

Anonymous Referee #2

Referee comment on "Validation of the Aeolus Level-2B wind product over Northern Canada and the Arctic" by Chih-Chun Chou et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2021-247-RC2>, 2021

We thank the reviewers for their careful reading of the manuscript and comments, which helped us improve the study.

General comments:

The authors compared the Aeolus wind measurements with Ka-radar, wireless soundings, ECCO, and ERA5 reanalysis data, and the results show that the Aeolus wind measurements are in good agreement with other wind measurements or reanalysis data. This paper has important implications for the application of Aeolus wind measurements in the Arctic, where wind measurements are currently scarce. However, the paper requires significant revision before publication, and the specific issues are described below.

1. The authors use other soundings and reanalysis data as a benchmark for comparison with the Aeolus wind data, but the data quality of the other data is not presented in the paper. Perhaps the other wind data also have large biases in the Arctic, and then the authors' use of them as a comparison benchmark would make none of the comparisons in this paper credible. Especially for high-altitude wind fields, the data sets may have different data quality performances at different heights. Moreover, the reanalysis dataset itself contains assimilation of existing sounding data. The authors need to fully justify the reliability of the data in each of the datasets used in this paper.

Thank you for pointing this out. To improve this aspect, we added a few more references on the validation of the datasets used in the study. For the radiosondes, Dirksen et al. (2014) described the uncertainty for the wind speed and Laroche and Sarrazin (2013) described the errors due to the radiosonde drifts. The following passage is added in Line 191 for clarification:

"Vaisala RS92 radiosondes (Mariani et al., 2018) were launched twice daily (45 minutes before synoptic times 00 and 12 Coordinated Universal Time (UTC)). They measure vector wind profiles with a vertical resolution of roughly 15 m depending on ascent speed, up to about 30 km above ground level. The data used (available at <http://weather.uwyo.edu/upperair/sounding.html>) is the processed radiosonde data provided at mandatory and significant pressure levels (which has a coarser resolution than 15 m). It takes about two hours to

reach 30 km altitude (around 10 hPa). The instrumental uncertainty for the wind speed is between 0.4 and 1.0 ms^{-1} and between 0.3 and 0.7 ms^{-1} for the zonal wind component (Dirksen et al., 2014). The error on the zonal wind component due to drift and elapsed time of the ascending balloon is between 0.5 and 1.0 ms^{-1} in the troposphere and UTLS (see Fig.5b in Laroche and Sarrazin, 2013). As a result, the total error for the zonal wind component from these sources of errors is between 0.6 and 1.2 ms^{-1} . Note that the radiosonde data are assimilated in the ECCO and ECMWF systems, which means that the ECCO-B and ERA5 errors are not independent of the radiosonde observation errors.”.

As for the data quality for in-situ measurements, Mariani et al. (2018) demonstrates reliable observations from the Ka-band radar and its application, and Mariani et al. (2020) validates the ground-based lidar to the radiosondes. A few more sentences are added in line 209:

“The uncertainty of the measurements depends on conditions, SNR, and decibel of the return signal. The average vertical wind profile bias to radiosonde is better than 0.27 ms^{-1} .”.

Furthermore, it is true that ECCO-B and ERA5 contain assimilation of existing sounding data; however, they also assimilate many other observations, for instance, from the satellites, aircraft measurements, and more. Therefore, we think that they are still credible sources of wind data to validate Aeolus winds.

2. The authors' presentation of the spatio-temporal matching process between the datasets is not clear enough. It is suggested that the data matching process be introduced as a separate section and the information about the data used in this paper be summarized. A clear and detailed data matching process is desired to be seen.

