We thank reviewer 1 very much for his/her careful reading and the corresponding valuable comments and suggestions, which are for sure very beneficial for the presented paper. We agreed on almost all. Please find below our reply to each comment (*italic*). Major comments if Reviewer 1

- line 2: Aeolus measures horizontal wind profiles => Aeolus measures profiles of a single horizontal wind component
  We changed the text accordingly.

- line 127: 30 x 2.7 = 81 not 87, please correct. It is more like 30 x 2.85.
  Thanks for the careful reading. We changed it accordingly and gave reference to the recent memorandum for wind retrievals (doi:10.21957/alift7mhr)

- line 152: "It is obvious that only two out of this four wind products are useful, namely the Rayleighclear and the Miecloudy product." Well, this is not that obvious. Also Rayleigh- cloudy can provide a useful wind because the L2B can correct for Mie contamination, in principle, using the scattering ratio as input. It is true that until now this has not been successful enough for operational use. So please update your text accordingly.

  You are completely right. I even have analysed the Rayleigh_cloudy winds by myself and found partly good agreement. Thus, text was changed accordingly.

- Figure 2 is from Raman-polarization lidar Polly, I guess? Please mention in the caption of the figure.
  Done!

- line **2**33: "In contrast, it was obviously not detected by the Aeolus measurement due to the fact that the Miecloudy wind is obtained practically only from return signals of the cloud and thus only from the height range at which the cloud was observed within this one range bin of 1 km thickness." I guess the problem here is that Aeolus only measures the wind at cloud top, so for a fair comparison you should compare with the radiosonde value at the cloud top. On the other hand, Aeolus cannot determine the exact location of the cloud top inside the vertical bin and the best one can do is to assign the Mie wind to the bin centre location, hence giving the large error the authors observe. Note that the Mie channel is much more sensitive to such height assignment errors than the Rayleigh channel (X. J. Sun, R. W. Zhang, G. J. Marseille, A. Stoffelen, D. Donovan, L. Liu, and J. Zhao, The performance of Aeolus in heterogeneous atmo spheric conditions using high-resolution radiosonde data, Atmos. Meas. Tech., 7 , pp. 2695-2717, 2014, doi:10.5194/amt-7-2695-2014)
  You are right with that. We meant the same, but obviously our non-native English prohibited the right message. We change this accordingly:
  "The disagreement is caused because with the Aeolus Mie algorithm, the wind speed at cloud top is measured but due to the range-bin thickness of 1 km, the top height of this cloud cannot be correctly assigned. Thus, the Mie wind speed measured at cloud top is assigned to the center of the 1 km thick range-bin disregarding the true top-height of the cloud. As a consequence, the agreement to the high resolution radio sonde profile is much better (almost identical values at 2.5 km) than to the radio sonde profiles binned to Aeolus resolution. The presence of cloud or aerosol layers in the measurement bins was already discussed prior launch by Sun et al. (2014) and it was shown that biases of more than 0.4 m/s can occur when the cloud top is not in the center of the range-bin. This statement is confirmed by our observations and shows that a higher vertical resolution is in principle preferable and valuable."

- line 268: "Thus, these wind measurements at this altitude should be neglected until the hot pixel correction is in place." Since figure 5b does not show altitude, it would be good to explicitly

mention to ignore the (dark) red colors, which indeed contradict with the Rayleigh-clear winds above (discontinuity).
We added: "(indicated by reddish colors just above the bluish colors in the lowermost profile)".

- line 318: "Due to its large vertical resolution". This is incorrect. Should be "coarse vertical resolution" or "large vertical bins"
corrected!

- Figure10a/11a. Data analysis from Aeolus have shown substantial differences between statistics from ascending and descending orbits. In figure 10a and 11a it would be interesting to indicate this by using different colors for the dots in the scatterplot.
Thanks for this very interesting advice. We updated the Figures accordingly and also added one more column in Table 1 indicating either ascending or descending orbit. Nevertheless, from our 6 observations in this early mission stage, no significant difference between the two orbit types are visible.
We added in the text: "The different colors indicate whether Aeolus had an ascending node (green) or descending node (red), i.e., if the measurement was taken at local evening or local morning, respectively. This separation is done because first long-term Cal/Val activities showed significantly differences in the determined biases of Aeolus wind measurements between the two different modes (Rennie and Isaksen, 2020; Geiß et al., 2019; Krisch et al., 2020). However, from our observations onboard RV Polarstern in the early mission phase of Aeolus, we do not observe a significant difference between the two modes with respect to the Rayleigh_clear winds." and regarding the Mie winds:
"As for the Rayleigh_clear winds, no difference between the Aeolus performance for ascending and descending orbits is found (Fig.11a)."

- line 345: "Considering the relatively small amount of measurements for this statistic, an almost Gaussian shaped distribution is found. Thus, one can conclude that the deviation between Aeolus and the radiosonde wind observation is normally distributed." You cannot conclude this based on the shape of the distribution in figure 10b. Please rephrase or otherwise apply a statistical analysis to test this Gaussian hypothesis.

  You are right, we deleted this statement at this stage and modified a later statement about this, so that the conclusion is significantly weakened: "As this is expected for a Gaussian distribution, one could assume, in accordance with the shape of the distribution shown in Fig. 10b, a normally distributed behaviour of the Rayleigh_clear wind deviations."

