann, R., Barja González, B., Mamouri, R.-E., Zamorano, F., and Ansmann, A.: Hemispheric contrasts in ice formation in stratiform mixed-phase clouds: disentangling the role of aerosol and dynamics with ground-based remote sensing, Atmospheric Chemistry and Physics, 21, 17 969–17 994, https://doi.org/10.5194/acp-21-17969-2021, 2021.

Radenz, M., Seifert, P., and O'Connor, E.: Classification data from Punta Arenas on 6 February 2020, https://hdl.handle.net/21.12132/1.

895   c06a2c60c7504072, 2022.

Rani, S. I., Jangid, B. P., Kumar, S., Bushair, M. T., Sharma, P., George, J. P., George, G., and Das Gupta, M.: Assessing the quality of novel Aeolus winds for NWP applications at NCMRWF, Quarterly Journal of the Royal Meteorological Society, 148, 1344–1367, https://doi.org/10.1002/qj.4264, 2022.

Ratynski, M., Khaykin, S., Hauchecorne, A., Wing, R., Cammas, J.-P., Hello, Y., and Keckhut, P.: Validation of Aeolus wind profiles using

900   ground-based lidar and radiosonde observations at La Réunion Island and the Observatoire de Haute Provence, EGUsphere, 2022, 1–33, https://doi.org/10.5194/egusphere-2022-822, 2022.

Ray, P. S. and Ziegler, C.: De-Aliasing First-Moment Doppler Estimates, Journal of Applied Meteorology and Climatology, 16, 563 – 564, https://doi.org/10.1175/1520-0450(1977)016<0563:DAFMDE>2.0.CO;2, 1977.

Reitebuch, O.: The Spaceborne Wind Lidar Mission ADM-Aeolus, in: Atmospheric Physics: Background – Methods – Trends, edited by

905   Schumann, U., pp. 815–827, Springer, https://doi.org/10.1007/978-3-642-30183-4_49, 2012.

Reitebuch, O., Huber, D., and Nikolaus, I.: Algorithm Theoretical Basis Document ATBD: ADM-Aeolus Level 1B Products, https://earth.esa.int/eogateway/documents/20142/37627/Aeolus-L1B-Algorithm-ATBD.pdf, last access: 5 May 2023, 2018.

Reitebuch, O., Lemmerz, C., Lux, O., Marksteiner, U., Rahm, S., Weiler, F., Witschas, B., Meringer, M., Schmidt, K., Huber, D., Nikolaus, I., Geiß, A., Vaughan, M., Dabas, A., Flament, T., Stieglitz, H., Isaksen, L., Rennie, M., de Kloe, J., Marseille, G.-J., Stoffelen, A., Wernham, D., Kanitz, T., Straume, A.-G., Fehr, T., von Bismark, J., Floberghagen, R., and Parrinello, T.: Initial assessment of the performance of the first Wind Lidar in space on Aeolus, EPJ Web Conf., 237, 01 010, https://doi.org/10.1051/epjconf/202023701010, 2020.

Rennie, M. and Isaksen, L.: The NWP impact of Aeolus Level-2B Winds at ECMWF, https://doi.org/10.21957/alift7mhr, 2020.

Rennie, M., Tan, D., Andersson, E., Poli, P., Dabas, A., De Kloe, J., Marseille, G.-J., and Stoffelen, A.: Aeolus Level-2B Algorithm Theoretical Basis Document (Mathematical Description of the Aeolus L2B Processor), https://earth.esa.int/eogateway/documents/20142/37627/Aeolus-L2B-Algorithm-ATBD.pdf, last access: 5 May 2023, 2020.

Rennie, M. P., Isaksen, L., Weiler, F., de Kloe, J., Kanitz, T., and Reitebuch, O.: The impact of Aeolus wind retrievals on ECMWF global weather forecasts, Quarterly Journal of the Royal Meteorological Society, 147, 3555–3586, https://doi.org/10.1002/qj.4142, 2021.

Santillan, D., Huber, D., Meringer, M., Reitebuch, O., Schindler, F., and Weiler, F.: VirES for Aeolus - Online visual analysis of Aeolus data, in: ESA Living Planet Symposium 2019, Milan, Italy, https://elib.dlr.de/130561/, last access: 8 May 2020, 2019.

Seifert, P., Radenz, M., Barja Gonzalez, B., Kalesse, H., Stratmann, F., Bühl, J., Teisseire, A., Vogl, T., Jimenez, C., Ohneiser, K., Schimmel, W., Wex, H., Ataei, F., Gong, X., Floutsi, A., Engelmann, R., Baars, H., Witthuhn, J., Ansmann, A., and Zamorano, F.: Aerosol, clouds, dynamics and their interaction over Punta Arenas, Chile (53°S, 71°W): A summary of two years of remote sensing and in-situ observations in the frame of DACAPO-PESO, in: 2020 AGU Fall Meeting, 2020.

Simonelli, G., Brandt, C., and Rezazad, M.: Aeolus First Year in Orbit Power System Performance, in: 2019 European Space Power Conference (ESPC), Juan-les-Pins, France, pp. 1–4, https://doi.org/10.1109/ESPC.2019.8932017, 2019.