Thank you for these suggestions. Although we considered introducing an additional section on this, we thought it would be best to revise Sect. 2.5 on the data matching and coincidence criteria. In doing this, we sought the same level of detail as provided by other publications including Belova et al. (2021), Martin et al. (2021), and Baars et al. (2020). Section 2.5 is revised to:

“For the ground-based validation, the criterion for coincidence of Aeolus overpasses is that the distance from the sites to the measurements be no more than 90 km (horizontal resolution of Rayleigh winds). Using this coincidence criterion, Aeolus overpasses are selected as targets for validation at Iqaluit three

times a week at around 21:50, 11:15, and 22:00 UTC, and at Whitehorse twice a week at around 02:25 and 15:30 UTC. The Aeolus measurements are compared to the reanalysis and in-situ measurements that are available in the nearest time. Temporal sampling for each product is as follows: Aeolus overpasses at Iqaluit and Whitehorse are as mentioned above; reanalysis data is provided hourly, on the hour; radiosonde data is from launches at 00 and 12 UTC, with a two-hour time-of-flight to 30 km as mentioned above; Ka-band radar data is provided via 15-minute scans. For example, if Aeolus overpasses selected as a target for validation at the Iqaluit site at 11:15 UTC, since the reanalysis data is sampled hourly, the radiosondes are launched at 00 and 12 UTC, and the Ka-band radar at Iqaluit scans every 15 minutes, the Aeolus HLOS profile would be compared to the reanalysis data and radiosonde measurements at 12 UTC and to the nearest scan by the radar. On the other hand, if the overpass time is 02:25 UTC, the profile would be compared to the ERA5 data at 02 UTC, the radiosonde measurements at 00 UTC, and, again, the nearest scan by the radar.”.

3. the authors' analysis of Figure 5 is not rigorous enough and the conclusions drawn from Figure 5 are not reliable enough, see Specific comments L296.

Please see the response below.

4. The authors need to clarify the practical significance of the discussion of the statistical distribution of the wind products themselves in Figures 7 and 8. It might be more valuable to discuss instead the distribution of the difference between the Aeolus data and the other data.

Thank you for the suggestions. Regarding Fig. 7, we mainly want to show the Aeolus sampling in different atmospheric layers and that the decomposition into different wind-component directions provides insight into understanding the meteorological conditions that the measurements are sampling, which might be helpful to better understand the dynamical characteristics of this data in both Aeolus and other products. We have slightly modified the text to better explain this analysis (line 375):

“Furthermore, some ascending and descending HLOS wind measurements cancel in the average owing to simply to the change of the angle of the LOS. To avoid this artefact and to add some insight into the wind features being measured, we also compare the projected HLOS wind vector into its zonal (positive to the east) and meridional (positive to the north) components. The

distribution of the zonal-component of the HLOS winds is shown in Fig. 7e and g for Aeolus and ECCC-B HLOS winds. By doing this decomposition, the distributions for ascending and descending measurements are brought into better agreement (Fig. 7f). We also notice that the HLOS winds can provide some information about the vertical variation of the HLOS winds that are projected onto the zonal direction (Figs. 7e and g). For example, for Aeolus the projection of HLOS into the zonal direction for the stratosphere, UTLS, and troposphere are +11.00 ms⁻¹, +4.00 ms⁻¹ and +1.00 ms⁻¹ respectively for this measurement period and these values (and the standard deviations of their distributions, see the figure legend for values) agree very well with ECCC-B (and ERA5 – not shown). The distributions have mean values that are positive because the winds are mainly westerly over the Arctic in the winter.”

Furthermore, we agree that it would be valuable to discuss the distribution of the difference between the Aeolus data and the other data. Thank you for the suggestion! Therefore, Fig. 8 now shows the means and standard deviations of the differences between Aeolus and ECCC-B and ERA5. The means of the differences therefore reflect the remaining bias between the datasets after the dynamic bias correction has been applied. The associated paragraph describing Fig. 8 is also revised. Starting at line 394:

“We compare the distributions of the differences between the Aeolus wind measurement data and the ECCC-B and ERA5 data during fall 2018, summer 2019, and winter 2020 over the Arctic, as summarized in Fig. 8, which shows the bias and standard deviations of the differences between Aeolus HLOS winds and the ECCC-B HLOS winds, and ERA5 HLOS winds, and their zonal and meridional projections. The measurements are decomposed into Rayleigh (red) and Mie winds (black). They are further decomposed into ascending (indicated with upright triangles) and descending (inverted triangles) measurements. The results, with the bias (the mean values of these differences for the different sampling used) being smaller than 0.7 ms⁻¹, are consistent with our bias correction method. The distributions of the differences in the ascending and descending measurements do not show a significant difference. The discrepancies in the meridional projections of the HLOS winds are smaller because Aeolus picks up mostly the zonal component of the winds due to the direction of the LOS.”