- line 353: Please mention that the MAD is less sensitive to outliers and equals 0.674 x STD for a perfectly Gaussian distributed stochastic variable. This gives a good handle on how to interpret the value of 3.33 m/s in the next line, i.e., it corresponds to 4.94 m/s error standard deviation, the metric more commonly used in error quantification and data assimilation.
Thanks, done:
"The median absolute deviation (MAD) of the distribution is used to calculate the random error of the Aeolus wind observations (Lux et al., 2020; Witschas et al., 2020) because it is less sensitive to outliers than the standard deviation. It is 67.4% of the standard deviation or the other way around, the scaled MAD (MAD times 1.4826) is identical to the standard deviation for a perfectly Gaussian distribution. The scaled MAD is thus used an indicator for the random error for Aeolus observations. The MAD is in case of the Rayleigh clear winds 3.26 m/s, the scaled MAD correspondingly 4.84 m/s."

- Table 2. Why differs the value of 3.26 for Rayleigh-clear MAD from 3.33 in the text? Please correct.
  Done, thanks

- line 364: "This is caused by the generally lower Rayleigh return signal compared to the Mie channel. Rayleigh scattering is orders of magnitude lower than the Mie scattering." The difference is not only SNR, it is also different interferometers and different type of processing for wind retrieval (peak fitting versus fitting of measured Rayleigh Response to temperature and pressure dependent (Rayleigh Response, Doppler shift) curves/tables). Please rephrase to something like: "This is mainly caused by the generally lower SNR of the Rayleigh return signal compared to the Mie channel, besides the different measurement and retrieval techniques".

  Done! Thanks for the suggestion!

- line 373: "Some instrumental effects, like the hot pixel issue, have not yet been corrected" In the mean time, the main reason for biases of Rayleigh winds has been found: temperature variations over the telescope which are not fully compensated for by the instrument. Please mention this tool
  We mentioned it and rephrased to:
  "Despite the mission requirements could not yet be achieved, the mission can be seen as success as it was already demonstrated that winds are globally observable from space by active remote sensing with sufficient quality to achieve a positive impact in NWP (Rennie and Isaksen, 2020; Martin et al., 2020). However, it is worth to mention again that the Aeolus data which was used is not yet the finalized data set for this space mission. In the meanwhile it was found, that slight temperature variations over the receiving telescope area are one of the main reasons for biases of the Rayleigh winds (Rennie and Isaksen, 2020; Krisch et al., 2020; Reitebuch et al., 2020). This effect and some other instrumental challenges, like the hot pixels issue, have not yet been compensated in the data of the early mission stage. Processor updates with several improvements have been taking place in the meantime and more are expected in the future to correct such effects, after which a reprocessing of the early Aeolus data set is foreseen.

- line 374: "Despite the mission requirements could not yet be achieved, the mission can be seen as success as it was already demonstrated that winds are globally observable from space by active remote sensing with an accuracy needed for assimilation in NWP." I would end with "..... active remote sensing with sufficient quality to demonstrate positive impact in NWP [Ref]". With a reference to results presented by ECMWF or ESA outreach publications.

  DONE, see above.

- line 397: "horizontal heterogeneity". I guess you mean: "horizontal atmospheric heterogeneity"? Please add.
  Thanks!

- line 421:"was not sufficient to capture the maximum wind speeds in relatively thin strong-wind regions, here discussed in terms of the example of the tropical jet stream". I would rephrase to: "was not sufficient to capture events of strong vertical wind-shear such as near the tropical jet stream". *Thanks for the suggestion! Sone!*

- line 434: you could add: "in fact, Rayleigh-clear winds have proven more beneficial for NWP than Mie-cloudy winds".
  Added. Thanks!

- line 436: "..... and random error of 3.3 m/s for the Rayleigh ...." This is very misleading as this value does not represent the usual STD but MAD, see comment above. So please translate this value to STD.
  Thanks for the hint! Done!

- line 438: "Some known instrumental effects and calibrations have not yet been implemented in the retrieval algorithms" Rephrase to: "In the meantime discovered instrumental and calibration imperfections were not yet implemented in the retrieval algorithms used for the 2018 autumn data set"
  Done!

- In this context, do you have plans to use reprocessed, unbiased, Aeolus data with the same radiosonde data set presented here? I would very much encourage the authors to write a follow-up paper, once the reprocessed data for the autumn 2018 period become available. If so, please mention in section 5.
  Yes, this is a good idea. We mentioned it now: "Once a final reprocessing has been taken place it could be worth to use the existent RV Polarstern data set to quantify the improvements of the algorithm updates."

Minor comments / typos ==========

- line 39: observation => observations*, done*
- line 62: chosen => selected *, done*
- line 89: around ≈;remove either 'around' or '≈' , done
- line 94: to retrieve wind retrievals => to retrieve winds*, done*
- line 100: correction => corrections*, done*
- line 110: the data must be available within 3 hours=> the data must be available within 3 hours after measurement time (timeliness). By the way, this is not true for ECMWF who wait about 5 hours before they start there analysis run. This is valid for medium-range forecasts. Mesoscale meteo centers need the data within 3 hours for operational use. *, thanks, good to know!*
- line 116: parameter => parameters *, done*
- line129: The currently applied method by ESA is the use of the scattering ratio => The currently applied method by ESA is the use of the scattering ratio, which is determined as part of the L1B processing (ref) and used as input for the L2B processing. *, done*
- Line135: comprised => comprises *, done*
- line 159: please explain DISC *, done*
- line 164/166: pixel => pixels , *done!*
- line 166: increase => increased *, done*
- line 246: the resolution is simply too low => the resolution is simply too coarse, *done*