Response to Referee Comment 2 on "*Rayleigh wind retrieval for the ALADIN airborne demonstrator of the Aeolus mission using simulated response calibration*"

We appreciate the referee's insightful and valuable comments on our manuscript AMT-2019-274. We thank the referee explicitly for careful proofreading, which significantly improve the clarity of the text and the readability for a broader audience. According to the suggestions and questions, the point-by-point response is attached in this file.

Major comments:

This paper presents an alternative technique for retrieving LOS wind estimates from the molecular channel of the Aeolus Airborne Demonstrator (A2D) using modeled response functions ("Simulated Rayleigh Response Calibration" or SRRC) derived using best-fit instrument models and the given atmospheric conditions (temperature and pressure) when available from other observations.

1. The SRRC approach provides some advantages over the "traditional" double-edge approach of measuring calibration response curves during the test process (the MRRC approach), but the authors could do a better job of explaining the reasoning behind this (vs. just listing numbers) at the beginning of the paper and in the abstract. The approach is a good idea, especially when faced with consistent Mie contamination during flight tests.

   R: revised.

   In the abstract, page 1, line 14-19, the reason why SRRC provides advantages over MRRC is added and revised as "…, However, differences in the atmospheric temperature profile between the location and time of the MRRC and the actual wind measurements are important sources of wind bias since the atmospheric temperature has a direct effect on the instrument response calibration. Furthermore, some experimental limitations need to be considered carefully to achieve a reliable MRRC. The atmospheric and instrumental variability thus currently limit the reliability and repeatability of this MRRC. In this paper, a procedure … to resolve these limitations of the A2D Rayleigh channel MRRC".

   In addition, related introductions are also added **in Sect.1, Page 3, line 14-23**, "Currently, only measured Rayleigh response calibrations (MRRC) are used for the A2D (Marksteiner, 2013; Lux et al., 2018; Marksteiner et al., 2018). However, the atmospheric temperature affects the Rayleigh-Brillouin line shape, and has a direct effect on the instrument response calibration (Dabas et al., 2008). Differences in the atmospheric temperature profile of the time and location when the MRRC was obtained and the actual wind measurements are important sources of wind bias, which are especially severe in case of large temperature differences. This is the reason why it is mandatory to consider the atmospheric temperature in the Aeolus level 2B procedure to retrieve reliable winds (Dabas et al., 2008; Rennie et al., 2017). Furthermore, some experimental limitations, which will be introduced specifically in Sect. 2.1, need to be considered carefully to achieve a reliable MRRC. Overall, the atmospheric and instrumental variability coming along with an MRRC limits the reliability and repeatability of A2D instrument response calibrations."

   Have other double-edge wind lidar researchers done anything similar to this before?

R: Yes, as shown in Table 1, there are several FPI-based direct detection wind lidar systems that are capable of measuring wind based on a measurement approach or a simulation approach. The black-marked parts use a simulation approach to obtain calibration response curves, which is similar to the SRRC method mentioned in this paper.

Table 1. Comparison of different FPI-based direct detection wind lidars

Lidar	Wavelength and system	Calibration approach	Instrument drift correction	References
OHP [a] Rayleigh lidar	532 nm, double FPIs	**Simulation**, FPI scanning	quick wind acquisition cycle strategy	Chanin et al., 1989; Garnier and Chanin, 1992; Souprayen et al., 1999a, 1999b
NASA [b] Rayleigh/Mie lidar	355 nm, three FPIs	**Simulation** FPI or laser scanning	locking etalon and servo-control system	Korb et al., 1992; Korb et al. 1998; Flesia and Korb, 1999; Flesia et al., 2000
USTC [c] Rayleigh lidar	355 nm, three FPIs	measurement and **simulation,** FPI scanning	locking etalon and servo-control system	Xia et al., 2012; Dou et al., 2014
ESA ALADIN	355 nm, double FPIs for Rayleigh channel	level 1B: measurement, laser scanning **level 2B: simulation,** laser scanning	internal reference path	Reitebuch et al., 2018; Rennie et al., 2017
DLR A2D	355 nm, double FPIs for Rayleigh channel	Measurement, laser scanning	internal reference path	Marksteiner, 2013; Lux et al.,2018; Marksteiner et al., 2018

[a] Observatory of Haute Provence, France
[b] National Aeronautics and Space Administration, U.S.
[c] University of Science and Technology of China, China. This lidar is mobile.

2. An explanation of the physical differences between the internal reference channel and the atmospheric channel would be helpful. For example, does the IRC have a different set of field angles into the FP etalons than provided by the telescope/receive path returns? Does the IRC only see narrowband light?