Siomos, N., Gkikas, A., Baars, H., Wandinger, U., Amiridis, V., Paschou, P., and EARLINET-Consortium: Investigating the performance of AEOLUS L2A products over Europe with EARLINET ground-based lidars, in: EGU General Assembly 2021 (Online), European Geosciences Union, https://doi.org/doi:10.5194/egusphere-egu21-12460, 2021.

Stoffelen, A., Pailleux, J., Källén, E., Vaughan, J. M., Isaksen, L., Flamant, P., Wergen, W., Andersson, E., Schyberg, H., Culoma, A., Meynart, R., Endemann, M., and Ingmann, P.: The Atmospheric Dynamics Mission for Global Wind Field Measurment, Bull. Amer. Meteor. Soc., 86, 73–88, https://doi.org/10.1175/BAMS-86-1-73, 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. R. Meteorol. Soc., 132, 1927–1947, https://doi.org/10.1256/qj.05.83, 2006.

Straume, A. G., Rennie, M., Isaksen, L., de Kloe, J., Marseille, G.-J., Stoffelen, A., Flament, T., Stieglitz, H., Dabas, A., Huber, D., Reitebuch, O., Lemmerz, C., Lux, O., Marksteiner, U., Weiler, F., Witschas, B., Meringer, M., Schmidt, K., Nikolaus, I., Geiß, A., Flamant, P., Kanitz, T., Wernham, D., von Bismarck, J., Bley, S., Fehr, T., Floberghagen, R., and Parrinello, T.: ESA's space-based Doppler wind lidar mission Aeolus – First wind and aerosol product assessment results, EPJ Web Conf., 237, 01 007, https://doi.org/10.1051/epjconf/202023701007, 2020.

Tabary, P., Scialom, G., and Germann, U.: Real-Time Retrieval of the Wind from Aliased Velocities Measured by Doppler Radars, Journal of Atmospheric and Oceanic Technology, 18, 875 – 882, https://doi.org/10.1175/1520-0426(2001)018<0875:RTROTW>2.0.CO;2, 2001.

Tan, D. G. H., Andersson, E., De Kloe, J., Marseille, G.-J., Stoffelen, A., Poli, P., Denneulin, M.-L., Dabas, A., Huber, D., Reitebuch, O., Flamant, P., Le Rille, O., and Nett, H.: The ADM-Aeolus wind retrieval algorithms, Tellus A, 60, 191–205, https://doi.org/10.1111/j.1600-0870.2007.00285.x, 2008.

Tukiainen, S., O'connor, E., and Korpinen, A.: CloudnetPy: A Python package for processing cloud remote sensing data, Journal of Open Source Software, 5, 2123, 2020.

Weiler, F., Kanitz, T., Wernham, D., Rennie, M., Huber, D., Schillinger, M., Saint-Pe, O., Bell, R., Parrinello, T., and Reitebuch, O.: Characterization of dark current signal measurements of the ACCDs used on board the Aeolus satellite, Atmospheric Measurement Techniques, 14, 5153–5177, https://doi.org/10.5194/amt-14-5153-2021, 2021a.

950   Weiler, F., Rennie, M., Kanitz, T., Isaksen, L., Checa, E., de Kloe, J., Okunde, N., and Reitebuch, O.: Correction of wind bias for the lidar on board Aeolus using telescope temperatures, Atmospheric Measurement Techniques, 14, 7167–7185, https://doi.org/10.5194/amt-14-7167-2021, 2021b.

Witschas, B., Lemmerz, C., Geiß, A., Lux, O., Marksteiner, U., Rahm, S., Reitebuch, O., and Weiler, F.: First validation of Aeolus wind observations by airborne Doppler wind lidar measurements, Atmospheric Measurement Techniques, 13, 2381–2396, https://doi.org/10.5194/amt-13-2381-2020, 2020.

955   Witschas, B., Lemmerz, C., Geiß, A., Lux, O., Marksteiner, U., Rahm, S., Reitebuch, O., Schäfler, A., and Weiler, F.: Validation of the Aeolus L2B wind product with airborne wind lidar measurements in the polar North Atlantic region and in the tropics, Atmospheric Measurement Techniques Discussions, 2022, 1–32, https://doi.org/10.5194/amt-2022-233, 2022.

Wu, S., Sun, K., Dai, G., Wang, X., Liu, X., Liu, B., Song, X., Reitebuch, O., Li, R., Yin, J., and Wang, X.: Inter-comparison of wind measurements in the atmospheric boundary layer and the lower troposphere with Aeolus and a ground-based coherent Doppler lidar network over China, Atmospheric Measurement Techniques, 15, 131–148, https://doi.org/10.5194/amt-15-131-2022, 2022.

960   Zuo, H., Hasager, C. B., Karagali, I., Stoffelen, A., Marseille, G.-J., and de Kloe, J.: Evaluation of Aeolus L2B wind product with wind profiling radar measurements and numerical weather prediction model equivalents over Australia, Atmospheric Measurement Techniques, 15, 4107–4124, https://doi.org/10.5194/amt-15-4107-2022, 2022.