Specific comments:

L82:1 December to 31 January 2020

Wrong time markings

Thank you, fixed.

L115: ECMWF has recently published the first reprocessed data (2B10)

Data links should be added after the data introduction.

Thank you, the link is added in Line 114: "ECMWF has published the first reprocessed data in fall 2020 (2B10; available at ftp://2018_aeolus_l2b:ecmwf@acquisition.ecmwf.int/), which covers the period between 24 June and 31 December 2019."

L125: The quality control recommendation following the Guidance for Aeolus NWP Impact Experiments (Rennie and Isaksen, 2019), including the threshold for L2B estimated observation errors.

The recommendation given by NWP is a threshold range. Authors should submit specific thresholds to be used when processing data.

Thank you for the comment. The threshold values are added in the text (line 125):

"The thresholds for L2B estimated observation errors during the FM-A period are 4.5 ms⁻¹ for the Mie winds and 6.6 to 11 ms⁻¹ for the Rayleigh winds, depending on the pressure level, and 5 ms⁻¹ for the Mie winds and 8.5 to 12 ms⁻¹ for the Rayleigh winds during the early FM-B period. For more details, please refer to Rennie and Isaksen (2020)."

L129: We further reject the outliers by excluding all the data when the difference between the observations and ECCC-B or ERA5 is greater than 30 m/s.

For the screening threshold of 30m/s, the authors need to give an explanation, either from data analysis or literature.

Thank you, the following explanation is added in Line 130: "This criterion was obtained from an initial comparison between Aeolus FM-A and ECCC-B (Laroche et al., 2019).".

L162: ECCC-B is then linearly interpolated to Aeolus measurement locations and times.

The process of linear interpolation needs to be clarified. In addition, the main comparison data in the latter section does not provide complete information on the vertical resolution, and due to the large number of datasets used in this paper, it is recommended to use a table after section 2.5 to summarize the important information of each dataset used.

Thank you for the comments. This passage is added in line 163 for clarification of the process of linear interpolation of ECCC-B:

“The data used to compare with Aeolus winds in this paper is the assimilated data that is linearly interpolated to Aeolus measurement locations and times. For the linear interpolation between the model’s grid points, the horizontal grid-spacing is 15 km and the vertical grid-spacing varies from approximately 100 m in the PBL to 1 km in the stratosphere (McTaggart-Cowan et al., 2019). The linear interpolation in time is between two consecutive model states, 15 min apart.”

We revised Sect. 2.5 (line 229) instead of adding a table after the section. It summarizes the temporal resolution of each dataset and clarifies the data matching process and coincidence criterion:

“For the ground-based validation, the criterion for coincidence of Aeolus overpasses is that the distance from the sites to the measurements be no more than 90 km (horizontal resolution of Rayleigh winds). Using this coincidence criterion, Aeolus overpasses are selected as targets for validation at Iqaluit three times a week at around 21:50, 11:15, and 22:00 UTC, and at Whitehorse twice a week at around 02:25 and 15:30 UTC. The Aeolus measurements are compared to the reanalysis and in-situ measurements that are available in the nearest time. Temporal sampling for each product is as follows: Aeolus overpasses at Iqaluit and Whitehorse are as mentioned above; reanalysis data is provided hourly, on the hour; radiosonde data is from launches at 00 and 12 UTC, with a two-hour time-of-flight to 30 km as mentioned above; Ka-band radar data is provided via 15-minute scans. For example, if Aeolus overpasses selected as a target for validation at the Iqaluit site at 11:15 UTC, since the reanalysis data is sampled hourly, the radiosondes are launched at 00 and 12 UTC, and the Ka-band radar at Iqaluit scans every 15 minutes, the Aeolus HLOS profile would be compared to the reanalysis data and radiosonde measurements at 12 UTC and to the nearest scan by the radar. On the other hand, if the overpass time is 02:25 UTC, the profile would be compared to the ERA5 data at 02 UTC, the radiosonde measurements at 00 UTC, and, again, the nearest scan by the radar.”