   R: revised. Thanks for your suggestion, we didn't explain it clearly. Yes, the atmospheric path and internal reference path differ in their field angles. The internal reference signal is coupled into the receiver via an optical fiber whereas the atmospheric signal enters the receiver via free beam bath through a set of different optics. This leads to a slightly different set of field angles on the FPIs for the internal path and the atmospheric path. During the ISR only the internal path, illuminated with spectrally narrow-band light from the laser is recorded, while for the IRC the internal path (with narrow spectral bandwidth from the laser) and the atmospheric path with broad spectral bandwidth molecular returns, but also narrow spectral bandwidth cloud, aerosol and ground returns is recorded.

The specific schematic of ALADIN Airborne Demonstrator (A2D) was shown in Fig. 1 in (Lux et al., 2018), which has been already referenced in the following added paragraph **in Sect 2, page 4, line 11-23**, see below:

"For each direct detection wind lidar system, the emitted laser frequency should be known to accurately derive the Doppler frequency shift. A zero Doppler shift reference determined by pointing to the zenith direction has been used to correct for the short-term frequency drift in previous studies (Souprayen et al., 1999b; Korb et al., 1992; Dou et al., 2014). But for the A2D, the internal reference path is specially used to measure the emitted laser frequency information. As shown in Fig. 1 in (Lux et al., 2018), a small portion of laser beam radiation is collected by an integrating sphere and coupled into a multi-mode fibre, then injected into the receiver via the front optics. The atmospheric backscattered signal is collected by a Cassegrain telescope and guided via free optical path propagation to the front optics and receiver successively. This path is called the atmospheric path. An electro-optic modulator is used to separate the atmospheric signal from the internal reference signal temporally in order to minimize the contamination of the internal reference signal with atmospheric signals and saturation of the detectors at short ranges (Reitebuch et al., 2009). Because of the different optical illumination of the internal path and atmospheric path resulting in different divergence and incidence angles on the FPIs, the response calibration curves for these two paths are slightly different. Note that ALADIN uses free path propagation rather than a fibre coupling unit for the internal reference path."

The related descriptions of the internal reference path and atmospheric path are also updated:

1. **In Sect 2.1, page 5, line 9-11**, "The ALADIN Rayleigh winds produced by the level 1B processor (Reitebuch et al., 2018) are based on a MRRC while the level 2B processor uses SRRC. Basically, MRRC includes two response calibration curves derived from internal reference path and atmospheric path, respectively."

2. **In Sect 2.2, page 6, line 25-28**, "The A2D SRRC based on this simulation approach promises an improvement in terms of A2D wind speed errors due to the limitations of A2D MRRC. Similar to MRRC, SRRC also includes two response calibration curves derived from internal reference path and atmospheric path, respectively."

3. The paper would also benefit from a short, clear discussion on the topic of Mie (aerosol) contamination on the Rayleigh calibration as the topic comes up several times in the paper. Present the reasons for the aerosol induced bias and reference the literature. This could be followed by cleaning up some paragraphs that vaguely refer the issue, without explaining it.
R: revised. Thanks for your suggestion. About the topic of Mie contamination on Rayleigh calibration, we have updated related paragraphs.

Firstly, **in Sect 2.1, page 5, line 23-27**, we discuss the reasons for the aerosol induced bias: "…, the spectrally narrowband Mie scattering which is not filtered out by the Fizeau interferometer will enter the FPIs and can be considered as Mie contamination of the Rayleigh signal. Because of the different spectral widths of the particle and molecular backscatter signal, the sensitivities of the FPIs on them are different. The Mie contamination on Rayleigh channel is one of the sources for systematic errors because it modifies the instrument response calibration curve, which should be avoided to ensure the representativity of pure Rayleigh response."

Then, we analyze the LOS wind velocity error induced by Mie contamination **in Sect 3 page 9 line 9-14** based on simulation results: "The LOS wind velocity error induced by Mie contamination ΔV is defined as the difference of LOS wind velocity under pure atmospheric molecular condition and atmospheric spectral condition with scattering ratio of ρ. Figure 2 shows the simulation for ΔV at T=223 K and P=301 hPa, where the x-axis and y-axis represent different response values and scattering ratios, respectively. Positive and negative ΔV represent the overestimation and underestimation of LOS velocity, respectively. An overestimation of the LOS velocity occurs at response values less than 0.2. Larger scattering ratios result in a larger overestimation. The difference can get up to 20 m s^{-1} in case of $\rho > 10$, if this Mie-crosstalk is not considered.

We also introduce the effect of Mie contamination correction on systematic error optimization in **Sect 3 page 9 line 14-19**: "According to previous studies (Dabas et al., 2008), it is implied that the Mie contamination correction could improve the quality of the Rayleigh wind in the cases of intermediate ρ, e.g. below 1.5, as in this case the Mie signal is not high enough to guarantee an accurate Mie wind measurement but rather becomes significant for the Rayleigh channel (Sun et al., 2014; Lux et al., 2018). The value of ρ, which is needed to correct for the Mie contamination in the Rayleigh channel, is obtained by analyzing the Mie channel signal (Flamant et al., 2017). "

Reference:

1. Flamant, P., Lever, V., Martinet, P., Flament, T., Cuesta, J., Dabas, A., Olivier, M., Huber, D.: ADM-Aeolus L2A Algorithm Theoretical Baseline Document, AE-TN-IPSL-GS-001, 5.5, 89 pp., 2017.

