Answers to Referees, 2nd review.

Reviewer 1

The authors have addressed all the points raised in my review and revised the manuscript accordingly. Therefore, I recommend publication of the manuscript after the following two minor technical corrections:

1. For the citation of webpages in the introduction (l. 65, l. 118, l. 138, l. 245) a representative title should be used instead of the URL. The latter is supposed to be included in the reference list together with corresponding title and access date.

These references are now identified as (Aeolus-ESA-Portal-ext, date) with "ext" defined as the two extensions "forecast" and "mission", respectively

2. In Table 2, third line from top, 7.2 km/s should be changed to 7.2 km·s-1 in order to be compliant with the journal's formatting style.

Identified modification in Table 2 has been made.

Reviewer 2

- Line 75: Is this really the first airborne HSRL? Sroga, Eloranta, et. al., 1983 - (airborne HSRL demonstration in 1980). Also see Shipley et al. 1983. Also important to recognize the work of Eloranta et. al with NCAR –the (HIAPER) GV-HSRL system.

Although we had the Shipley et al. paper high in our tablets, we omitted the airborne aspect of this development. The work of Ed Eloranta was also more associated with ground-based measurements. Thank you to have refreshed our memories, and our apologies to these authors. We have modified the sentence as

After the first pioneering developments using Fabry-Perot interferometers (Shipley et al, 1983; Sroga et al, 1983), the first operational airborne HSRL systems were developed in the USA at U. Wisconsin [Eloranta et al., 2008] and at NASA [Hair et al., 2008], as well as in Europe at DLR [Esselborn et al., 2008], all these systems being based on the Iodine cell absorption technique.

And added three more references

Shipley, S. T., D. H. Tracy, E. W. Eloranta, J. T. Trauger, J. T. Sroga, F. L. Roesler, and J. A. Weinman, "High spectral resolution lidar to measure optical scattering properties of atmospheric aerosols. 1: Theory and instrumentation," Appl. Opt. **22**, 3716-3724, 1983.

Sroga, J. T., E. W. Eloranta, S. T. Shipley, F. L. Roesler, and P. J. Tryon, "High spectral resolution lidar to measure optical scattering properties of atmospheric aerosols. 2: Calibration and data analysis," Appl. Opt. **22**, 3725-3732, 1983.

Eloranta E. W., I. A. Razenkov, J. Hedrick and J. P. Garcia, "The Design and Construction of an Airborne High Spectral Resolution Lidar," *2008 IEEE Aerospace Conference*, pp. 1-6, doi: 10.1109/AERO.2008.4526390, 2008.

• Line 77-84 – this addition of more background information on the LNG is quite helpful, especially on connection with the new Appendix C.

Thank you. We thought a brief historical information could be useful.

• Lines 270-274: The cat-eye arrangement described in Tucker et al. 2018 (based on Wang et al. 2000) provides an 8 mrad FOV for a QMZ with a 90 cm OPD (challenging for a solid MZI) ; and the description indicates it does not require delicate mechanical stability for the same reasons that QMZs do not require delicate laser frequency stability if one is capturing the reference phase.
Suggest replacing:

"An alternative design of a field-compensated MZI is given by the cat-eye arrangement (Tucker et al., 2018). This fully reflective design eliminates the light path through the glass which can cause a wavefront distortion but leads to a bulkier arrangement which requires a delicate mechanical stability. Additionally, as discussed in section 3.4, an all-prism MZI design with a small OPD can achieve high quality and the thermally induced wavefront distortion can be easily controlled."

with

"An alternative approach to field-compensating an MZI is to use cat-eye arrangement as demonstrated for a 90 cm OPD in Tucker et al., 2018.**"**

We have been asked by Reviewer 1 to compare the advantages and disadvantages of the proposed design with OAWL. So we tried to give the main lines of this comparison, in a few words. We agree that a 90 cm OPD with an all-prism MZI is unrealistic, but we actually address Doppler measurements relying mainly on molecular scattering which require a much smaller OPD. We do not fully agree with the reviewer when he proposes to shorten the explanations on the selected designs, but we agree it should be a more balanced presentation. In our opinion, the main difficulty of a cat-eye design lies in its mechanical stability (even with a monitoring of the reference during operation) while the main disadvantage of the all-prism design lies in its thermal sensitivity. To be fair, and more explicit, we replace the two sentences by:

An alternative design of a field-compensated MZI is given by the cat-eye arrangement (Tucker et al., 2018). This fully reflective design eliminates the light path through the glass which can cause a wavefront distortion but leads to a bulkier arrangement which requires a high mechanical stability. On the opposite the all-prism MZI design is insensitive to vibrations but is sensitive to temperature gradients. Nevertheless, as discussed in section 3.4, for a small OPD the all-prism design can achieve high quality and the thermally induced wavefront distortion can be easily controlled.

- Line 412-419: While this new discussion on speckle is highly useful, it appears to interrupt the discussion on sampling and accumulation. Perhaps move it to just before the line starting with "Appendix C shows measurements performed with…."

Agreed. The sentence is moved in the newly revised version.

- Line 423: Regarding this sentence "The signals are summed on 14 elementary samples corresponding to a total of Ns=700 shots for an observation horizontal resolution of 50 km" –

  - What is an "elementary sample" – this appears to be the first mention of the term. Some readers might understand that it refers to a 50 shot ACCD accumulation, but please clarify.
  - Assuming that elementary samples refers to sets of 50 shots (0.5s accumulations at 100Hz PRF), then if the signals are summed over that time (14 x 0.5s = 7s) does the laser frequency actually need to be stable over 700 shots (7 seconds)?
  - Perhaps the word "summed" could be replaced with "reference phase adjusted and accumulated"?

