Response to Referee Comment (RC2) on

Airborne wind lidar observations over the North Atlantic in 2016 for the pre-launch validation of the satellite mission Aeolus (https://doi.org/10.5194/amt-2018-19)

We are grateful for the referee's very valuable and helpful comments on our manuscript AMT-2018-19 which will certainly improve its quality and readability for a broader audience. According to the suggestions and questions, the below aspects will be discussed in more detail in a revised version of the paper.

Comment #1:

Page 3, line 30: Given that other characteristics of the transmitter are stated, it would be useful to know the bandwidth or frequency uncertainty of the transmitted pulse.

Response to Comment #1:

The bandwidth of the transmitted laser pulse (50 MHz) and the frequency stability (≈ 3 MHz rms, UV) will be mentioned in the revised manuscript. More detailed information on the frequency stability of the laser transmitter is provided in Lemmerz et al. (2017).

Comment #2:

Page 4, line 3: Why are the receiver electronics triggered 60 microseconds before laser pulse emission?

Response to Comment #2:

The long lead time of the detector electronics is due to an electronic preconditioning process of the accumulating charged coupled device (ACCD) arrays which require a trigger signal being provided 61.4 µs prior to each laser pulse acquisition. The same ACCDs are used for the satellite instrument, but here the long round-trip laser pulse travel time from the satellite to the first atmospheric range gate (≈ 2.5 ms) do not cause an issue from to the long preconditioning process.

Comment #3:

Page 4, line 25: It wasn't clear to me exactly what the EOM does. Does it simply switch between the internal reference and the atmospheric signal after the pulse has been transmitted?

Response to Comment #3:

The EOM in the front optics is used to avoid saturation of the ACCD by blocking the atmospheric path for several µs after transmission of the laser pulse, thus preventing strong backscattered light produced close to the instrument (up to about 1 km) from being incident on the detectors. In this way, the EOM separates the internal reference signal, which is guided to the receiver via a multimode fiber and registered in range gate #4, from the atmospheric signal, which enters the receiver on a free-space path including the EOM and is registered in range gates #5 to #24. The EOM is specific to the airborne instrument and not needed on the satellite.

Comment #4:

Page 8, line 18: It would be useful to know why the response function for the satellite instrument is only performed for a single range gate. How is that gate determined? Is the satellite attitude varied so that instrument looks vertically?

Response to Comment #4:

For the satellite instrument the atmospheric Rayleigh response function is derived after adding the return signals obtained from a number of range gates in the upper troposphere (e.g. between 6 km and 16 km) in order to increase the signal-to-noise ratio. The selection of the appropriate range for averaging is performed during on-ground processing and the information for each single range gate is still included in the downlinked raw data. In the satellite wind retrieval for the L2B product, a Rayleigh-Brillouin line shape model is used in combination with atmospheric temperature and pressure profiles from a NWP model (e.g. from ECWMF) to account for the altitude-dependence of the Rayleigh response over the entire vertical measurement range from ground to the lower stratosphere (Dabas et al., 2008). Hence, in contrast to the A2D, only one set of Rayleigh response calibration parameters is determined for a large vertical range covering multiple range gates. For the A2D a vertical profile of Rayleigh response parameters can be determined, because the SNR is sufficiently high for the airborne instrument. In this regard, the sentence on page 8 “For the satellite instrument, the response function is derived for only one atmospheric range gate” is indeed misleading and will be paraphrased accordingly in the revised manuscript.

As for the second question, the satellite instrument, like the A2D, is operated in nadir-pointing geometry during the response calibration procedure. For this purpose, the whole satellite is rolled by 35°.

Comment #5:

Page 9, line 20: I found this paragraph a bit confusing. What does "a visual inspection of intensities" mean, and why does a summation lead to an underestimation of the actual ground signal? Because the paper places quite a bit of emphasis on the improvement of the new technique for dealing with ground return, I think a better characterization of the old technique is necessary.

Response to Comment #5:

The old ground detection scheme was based on a rough analysis of the curtain plot depicting the Rayleigh and Mie signal intensities after range-correction and normalization to the integration time of each range gate, as shown for the response calibration #6 in Fig. 4(a) and (b). Here, high signal intensities related to strong ground return become visible as white bins, as the intensity exceeds the maximum of the respective colour scale. For this particular example, range gates #21 to #23 would have been (subjectively) selected as ground range gates in the old scheme (by visual inspection by an experienced data analyst), since most of the white bins are found in these three range gates. Only ground signals contained therein would be summed up. However, the new ground detection method based on the signal gradient approach reveals that ground signals are also contained in adjacent range gates, as displayed in the Rayleigh and Mie ground masks in Fig. 4(c) and (d). For instance, the Mie ground mask features two white bins in range gate #20. In other calibrations or wind scenes, the number of disregarded ground bins can be much larger. Thus, the old ground detection scheme generally involves an underestimation of the actual ground signal, unless a very ample selection of ground range gates is performed. However, the latter approach would significantly increase the atmospheric contribution to the summed ground signal, especially in case of varying ground elevation during the investigated response calibration or wind scene. It should also be noted, that in previous airborne campaigns, the old ground detection scheme was acceptable, since the response calibrations were performed over flat terrain, e.g. sea ice, so that ground signals were almost completely contained in only one range gate. The differences between the two ground detection methods will be elaborated as above in the revised manuscript.