Specific comments:

Some additional proofreading for English language/grammar should catch some minor errors. Remaining comments listed by page/line#. A "Fair" rating is listed under Scientific Quality because it's not quite at the "Good" level with respect to referencing related work and being clear on the issues addressed, but with minor improvements as listed above and in the following comments, it will likely be above good. Overall, this is an interesting and useful paper for the field of double-edge direct detection Doppler wind lidar systems.

1. Page 2 line 24-26: This sentence describing Aeolus is awkward. Perhaps reword as "The novel combination of these two techniques, integrated for the first time into a single wind lidar, expands the observational altitude range from the ground to the lowermost 30 km of the atmosphere."

   R: revised. Please see **Sect 1 page 2, line 28-30**: "The novel combination of these two techniques, integrated for the first time into a single wind lidar, expands the observational altitude range from ground to the lowermost 30 km of the atmosphere."

Line 29: Can delete the words, "as well" from the end of the sentence since it begins with "Furthermore".

R: revised. Please see **Sect 1 page 3, line 1-3**: "Furthermore, as the first high spectral resolution lidar in space (Ansmann et al., 2007; Flamant et al., 2008), ALADIN has the potential to globally monitor cloud and aerosol optical properties to contribute to climate impact studies."

2. Page 3 Line 9: Can the authors expand a little bit on the causes of "...the atmospheric and instrumental variability" for readers not familiar with the observation approach. For example, how atmospheric pressure/temperature impact the MRRC and what varies in the instrument (temperature impacting alignment? Variations in the field of view/field angles entering the etalon? Etc.?)

   R: revised. Please see **Sect 1 page 3, line 15-22, see also reply to major comment 1**

   "However, the atmospheric temperature affects the Rayleigh-Brillouin line shape, and has a direct effect on the instrument response calibration (Dabas et al., 2008). Differences in the atmospheric temperature profile of the time and location when the MRRC was obtained and the actual wind measurements are important sources of wind bias, which are especially severe in case of large temperature differences. This is the reason why it is mandatory to consider the atmospheric temperature in the Aeolus level 2B procedure to retrieve reliable winds (Dabas et al., 2008; Rennie et al., 2017). Furthermore, some experimental limitations, which will be introduced specifically in Sect. 2.1, need to be considered carefully to achieve a reliable MRRC. Overall, the atmospheric and instrumental variability coming along with a MRRC limits the reliability and repeatability of A2D instrument response calibrations."

   Line 12: update to read, "It is based on an accurate theoretical model of the FPI transmission function...."

   R: revised. Please see **Sect 1 page 3, line 24-25:** "It is based on an accurate theoretical model of the FPI transmission function and the molecular Rayleigh backscatter spectrum."

   Line 28: edit to read, "Table one lists FPI-based direct detection wind lidar systems that are capable of measuring wind information...." Note that not all existing FPI systems that can be modeled this way are listed in the table, there are others in existence.

   R: revised. Please see **Sect 2 page 4, line 8-10**: "Table 1 lists several FPI-based direct detection wind lidar systems that are capable of measuring wind information based on measurement or simulation approach."

3. Page 4 Line 10 – Should be "atmospheric conditions"

   R: revised. Please see **Sect 2.1 page 5, line 3-4:** "…, the calibration procedure can be carried out frequently based on atmospheric conditions (Dou et al., 2014; Liu et al., 2002),…"

4. Page 5 Line 19: fix to read, "...the transmission functions of the FPs for the atmospheric path

are slightly different compared to ...” Then please explain why this is (physics causing the differences).

R: revised. Please see **Sect 2.2 page 6, line 19-24.**

“However, the transmission characteristics of the FPIs for the atmospheric path are different from the transmission curves registered on the internal reference path during the instrument spectral registration because of a slightly different illumination of the beams in the respective paths due to different divergence and incidence angles on FPIs (Reitebuch et al., 2009).”

Line 24: “regardless of measurement or simulation method, any angular alignment drift will change the incidence angles on FPIs, and hence change their transmission characteristics.” Technically, the FPI transmission characteristics should be a function of incidence angles, field of view, temperature, pressure, thickness or gap length, finesse, etc. so perhaps the better term here (and elsewhere) is to say that “any angular alignment drift will change in the incidence angles on the FPIS, resulting in a different transmission value.” (or something similar).

R: revised. Thanks for your comments, it has been revised as “Furthermore, the FPIs transmission characteristics should be a function of incidence angles, field of view, temperature, pressure, thickness, and so forth, regardless of measurement or simulation method, any angular alignment drift will change in the incidence angles on the FPIs, resulting in a different transmission value.” Please see **Sect 2.2, page 6 line 30-31 to page 7 line 1-2**.