L165: Reanalysis ERA5

For the ERA5 dataset, assimilation relationships between it and the other datasets used in this paper should be added.

Thank you. The following passage is added in Line 195 for clarification: "Note that the radiosonde data are assimilated in the ECCO and ECMWF systems, which means that the ECCO-B and ERA5 errors are not independent of the radiosonde observation errors."

ERA5 only contains assimilation of radiosonde data and other observational data, but not the in-situ measurements used in the study. Therefore, even though ERA5 errors are dependent to the radiosonde observation errors, it is still worth comparing both datasets since ERA5 contains information from many other observations as well. An example of other study that uses both datasets is Chen et al. (2021).

Reference: 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, *Atmos. Chem. Phys.*, 21, 11489–11504, <https://doi.org/10.5194/acp-21-11489-2021>, 2021.

L190-L210:

From the latter, there is very little overlap between Ka-band radar or LIDAR with Aeolus data. The main object of comparison is the radiosondes, so its data should have been presented in more detail. In particular, the horizontal drift problem of radiosondes, which seems to be not taken into account by the authors, may also lead to a large bias. In addition, is the introduction and use of Ka-band radar and ground-based LIDAR data in this paper meaningful? It seems that the absence of these two does not affect the logic and conclusions of this paper.

We acknowledge the little overlap between Ka-band radar or lidar with Aeolus data; however, they are still valuable because they are entirely independent from the Aeolus winds. Unlike the radiosondes, they are not assimilated in ERA5 and ECCO-B which are used in the correction of Aeolus winds and quality control process. Thus, the comparison would still add some value to this validation work. The retrieval of the radar wind profile is included because we have some comparison between Aeolus and radar in Figs. 3 and 4. The Aeolus wind is not compared to the ground-based is not

included, so its retrieval is not explained in the paper. However, since its measurements are shown in Fig. 2, a brief introduction to the instrument is still included.

As you mentioned, the horizontal drift problem of radiosondes has not been considered. So, we added more details on the data quality of radiosondes in Line 191 for clarification:

“Vaisala RS92 radiosondes (Mariani et al., 2018) were launched twice daily (45 minutes before synoptic times 00 and 12 Coordinated Universal Time (UTC)). They measure vector wind profiles with a vertical resolution of roughly 15 m depending on ascent speed, up to about 30 km above ground level. The data used (available at <http://weather.uwyo.edu/upperair/sounding.html>) is the processed radiosonde data provided at mandatory and significant pressure levels (which has a coarser resolution than 15 m). It takes about two hours to reach 30 km altitude (around 10 hPa). The instrumental uncertainty for the wind speed is between 0.4 and 1.0 ms^{-1} and between 0.3 and 0.7 ms^{-1} for the zonal wind component (Dirksen et al., 2014). The error on the zonal wind component due to drift and elapsed time of the ascending balloon is between 0.5 and 1.0 ms^{-1} in the troposphere and UTLS (see Fig.5b in Laroche and Sarrazin, 2013). As a result, the total error for the zonal wind component from these sources of errors is between 0.6 and 1.2 ms^{-1} .”

L246: On 24 September, Aeolus measures westerly winds in reasonable overall agreement with the other data.

The agreement between Aeolus and the other data in Figure 2b does not seem to be obvious, and the deviation in some data points is already close to 50%.

Thank you for pointing this out. We agree that the agreement is not that obvious.

The sentence (line 266) is revised: “On 24 September, although a few of the Rayleigh measurements have a deviation close to 50%, Aeolus still measures westerly winds in reasonable overall agreement with the other data.”.