Comment #6:

Page 10, line 31: I assume that delta h is computed from the DEM data for Table 2, although it isn't totally clear to me from the text.

Response to Comment #6:

Correct. As described on page 9, lines 15ff., ΔH is the "height difference between a reference ground elevation during one measurement and the upper bin border of the highest (or first) range gate that contains ground signals […]. The reference ground elevation per measurement is derived from the digital elevation model (DEM) ACE2, providing elevation data at a resolution of 9 arc seconds (300 m x 300 m at the equator) (Berry et al., 2010)."

Comment #7:

Page 16, line 25: Because the coherent Doppler lidar measures return from aerosols, and the Rayleigh channel is sensitive to the presence of aerosols, is a comparison between the two valid? Doesn't this potentially overestimate the error in the Rayleigh winds, unless the effect of aerosol on that channel is negligible? Perhaps this should be discussed some more.

Response to Comment #7:

We agree with the referee, that the comparison of A2D Rayleigh winds with the 2-µm DWL is limited to atmospheric regions, where cloud and aerosol backscattering occurs. However, it should be noted that the coherent 2-µm DWL is very sensitive even to weak particulate backscatter return, due to its coherent detection principle with small bandwidth. In addition, since the coherent DWL is deployed on the aircraft, the atmospheric altitudes with low aerosol backscatter are located in near range gates, which do not suffer from the R^2-dependency of the signal and strong aerosol extinction (as it would be the case for ground-based coherent DWL). Hence, 2-µm DWL winds are even available for low scattering ratios (<1.1), where no significant aerosol-contamination of the A2D Rayleigh winds can be expected. Furthermore, as described in section 4.1.2, Mie-contaminated bins in the Rayleigh data are identified by a signal threshold approach and excluded from the Rayleigh wind curtain. Such range bins thus do not enter the statistical comparison with the 2-µm DWL winds. Also, the Rayleigh winds are only considered as valid (and enter the statistical comparison with the coherent DWL) in case that no valid winds are detected from the A2D Mie channel by using a SNR threshold on the Mie channel. Thus, we agree that valid A2D Rayleigh winds could be contaminated by narrowband aerosol backscatter, but we consider this effect as not dominating the systematic and random error.

Comment #8:

Page 17, line 1: The removal of the "bad" measurement from the Rayleigh dataset is done based on comparison with the coherent lidar. How would this be handled for the satellite measurement, where no comparison source will be available, to avoid sending bad data to the assimilation algorithms?

Response to Comment #8:

The removal of "outliers" from the comparison dataset is performed to obtain a probability density function of wind speed differences (Fig. 9 (b) and (c)) which is closer to a Gaussian distribution. In this way, the standard deviation represents a good measure of the random error. However, it is certainly necessary to state the number or percentage of outliers removed from the dataset. We consider "outliers" or "gross errors" as being uniformly distributed over the wind speed measurement range which add to the Gaussian-distributed random errors. Indeed, the error model for Aeolus, as described in the Mission requirement document MRD (ESA, 2016) does separate these two different errors and defines a requirement for Aeolus on the probability of gross outliers (< 5%). Hence, we think that this approach for a statistical comparison is justified. For the satellite L2B data products, the estimation of the random error is provided for each observation and could be used as additional QC parameter. In addition, NWP centres usually apply a QC (or even variational QC) during the assimilation of the wind products by comparing it with best guess values (background) from the model.

Comment #9:

Page 18, line: 23: Why exactly does a heterogeneous cloud structure contribute to random error? Is it a range-weighting effect in the presence of shear, or some other optical effect? Probably a bit more explanation is needed here.

Response to Comment #9:

Yes, it is the combination of the presence of a strong backscatter gradient (e.g. cloud boundaries) and strong wind shear which is not resolved due to the coarse vertical resolution of A2D and Aeolus. It could be also referred as a Mie height assignment error. As stated in the text, the impact of heterogeneous cloud structure on the Mie random error is comprehensively explained in Sun et al. (2014). The authors found that assigning the measured wind to the centre of the measurement bin introduces an error when particles and/or molecules are not uniformly distributed inside the bin, which is generally the case. Hence, the error originates from a range-weighting effect and is particularly large in the presence of shear in the sampling volume. Assuming a constant wind-shear with typical amplitude of 0.01 s^{-1} over the bin, the random error of the Mie wind scales inversely proportional with the thickness of a particle layer randomly positioned inside the bin, reaching 2 $m \cdot s^{-1}$ for a bin size of 1000 m and a layer thickness of 300 m.

Comment #10:

Figure 4: The caption refers to "orange bins" in (c) and (d). I’m not sure what bins this refers to; I don’t see any orange bins.

Response to Comment #10:

In a previous version of the manuscript, the ground bins in Figure 4(c) and (d) were indicated in orange. The colour of the bins was changed to white without adapting the figure caption. The caption will be corrected accordingly in the revised version. Sorry for the inconvenience.