5. Page 6 Line 5: This is an unusual mix of variables (wavelength and frequency shift), but ok. Line 17 and 19: The authors state that Equations 3 and 4 represent convolutions, but this is not mathematically so. These are integrations over frequency of the product of the FPI transfer function times the specific input spectrum value at that frequency. Likewise, integrating this product over only one free spectral range implies that the authors assume there are never any signals outside the etalon FSRs (e.g. where the etalon can start to transmit again). This may be practically true for most applications/wind speeds/platform pointing motions, etc. but should at least be stated as an assumption.

   R: revised. Please see the updated **Equation 3 and 4 in Sect 3, page 7, line 22-24.**

   Still we state that the equation describes the convolution of the respective functions, as we first calculated the intensity values for all f_i and thus calculate a function of the transmitted intensities depending on f_i. Afterwards this function can be used to calculate the transmitted intensity for a respective frequency f_i. Thus, mathematically, this is not only the product but indeed the convolution of the respective functions.

$$I_{A,B,INT}(f_i) = \int_{-\infty}^{+\infty} T_{A,B,INT}(f) S_i(f_i - f) df$$

$$I_{A,B,ATM}(f_a) = \int_{-\infty}^{+\infty} T_{A,B,ATM}(f) S_a(f_a - f) df$$

6. Page 7 Line 5: defects could be in the FPI mirror surface(s) (plural) right?

   R: revised. Please see **Sect 3, page 8 line 15-16**: “…, however, small defects on the FPI mirror surfaces or of the illumination of the FPI could result in small deviations, …”

   Line 7: Why not also mention/reference the works of Spinhirne, McGill.

R: revised. Thanks for your suggestion, the related reference has been added in the revised manuscript. Please see **Sect 3, page 8 line 15-16**.

"However, small defects on the FPI mirror surfaces or of the illumination of the FPI could result in small deviations that have to be considered for an accurate analysis (McGill et al., 1998)."

Reference:

McGill, M. J., and Spinhirne, J. D.: Comparison of two direct-detection Doppler lidar techniques. Opt. Eng., 37 (10), 2675-2686, https://doi.org/10.1117/1.601804, 1998.

Line 9: R is the mean reflectivity of the etalon mirrors? (again, plural?)

R: revised. Please see **Sect 3, page 8 line 13-14**: "R is the mean reflectivity of the mirror surfaces and"

Line 14: Suggest instead to say, "An easily calculated analytical expression...."

R: revised. Please see **Sect 3, page 8 line 19-20**: "An easily calculated analytical expression of Tenti S6 line shape model for,…."

Lines 16-21: The paper might read more easily if this paragraph was moved up earlier in the discussion.

R: revised. The sentence "The spectrally narrowband Mie scattering which is not filtered out by the Fizeau interferometer will enter the FPIs and can be considered as Mie contamination of the Rayleigh signal." has been moved to **Sect 2.1 Page 5 Line 23-24**. We didn't change the position of the rest of the paragraph in order to read more easily, because the mentioned variables need to be described firstly.

Line 21: The "magenta" filled area appears more "pink" – perhaps use that term instead, or "light magenta"

R: revised. It has been revised as "light magenta". Please see **Sect 3 Page 9 Line 8**.

7. Page 8 Line 1: Here the authors could clarify for the readers not familiar with double edge approach why the biases are worse when Mie signal is significant but not good enough to measure winds using the Mie channel. Can this be shown somehow in Figure 2?

   R: it has been revised as "According to previous studies (Dabas et al., 2008), it is implied that the Mie contamination correction could improve the quality of Rayleigh wind in the cases of intermediate ρ, e.g. below 1.5, as in this case the Mie signal is not high enough to guarantee an accurate Mie wind measurement but rather becomes significant for Rayleigh channel (Sun et al., 2014; Lux et al., 2018). Please see **Sect 3 Page 9 Line 14-17**.

   Line 4 (paragraph 2): clarify that the procedure is done assuming no Mie interference (or otherwise?)

   R: revised. Please see **Sect 3 Page 9 Line 29 to page 10 line 1:** "It is noted that the procedure is done assuming no Mie interference."

Line 9: The text says that the red-square marks +/- 850 MHz, but the figure looks like its closer to 1 GHz. Please rectify one or the other to match.

R: Figure 3 has been updated in the revised manuscript.

