

## **General comments:**

The manuscript of Iwai and coauthors reports the comparison of Aeolus HLOS winds over Japan with wind profiler measurements, two ground-based doppler wind lidar and GPS radiosondes. The study is well structured and is found as an important contribution to the Aeolus special edition. The technical set up and statistical methods are comprehensively described. It is appreciated that the systematic and random error estimations are provided for three independent reference data sets and that the results are discussed with respect to other Cal/Val studies. The topic of the submitted paper fits very well to the scientific purposes of the AMT and can be published after addressing some minor comments and suggestions which are listed below.

## **Specific comments:**

### Abstract:

For an easier readability, I would suggest to round numbers to one decimal place in the abstract

Line 23 – 25: Not well understood. Was the bias during the 2B10 period negative (Rayleigh) and gradually fluctuating (Mie) or was it close to zero? Please try to formulate the message more clearly.

### Introduction:

The first paragraph of the introduction needs to be revised. The order of statements of the first four sentences (Line 38 – 42) seems illogical. I suggest to start with the importance of accurate NWP for agriculture, transportation, etc. and then explain the importance of wind field measurements for NWP and furthermore also for climate studies, air quality monitoring, etc. The sentences in Line 43 –44 are kind of misleading. Are you here referring to the entire global observing network or do you mean, that measurements such as radiosondes, wind profilers (WPRs), ground-based Doppler wind lidars (DWLs), and aircrafts provide accurate vertical profiles and that they are limited from the global perspective? If the latter is the case you have to distinguish between single level aircraft measurements at flight level and aircraft ascents. In the following part I suggest to make the limitations of wind measurements from satellite radiances and AMVs more clearly. I miss the information that satellite radiances only provide information about the mass field which leads to particularly strong restriction in the absence of geostrophic balance. The limited accuracy of AMV single-level winds is mainly caused by significant systematic and correlated errors due to uncertainties of their height assignment (see., Folger and Weissmann, 2014 or Bormann et al., 2003).

Line 50 - 51: Please be precise - Aeolus is the name of the earth explorer mission/satellite and not the name of the DWL. Regarding the following sentence this can be confusing.

Line 57 – 59: For me the expression ‘and so forth’ implies that further purposes are obviously. Is this really the case? Furthermore, I suggest to put some reference here (e.g. ESA, 1999: The four candidate Earth Explorer core missions - Atmospheric Dynamics Mission or ESA, 2001: ADM-Aeolus Mission Requirements Document)

Since some phrases are very similar to Belova et al. 2021, I suggest to add this study to your reference

### Section 3:

I suggest to already mention that you use the vertical range bins of Aeolus as vertical collocation criteria in the paragraphs about the horizontal and temporal collocation criteria when comparing the Aeolus winds to WPRs, CDWLs and RSs. This is a question the reader may asks himself at this point.

In section 2 about the Aeolus data and in the abstract, you mention the three validation periods, whereby you write that the first period is from 1 October to 18 December. In Section 3 you write that you are using the WPRs and CDWLs for the comparisons from 18 October 2018 to 15 and 11 May 2019. In table 4 you also have the 20 December 2018 as validation day for comparisons with radiosondes. So, you also used the Aeolus data with baseline 2B02 till May 2019? Please correct for this consistently in the whole paper.

### Section 4:

Line 196: Please shortly describe how the Aeolus azimuth angle is defined.

Line 213: Maybe add some additional validation studies like Lux et al., 2020 or Martin et al.,2021 as reference.

## Section 5:

Numbers are sometimes rounded to two decimal places, sometimes to only one decimal places. As your plots and tables are showing the statistical results which are discussed in section 5 as numbers rounded to two decimal places, I suggest to also do so in the text

Line 261: Actually, you cannot compare your estimates of systematic and random errors for HLOS winds with the values of Lux et al. 2020. The A2D measures LOS with off-nadir of  $20^\circ$  so that it is not possible to calculate HLOS (in Aeolus direction). Furthermore, the azimuth directions are not identically. The paper concentrates on a comparison in LOS space which is not equivalent to comparisons in HLOS space. You have to multiply it with a factor of  $1/\sin(37)$ , see section 4.2 of Lux et al. 2020.