Figure 2 is also revised. There were two Aeolus measurements at the same level due to the collocation criteria. Figure 2 is now showing the nearest profile from Aeolus to the sites only.

L249:

Figures 3 and 4 are identical and need to be modified. The comparison of Ka-band radar in Figure 3d needs to be addressed for its significance since there are only about 10 data points. Also, the information on the fitted straight lines, standard deviation, and sample size in the figure can be added, while the information in the supplemental Figure S1 will be placed in the original text in the form of a table.

Thank you for pointing out this error and for the suggestions.

Figure 4 is now fixed and a F-test is performed as mentioned above (see response for L190-210).

Table 1 is added and Fig. S1 is removed (its information can be found in Table 1). Table 1 shows the adjusted r-squared and slope of the fitted line for the in-situ comparison.

L270:

The data consistency performance of the two sites is different, and the data consistency of ERA5 and ECCC-B with Aeolus in the Whitehorse site didn't change, and the conclusion given by the authors as well as the explanation is not reasonable enough.

Thank you for the comment. We agree that the data consistency performance of the two sites is different. We added some explanations on the differences. It is mainly due to the topography of the sites and the adjusted r-squared is sensitive to the wind speed range. The paragraph in line 291 is revised:

“Overall, the datasets show strong consistency. ECCC-B and ERA5 are highly mutually consistent (Table 1; with adjusted r-squared greater than 0.97) and therefore show similar consistency with Aeolus (Figs. 3a-b and 4a-b). It can be seen that Aeolus Mie winds are less consistent with ECCC-B, ERA5, and radiosondes at Iqaluit than the corresponding observations at Whitehorse and for the Rayleigh winds. One possible reason for this relates to the fact that the Mie channel samples winds in the lower atmosphere where winds are harder to assimilate or measure due to topography. Since Iqaluit is situated in tundra valleys with rocky outcrops that can cause increased variability in the wind field while Whitehorse is situated in large valleys with less wind variability due to topography, terrain effects might account for the difference in consistency. In addition, the overall range extent of the HLOS wind samples is between -25 to 25 ms^{-1} at Iqaluit and -45 to 45 ms^{-1} at Whitehouse and r-squared is sensitive to

the range of data (note the denominator of the second term in Eq. (5)). Overall, Aeolus data show good agreement with these three datasets with adjusted r-squared greater than 0.8.”.

L275-L292:

The authors' discussion of Ka-band radar consistency with Aeolus and its causes seems unnecessary for this paper, as there are too few overlapping data points and no valid conclusions are drawn.

As mentioned above (see response for L190-210), we acknowledge the little overlap between Ka-band radar or lidar with Aeolus data; however, they are entirely independent from the Aeolus winds. Unlike the radiosondes, they are assimilated in ERA5 and ECCO-B, which is used in the correction of Aeolus winds and quality control process. Thus, the comparison would still add some value to this validation work.

L296:

Are the sample sizes in Figure 5 the same for the three time periods? The use of the expressions summer, fall, and winter is not rigorous and should be specific to dates. In addition, the discussion of Figure 5 is inadequate. If solar radiation is used to explain Figure 5, then why did the overall performance of fall 2018 be better than that of winter 2020? The Mie channel also performed better in summer 2019 than in winter 2020, and the Mie channel is also influenced by solar background radiation. The authors seem to have overlooked some phenomena in their haste to reach conclusions.

Thank you for the suggestions. Number of measurements (N) and number of profiles (p) are now provided in Table S1 to S3. Table S1 is for the validation at the sites, Table S2 for validation over the Canadian Arctic, and Table S3 for validation over the pan-Arctic. And the dates of the seasons are added to Fig. 5.