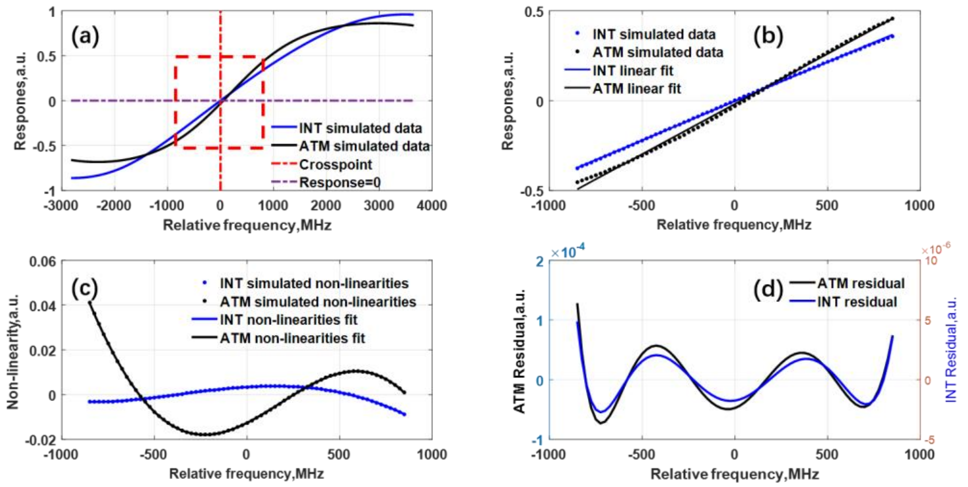

Figure 3: (a) The Simulated Rayleigh Response Calibration (SRRC) for internal reference (INT, blue line) and atmospheric return (ATM, black line), the frequency of the filter cross point is marked with a red dotted line, (b) INT (blue dots) and ATM (black dots) response and corresponding linear least squares fit (blue line for INT, black line for ATM) calibration with a frequency interval of ±850 MHz, where relative frequency is used instead of absolute frequencies, (c) the non-linearities of simulated (dots) and fitted (lines) response functions from INT (blue) and ATM (black). (d) response function residuals from INT (blue line) and ATM (black line).

Line 22: clarify that the “Then the fit of the SRRC for the internal reference and atmospheric paths can be expressed as a sum of a linear fit plus a 5th order polynomial:”

R: revised. It has been revised as “Then the fit of the SRRC for the internal reference and atmospheric paths can be expressed as a sum of a linear fit and a 5^{th} order polynomial fit, that is,” Please see **Sect 3 Page 10 Line 16-17**.

8. Page 9 Line 5: replace “In the frame of...” with “As part of..” The rest of this paragraph would benefit from additional proofreading for English grammar.

   R: revised. Please see **Sect 4 Page 11 Line 2: “**As part of the North Atlantic Waveguide and Downstream Experiment (NAWDEX) carried out in, …”

   Line 19-20: Suggest a rewrite to read “Time-space matching datasets between dropsonde and A2D can be used as both references to validated A2D wind measurements and to provide essential....”

   R: revised, it has been revised as “Time-space matching datasets between dropsonde and A2D can be used as both references to validate A2D wind measurements and to provide essential atmospheric temperature and pressure profiles for SRRC in this study.” Please see **Sect 4 Page 11 Line 16-18.**

   Line 23-24: This sentence repeats a little bit of what was written before, now referring to “illumination properties” - can you be more specific? Is this a function of differences in the spatial (e.g. the pupil) distribution or in the field (e.g. angular) distribution?

R: revised. It has been revised as "It is noted that the transmission functions of the FPIs are reproducible, and the transmission characteristics are different for the internal reference and atmospheric path due to the difference of the illumination of the FPIs in these two paths. This refers to the difference of divergence and incidence angles on the FPIs for the respective paths. It is both a difference in the spatial as well as in the angular distribution of the light. In particular, the use of a multimode fibre in the internal reference path gives rise to speckles, resulting in an intensity distribution which is markedly different from that of the atmospheric path." Please see **Sect 4 Page 11 Line 21-26.**

9. Page 10 Line 13-14: The authors state that the, "measured response values obtained from A2D wind velocity measurement mode are brought into the fitted SRRC...." What does "brought into" mean here? Is this a mapping? What is the process for doing this?

   R: revised. Change "brought into…" to "combined with…", the specific process for doing this is marked with red-line square in the figure below. Please see **Sect 5 Page 12 Line 15.**