To be more explicit, we added some information as

The signals are first accumulated on the ACCD over 50 shots forming an elementary sample. Then 14 elementary samples are reference adjusted and accumulated, corresponding to a total of *Ns* = 700 shots, for an observation horizontal resolution of 50 km and the resulting total SNR is

$$SNR = \frac{S_{tot}\sqrt{N_s}}{\sqrt{S_{tot} + S_b}}$$

- Figure 4: While two of the figure subplot titles now list the wavelength in their title, it would still be helpful to indicate the wavelength (for all the profiles and subplots) in the figure caption.

OK

- Lines 543-545: The second half of the sentence is unclear: "… but in regions where the aerosol load is significant (R ≥ 2) the contributions of the uncertainty on βmol and Mmol are of the same order of magnitude" The contributions of which uncertainty on Bmol and Mmol are of the same order of magnitude? Do the authors mean to say "…but in regions where the aerosol load is significant (R ≥ 2), contributions to the total error from uncertainty in βmol and Mmol are of the same order of magnitude."

OK, we propose to rewrite as

In clear air the uncertainty on M_{mol} dominates the total error on the backscatter and extinction coefficients (see Appendix D, Eqs D1 and D2), but in regions where the aerosol load is significant ($R_\beta \geq 2$), contributions to the total error from uncertainty in β_{mol} and M_{mol} remain small and comparable in magnitude.

•	Appendix C: It is important that that authors have added a new section to show validation of the instrument performance model.

○ Regarding Figure C1b - Is the standard deviation (sigma) estimate (provided in equation A-17) a minimum sigma, or a lower bound on sigma? If so, how is it that the measured sigma is sometimes lower than the calculated sigma? Wouldn't one expect that with the added natural atmospheric variability, the measured standard deviation would always be larger than the model?

Equation A-17 gives the theoretical standard deviation $\sigma(V_{LOS})$ caused by detection noise. It is the standard deviation that would be obtained on an infinite series of measurements. In Fig. C1-b each point of the observed standard deviation is an estimate of this standard deviation obtained on only one hundred measurements. This standard deviation estimate is itself a stochastic variable. Its mean value would reach $\sigma(V_{LOS})$ for an infinite averaging (in the absence of additional variability). It is thus not impossible to obtain standard deviation estimates lower than the theoretical standard deviation even if an additional cause of variability is brought by the atmosphere (as mentioned in the text, the average of the observed standard deviations is slightly higher than $\sigma(V_{LOS})$, due to this additional variability).

○ Line 808: The equation A-17 used in the comparison shown in Figure C1b does not include "N" (number of shots, number of samples) used in the estimate. What parameters were used in A-17 for SNR, Mo, Matm, and how were the accumulations accounted for in the measurements vs. the model? This can probably be addressed with a few simple parameters including SNR, Mo, and Matm used in A-17, as well as a 1/sqrt(N) factor to account for the various pulses used in the estimates.

We believe that the measurement conditions are clearly defined. Equation A-17 includes the detection SNR which itself depends on the number of accumulated shots N (100) and on the vertical resolution (100m). As discussed in section 3.4 the intrinsic modulation factor M_0 of the HSRD-LNG is 0.65. Actually, the SNR is derived directly from the recorded signals, as well as the observed M_{atm}. These parameters, specific to the measurement conditions, are used in Eq. A-17 for the calculation of $\sigma(V_{LOS})$.

○ If the authors are short on room, the authors could remove lines 814-833 and just reference the Bruneau 2015 (or 2020 conference paper "Operation of the airborne 355 nm high spectral resolution and doppler lidar LNG) instead. Unless perhaps other reviewers have requested validation of the general QMZ approach It's not clear that the newly added radiosonde comparisons for the airborne campaign are critical to the main paper.

The statistics resulting from the comparison between lidar and dropsonde measurements have not been presented previously. Measurement bias is an important parameter in the assimilation process, requiring potential corrections as it is now the case for Aeolus. The histograms of figure C2-d and table C1 show that wind measurements can be carried out by the QMZ technique with a very limited bias. These results are related to the validation of the statistical error (bias and precision), and we believe they are of great interest in the context of the paper to show that the measurement bias with a QMZ can be very low.

○ Figure C3: like for Figure C1b, it's unclear how the measured standard deviation can be smaller than the ideal calculated standard deviation unless the calculation is itself based on an uncertain variable (e.g. SNR?). Which parameters

Same answer as above.

All suggested modifications to improve English have been implemented in the text.

Additional modifications

Last modifications have been made the text.

In the appendix C, one should read LOS instead of HLOS (now modified as)

Taking the sine fit as reference, we can compute the standard deviation of the LOS measurements and compare it to the standard deviation calculated with Eq. A17. Figure C1b shows the comparison of the measured and calculated LOS wind speed standard deviations computed at different altitudes on three different VAD measurements taken at 16:37, 17:34 and 18:00 UTC (for a total of 98 measurements).

The reference to the work of Herbst and Vancken (2016) on UV Doppler lidar using a Michelson interferometer has been added in section 2 and in the reference list.

Herbst, J., Vrancken, P.: Design of a Monolithic Michelson interferometer for Fringe-Imaging in a Near-Field, UV, Direct Detection Doppler Wind Lidar, Applied Optics, 55, 25, 6910-6929, 2016.