Thank you for the question. Please note that the range on the y-axis is from 0.7 to 1.0. The difference may look large on the plot, but the change is almost insignificant. The 99% confidence level on the adjusted r-squared is added on the figure. The range of the adjusted r-squared for Mie winds is almost overlapping between the seasons. The following passage is added in Line 337:

“We also note a slight drop in consistency of the Mie winds for the mid-FM-B period, which took place in winter 2020: for instance, the adjusted r-squared and their 99% confidence intervals, between Mie winds and ECCC-B, are 0.92 ± 0.03 during fall 2018, 0.91 ± 0.01 during summer 2019, and 0.87 ± 0.02 during winter 2020. This decrease in the consistency is almost insignificant.”

L355: Since ECCC-B and ERA5 are mutually consistent

This may need to be more fully demonstrated.

Thank you for pointing this out. We have tried make this consistency clearer with the following revision in line 291:

“Overall, the datasets show strong consistency. ECCC-B and ERA5 are highly mutually consistent (Table 1; with adjusted r-squared greater than 0.97) and therefore show similar consistency with Aeolus (Figs. 3a-b and 4a-b).”

L359-L364:

The conclusion of this paragraph does not need to be obtained by data analysis. It is just a mathematical law. The discussion of the longitudinal and latitudinal components also seems to be unnecessary.

Thank you for the comment. We agree with your point. Thus, this paragraph is removed.

Figure 8:

What is the significance of discussing the mean and standard deviation of the wind speed samples themselves? Clarification by the authors is needed. It may be more meaningful to discuss the distribution of the differences between the Aeolus wind measurement data and the ECCC-B and ERA5 data.

Please see above. Figure 8 is now showing the distributions of the differences.

L377: Figure 8 shows an overall agreement between Aeolus, ECCC-B, and ERA5

The proof of consistency between data by comparing data distribution characteristics only is not enough.

Thank you for the comments. Please see above. Figure 8 is now showing the distributions of the differences.

L391: Figure 9 shows that Aeolus data consistently has more structure than ECCC-B during all three periods and for both Rayleigh and Mie winds.

What does "more structure" mean here? Please explain.

Thank you for this question, by "more structure" we mean that the Aeolus' normalized standard deviations are greater than 1.0 for all three seasons and for both channels, and we explain this in the following revision in line 411:

"Figure 9 shows that Aeolus data consistently has greater standard deviations than ECCC-B during all three periods and for both Rayleigh and Mie winds: its normalized standard deviations are typically within 1.05 to 1.40."

L395: During the boreal summer period, the data in the stratosphere seem to agree less with the ECCC-B data, reflecting reduced sampling, solar background noise that is most effective during summer as mentioned earlier, and other possible errors (Reitebuch et al., 2020).

The derivation of this conclusion is not rigorous, there are many possible reasons for the decrease in the correlation between the Aeolus data and ECCC-B data in the stratosphere in summer, and it is also possible that it is caused by changes in the atmospheric environment in summer, thermal changes in the telescope, etc. I don't think we can make speculation on the cause from Figure 9. But the authors seem to attribute it to solar radiation in the abstract.

Thank you for the comment. In the text, we mentioned that there are other possible errors, but we only mentioned the solar background noise in the abstract – this is our mistake. We added in the abstract that the decrease in the consistency for Rayleigh winds in the summer may be due to other possible errors as well.

L402: For this reason, in the next paragraph, where we investigate the spatial distribution of the consistency in the lower and upper atmospheric regions, we exclude the Rayleigh winds in the PBL and the Mie winds in the stratosphere.

It would have been more convincing if the authors had made this data trade-off from the analysis of data used in this paper. Although Aeolus was designed to complement the dual channels at altitude, there are many cases where the Mie channel has higher data volume and data quality at altitude than at lower altitudes, especially in summer. Simply removing the data would be detrimental to the subsequent analysis.

Thank you for the comment. The choice of removing the Rayleigh winds in PBL and Mie winds in the stratosphere mainly comes from Fig. 9. There are two seasons out of the three seasons of study that there is no Rayleigh measurement in the PBL or the normalized standard deviation is greater than 2.2, and one season that there is no Mie measurement in the stratosphere or with normalized standard deviation greater than 2.2. The Rayleigh winds in the PBL are really noisy from the Taylor diagrams and from Rennie and Isaksen, 2020. Therefore, for the consistency across the periods of study, we remove the layer that provides very little and/or noisy data.