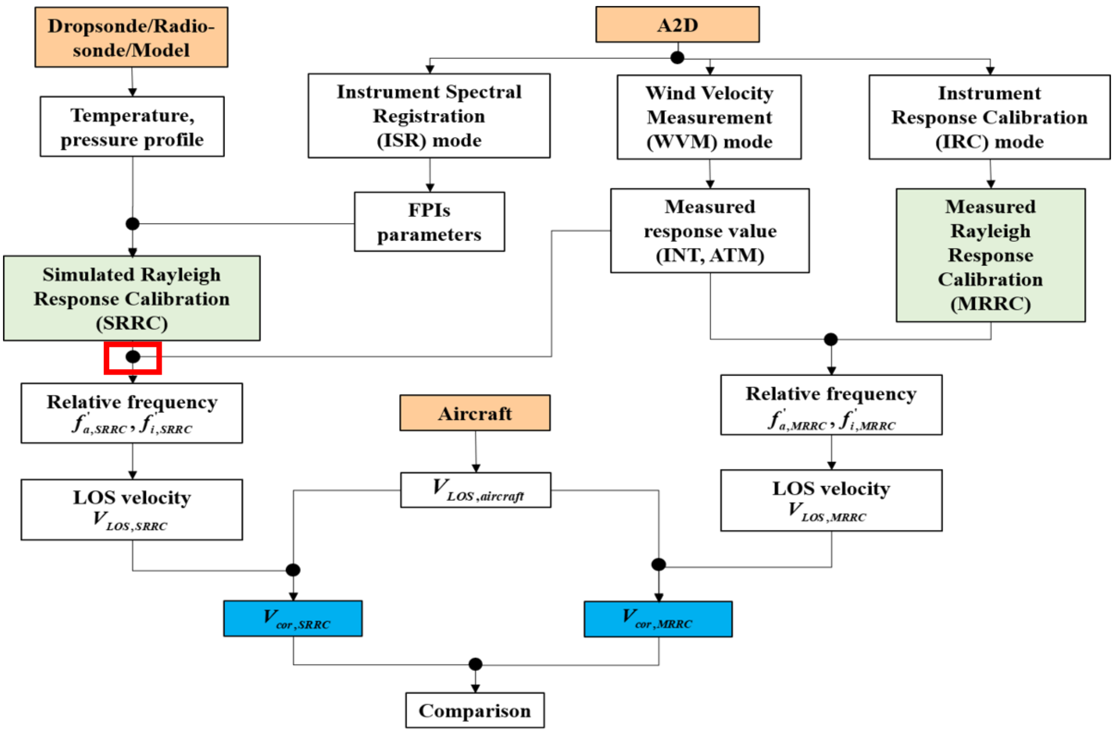

Figure 4: Flowchart of LOS velocity retrieval and comparison between A2D SRRC and MRRC.

Line 19-20: Add "and possible vertical velocity components" to the end of this first sentence.

R: revised. Please see **Sect 5 Page 12 Line 20-21: "…,** wind but also the contribution from the aircraft flight velocity and possible vertical velocity component."

10. Page 11 Paragraph 5.1: The figures described here would benefit from a diagram showing the campaign configuration.

    R: revised. The specific parameters of FPIs during different campaigns has been listed in Table 3.

Line 8-10: This mentions of the difference between ATMG and INTG due to different illumination. The reasoning for this should be described earlier in the paper and referenced back.
R: revised. The reasoning for this has been described in **Sect 2 Page 4 Line 14-22:**
"As shown in Fig. 1 in (Lux et al., 2018), a small portion of laser beam radiation is collected by an integrating sphere and coupled into a multi-mode fibre, then injected into the receiver via the front optics. The atmospheric backscattered signal is collected by a Cassegrain telescope and guided via free optical path propagation to the front optics and receiver successively. This path is called the atmospheric path. An electro-optic modulator is used to separate the atmospheric signal from the internal reference signal temporally in order to minimize the contamination of the internal reference signal with atmospheric signals and saturation of the detectors at short ranges (Reitebuch et al., 2009). Because of the different optical illumination of the internal path and atmospheric path resulting in different divergence and incidence angles on the FPIs, the response calibration curves for these two paths are slightly different."

The authors seem to change terminology back and forth throughout this section (and the corresponding figures) which makes reading the section slightly more challenging. Specifically, on page 8 the terms defined in Equations 8 and 9 are referred to as beta= sensivity and alpha=intercept, but in Figures 7 and 9 only the terms sensitivity and intercept are used. Perhaps adding the variable names beta_ATM and delta-alpha_ATM to the captions for figure7 and 9 would help. Likewise add the descriptive terms to the caption for figure 8.
R: revised.

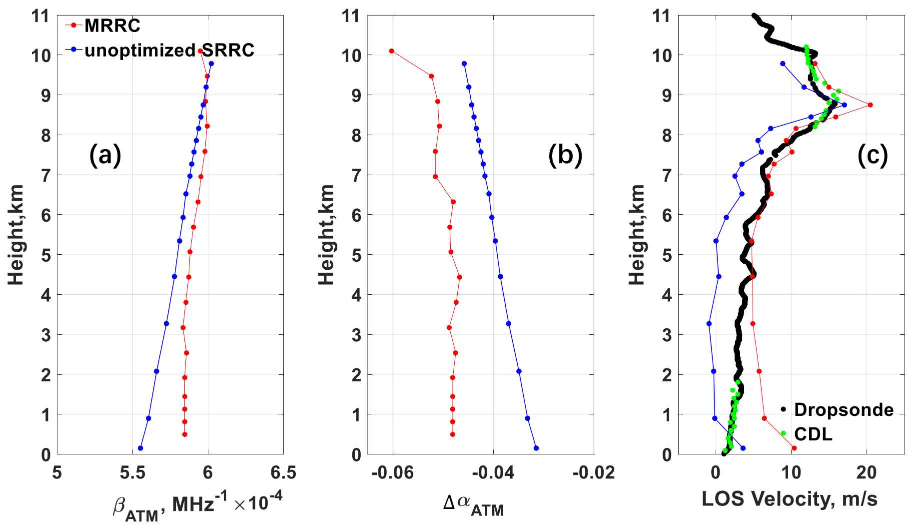

Figure 7: Case study using dropsonde data on 08:27:07 UTC, 23 September 2016: Comparison of (a) sensitivity β_{ATM} (MHz^{-1}) (b) $\Delta\alpha_{ATM}$ (c) LOS velocity between results from A2D Rayleigh channel MRRC (red) and not optimized SRRC (blue). The LOS velocity from dropsonde (black) and CDL (green) are also presented in Fig. 7 (c).