In the first part of your paper (section 2), you stress that you are investigating three different validation periods. However, in section 5.1.1 and 5.1.2 the reader may wonder why you are only talking about two, the 2B02 and 2B10, periods. Aren't there any differences in the statistics between the reprocessed data set and the data from April 2020 on? Shortly explain why you are combining the two time periods, refer to section 5.1.3 where you are showing the time series or consider to just talk about two validation periods in section 2 – the 2B02 and the 2B10 period while the 2B10 period is composed of the M1 bias corrected observations and the reprocessed data set.

In the end of section 5.1.1 you could point out that the reduced bias of the 2B10 period compared to 2B02 is most likely due to the M1 bias correction.

Line 313: There is a difference in bias for Rayleigh-clear 2B10 between ascending and descending of 0.6 m/s. Is this really not significant?

Figure 6 and 7: A second axis with the number of compared data points could be nice. But this is only a suggestion.

In section 5 you are sometimes disarranging tenses. For example, line 329 and line 330 should be in present like in the sentences before and after. Please look through your paper once again carefully with focus on tenses.

Line 332: This statement kind of confuses the reader at this point. Improved compared to what?

Section 5.1.2, Figure 7: I find it prominent that the bias for is tending to get more negative with height, while for ascending it is kind of the other way around. The negative trend with height for the descending HLOS winds is also visible in Figure 6. Maybe consider to mention this.

Line 349: Towards the end of 2B02 period (after April 2019) it looks more like a negative trend - especially for descending orbit.

Line 353: I suggest to add a reference about the laser performance here:

Reitebuch et al., (2020b), Assessment of Aeolus performance and bias correction results from the Aeolus DISC, Aeolus Cal/Val and Science Workshop 2020, <https://nikal.eventsair.com/QuickEventWebsitePortal/2nd-aeolus-post-launch-calval-and-science-workshop/aeolus>

Reitebuch, Oliver, Christian Lemmerz, Oliver Lux, Uwe Marksteiner, Stephan Rahm, Fabian Weiler, Benjamin Witschas, et al. (2020). "Initial Assessment of the Performance of the First Wind Lidar in Space on Aeolus". Edited by D. Liu, Y. Wang, Y. Wu, B. Gross, and F. Moshary. EPJ Web of Conferences 237: 01010. <https://doi.org/10.1051/epjconf/202023701010>

Also consider, that changes in the Aeolus range bin settings can lead to changes in the random error.

Line 354: It is the same for the descending orbit. Furthermore, you already mentioned the overall negative bias for Rayleigh winds in line 351. So, maybe consider to just remove this sentence.

Line 358: In fact, the Mie signal does depend on the laser energy, I think. But it depends also on the presence of aerosols or hydrometeors (see Martin et al. 2020 end of section 3.1.). That's why the effect is not visible in the random error that prominent compared to the Rayleigh winds. Please correct me if I am wrong.

Line 360: Approached zero towards September 2020, in my view.

Line 370: I don't really see an increase in Rayleigh bias with larger scattering ration. Isn't the bias more fluctuating and almost the same for scattering ration 1.1 and 1.4?

For the profiles of the comparisons of Aeolus with radiosondes (Fig. 13 and 14 (a),(b)), it would be nice to have the information of orbit phase in the caption.

Line 457: In my opinion the Mie bias is not almost the same as the Mie bias obtained by Baars et al 2020. Its 1.2 m/s larger. Probably due to small number of radiosonde launches of Baars et al, different location and meteorological conditions or different distance between the measurements?

Line 477 – 479: The sentence is not completely right in my opinion. Martin et al 2021 estimated the representativeness error by considering different error sources, not the representativeness error and error sources. In your case the representativeness errors result from the different measurement geometries and from the collocation criteria (spatial and temporal displacement).