Figure10, Figure11:

The calculation process of RMSD in this paper should be clearer, and it is better to give the formula and the range of RMSD to be considered "consistent". In addition, the radial mutations in these two figures seem to be inconsistent with common sense, especially in Figure 11e, which I hope the authors can explain. Also, the same color scale should be used for all subplots. The number of data samples used for different subplots should be provided, because the valid sample size may vary greatly for different seasons at different altitudes. In addition, the data density of Aeolus is also different for different latitudes, how did the authors deal with this point, and does this lead to lower quality data for lower latitudes?

Thank you very much for highlighting this important issue and for the suggested improvements to the presentation. This was very helpful! We have learned that the radial pattern was a spurious result arising from our choice of grids. We corrected this by transforming our data to the EASE (Equal-area scalable earth) grid, described at the NSIDC website (<https://nsidc.org/data/ease>). It is now corrected, and the following explanations are added in Line 436:

"Since the measurement density differs depending on the latitude, the RMSD of the profiles are calculated over nearly equal surface area, using the Equal-Area Scalable

Earth (EASE) Grids (Brodzik et al., 2012). Each grid cell is around 104 km² which is approximately the square of the along-path resolution of Aeolus Rayleigh winds.”

Panels in Figs. 10 and 11 are now sharing the same colorbar.

L446: No significant improvement is seen here because we have implemented a weekly updated dynamic bias correction to the near real time data.

How do the authors explain the slightly higher RMSD of 2B10 data compared to 2B06 in Figure 10 and Figure 11? What is the meaning of the dynamic bias correction mentioned in the paper, please elaborate, and what is the difference between this correction and the reprocessed data correction in the L2B product?

Thank you for pointing this out. We realized that the averaged RMSD shown in the subtitles are dependent on the grids chosen. Since we are now using different grids (EASE 2) (please see response for Figs. 10 and 11), the averaged RMSD also changed. To avoid misleading results, the average is removed from the subtitles and the paragraph in line 469 is revised:

“Note that the first reprocessed data, 2B10, only overlaps with one of the three periods of study: August to September 2019. The estimated observational errors have decreased compared to the 2B06 data (Figs. S1 and S2) since the bias due to the M1 mirror temperature dependence is updated on a daily basis and the dark current signals have been removed using improved quality control. However, we do not see the same improvement in the O-B statistics between 2B06 and 2B10 products over the Arctic region.”.

L476: We have found some initial evidence that the estimated error product is also a good predictor of RMSD between Aeolus and the reanalysis, which could be useful for constraining future forecasts.

The "estimated error product" itself is used to estimate the difference between the Aeolus wind product and the true wind field. If it does not predict the RMSD between Aeolus and reanalysis data, then the reanalysis data deviates from the real wind field. I don't understand the purpose of the author's statement, perhaps more information about "constraining future forecasts" is needed.

We agree that this sentence should be clarified. What we meant by constraining future forecasts is that the assimilation of Aeolus winds, in an optimal way, would improve analyses and forecasts. Consequently, the analyses and forecasts would be closer to the true atmospheric state. One way to get the best from assimilating Aeolus winds

into forecasting systems is to use most accurate observation errors for these data. We found that the L2B estimated error product could be useful for specifying the HLOS wind observation errors since there is a good correspondence between the spatial variability of the time-averaged L2B estimated error products (Figs. S1 and S2) and that of the RMSD between Aeolus and ECCC-B (Figs. 10 and 11). The sentence has been changed as follows:

"We found that the spatial variability of the time-averaged L2B estimated error product is in good agreement with the spatial variability of the RMSD between Aeolus and ECCC-B HLOS winds over the Arctic region. This validates the use of L2B estimated error product as a predictor for the HLOS wind observation errors in data assimilation systems, as proposed by Rennie et al. (2021), to obtain optimal positive impacts on forecasts from assimilating Aeolus winds."