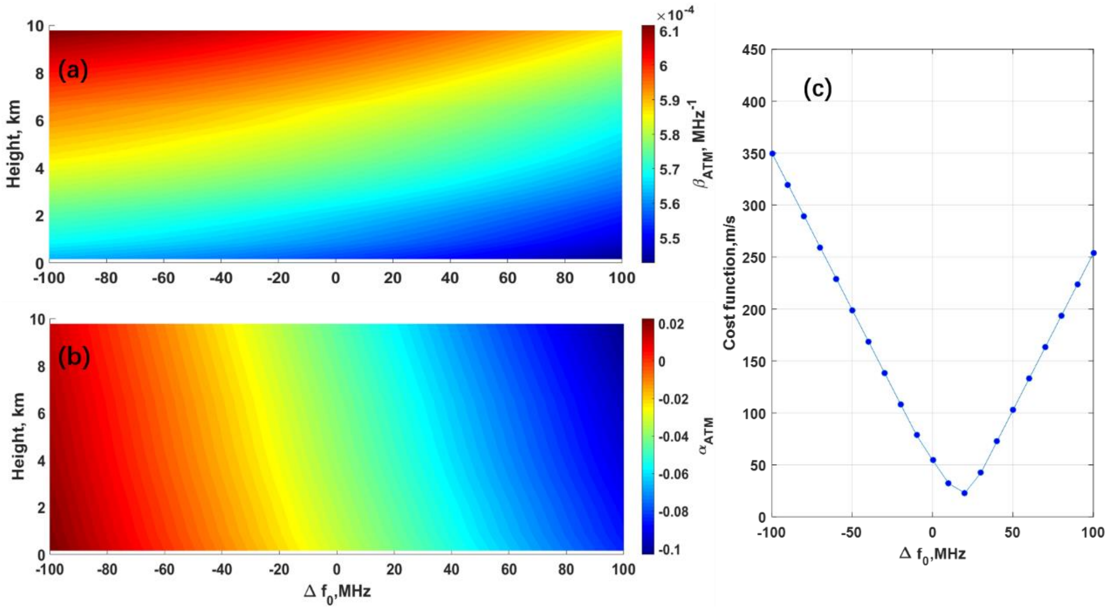

Figure 8: The effect of the centre frequency offset Δf_0 of filter A and B for atmospheric path on atmospheric response (a) $\boldsymbol{\beta_{ATM}}$ (b) $\boldsymbol{\alpha_{ATM}}$ and (c) corresponding cost function $F(\Delta f_0)$.

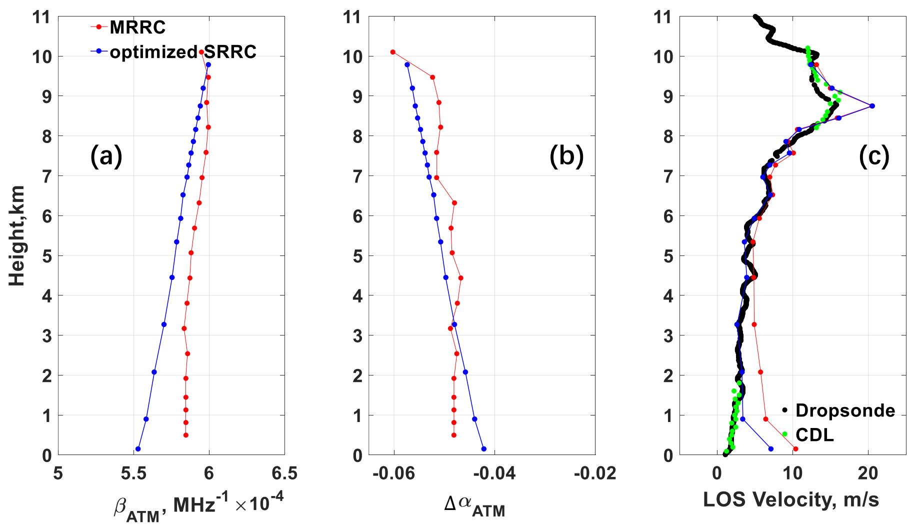

Figure 9: Case study using dropsonde data on 08:27:07 UTC, 23 September 2016: Comparison of (a) sensitivity $\boldsymbol{\beta_{ATM}}$ (MHz^{-1}) (b) $\boldsymbol{\Delta\alpha_{ATM}}$ (c) LOS velocity between results from A2D Rayleigh channel MRRC (red) and optimized SRRC (blue). The LOS velocity from dropsonde (black) and CDL (green) are also presented in Fig. 9 (c).