Since you use the scaled MAD as estimate for the random error in your paper, I suggest to name the sigma Aeolus 'Aeolus observational error' instead of 'Aeolus random error'. Otherwise, it could be misleading.

### Summary

Line 502: You say that you investigated the bias dependence on latitude, but there is no such part in the manuscript.

Line 521: Again, in my opinion the bias only is getting close to zero towards September 2020. In the early summer months 2020 it is as close to zero than in 2019.

I miss a nice closing sentence to round off the whole study. You could mention that despite you used three independent data sets with quite different sample sizes the differences in bias estimation are not greater than 1 m/s (except of Rayleigh 2B10). What's your conclusion of the good concordance between the three independent validation data sets? What's the importance of this study for the Aeolus mission and the use of the Aeolus HLOS winds in NWP? You can also give kind of an outlook or suggestions for improvements of your study.

### **Technical corrections:**

Line 9: check spelling 'onboard'

Line 16 - 18: Avoid duplication. Two sentences with the same message. I would suggest the following wording: "The statistical comparisons for the baseline 2B10 period show smaller biases,  $-0.8 - 0.5 \text{ m s}^{-1}$  for the Ryleigh clear and  $-0.7 - 0.2 \text{ m s}^{-1}$  for the Mie cloudy winds."

Line 57: please correct 'purposes'

Line 60: please correct 'requires'

Line 101 – 106: for better understanding consider to rephrase like this: "In this study, we used three different periods during the processor baseline 2B02 and 2B10 periods to assess L2B data products: 1 October to 18 December 2018 (2B02), 28 June to 31 December 2019 (2B10) and 20 April to 8 October 2020 (2B10). The first period with baseline 2B02 was within the commissioning phase. The L2B data products with the 2B10 baseline include a bias correction for ALADIN's telescope primary (M1) mirror temperature variation (Rennie and Isaksen, 2020; Rennie et al. 2021\*) and have been available for new observations since April 2020. The L2B winds from 28 June to 31 December 2019 are a homogeneous reprocessed dataset using also the 2B10 processor version." \*I suggest to add the reference Rennie et al. 2021 for more information about the M1 bias correction.

Line 126: please add a space between number and unit (294 m)

Line 198: it seems that the spacing between the equation environment and the text is missing

Line 205: consider to rephrase as follows: "...estimated HLOS errors are 2.3 m/s during both, baseline 2B01 and 2B10."

Line 228: it seems that the spacing between the equation environment and the text is missing

Line 267 and 292: consider to remove the term 'large' to not sound judgmental

Line 294: almost same sentence than Line 271. I suggest to rephrase as follows: "Again, the discrepancies may be caused by ..."

Line 356: wrong figure reference - Fig. 9d not Fig 9c

Line 440: please modify: "...the reason for that.."

Line 440 – 441: consider to rephrase: "Potentially, large horizontal wind gradients in this height region have an influence on the differences.

Line 482: consider to rephrase; "The Aeolus observation error considering the representativeness error in addition to the radiosonde observational error can be calculated as follows:..."

Line 503 and 511: I think there is no need for a paragraph here. Rather, I suggest to put a paragraph before you start to summarize the WPR validation in line 500, like you also did before the CDWL validation and before the radiosonde validation summary. That seems more conclusive for me.

Line 504 – 505: could remove this sentence. It's kind of repetition of the sentence in line 499 – 500.

Line 513: be careful with using the expression significant. Actually, this should be based on a statistical significance test. Better remove 'significantly' here.

Line 518 – 519: this does not make sense. Either "...the bias of Rayleigh-clear HLOS winds were generally negative..." or "...the systematic differences between the Rayleigh clear and WPR HLOS winds were generally negative..."

Line 520: consider to modify: "... and did also not show a clear seasonal trend."

Line 533: next line before the minus symbol

Line 534: again, maybe reconsider the use of significantly here