Line 10: What is the source of the atmospheric signal in the internal path on airborne testing (INTA)? Is there a delay in the internal reference path that causes the INTA signal to overlap with near field returns due to early overlap? Does multiple scattering play a role in these early returns?

R: revised. As the telescope and optical receiver is coupled via free optical path (and not via a fibre) the mechanical integration of the A2D inside the aircraft leads to small variation in position and incidence angle on the spectrometers for each deployment. The atmospheric contamination of the internal reference signal is caused by the limited suppression efficiency of the electro-optical modulator incorporated in the A2D front optics. This leads to a leakage of atmospheric backscatter being incident on the Rayleigh accumulated charge coupled device (ACCD), during the acquisition time of the internal reference signal. Please note that the internal path signal is recorded with the same ACCD detector as the atmospheric path signal, and a temporal resolution of 4.2 μs is used for the internal path signal. For the internal calibration that was performed on ground, the receiver was blocked and only the internal reference signal is used. For that reason, there is no contamination by atmospheric signal here. The related description has been added in **Sect 5.1 Page 13 Line 11-17**.

Lines 13-20: There are numerous papers discussing modeling of FPI performance. Perhaps some of these could also be referenced:
Jack A. McKay and David J. Rees "High-performance Fabry-Perot etalon mount for spaceflight," Optical Engineering 39(1), (1 January 2000). https://doi.org/10.1117/1.602361
P. D. Atherton, N K. Reay, J. Ring, and T. R. Hicks "Tunable Fabry-Perot Filters," Optical Engineering 20(6), 206806 (1 December 1981). https://doi.org/10.1117/12.7972819
J.A. McKay and David Rees , "Space-based Doppler wind lidar: Modeling of edge detection and fringe imaging Doppler analyzers"
Others by McKay, and Spinhirne, McGill, Gentry, etc.
R: revised. The related references have been added in the revised manuscript. Please see **Sect 5.2 Page 13 Line 20-21: "**The modelling of FPIs performance has been studied in the previous studies (McGill et al., 1998; McKay et al., 2000a; McKay et al., 2000b)."

References:
1. McGill, M. J., and Spinhirne, J. D.: Comparison of two direct-detection Doppler lidar techniques. Opt. Eng., 37(10), 2675-2686, https://doi.org/10.1117/1.601804, 1998.
2. McKay, J. A., and Rees, D. J.: High-performance Fabry-Perot etalon mount for spaceflight. Opt. Eng., 39 (1), 315-319, https://doi.org/10.1117/1.602361, 2000a.
3. McKay, J. A., and Rees, D.; Space-based Doppler wind lidar: modeling of edge detection and fringe imaging Doppler analyzers. Adv. Space. Res., 26(6), 883-891, https://doi.org/10.1016/S0273-1177(00)00026-0, 2000b.

Line 21: What is meant by the phrase, "Different from ALADIN" ? Were the ALADIN transmission curves (internal and atmospheric paths) never measured?
R: revised. For ALADIN, only the transmission curve in the internal reference path is measured during instrument spectral registration, and the transmission curve in the atmospheric path is modelled by a convolution of an Airy function and a tilted top-hat function (Dabas and Huber, 2017).

The related description has been revised as "Different from ALADIN, where only the transmission curve in the internal reference path can be measured during instrument spectral registration…" Please see **Sect 5.2 Page 13 line 28-29.**

11. Page 12 Equations 15 and 17 define variables "A" and "B" for the Atmospheric and Internal

paths, but this terminology is confused with the use of those variables as names for "Filter A" and "Filter B" (the two edge filters) per the labeling in Figures 1, 5, etc.

R: revised. The variables "A", "B" in equations 15 and 17 have been revised as "M", "N", respectively. Please see **Sect 5.2 Page 14 line 9-10:**

12. Page 13 Lines 11-13: This information could also be included in a previous section on the impact of angles on FPI transmission functions.

R: revised. The related introduction has been added in revised manuscript **Sect 2.2 Page 6 line 20-23:** "Furthermore, the FPIs transmission characteristics should be a function of incidence angles, field of view, temperature, pressure, thickness, and so forth, regardless of measurement or simulation method, any angular alignment drift will change in the incidence angles on the FPIs, resulting in a different transmission value." Please see **Sect 2.2, page 6 line 30-31 to page 7 line 1-2**.

Line 13: "Assuming the center frequencies of filter A and B have the same offset..." Are there any challenges to this assumption? If angles get larger, does the center frequency shift more for A vs. B? A diagram (or a reference to a paper with a diagram) of the two paths through the system might help confirm that the offset is the same.

R: revised. The Rayleigh spectrometer is composed of two FPIs which are sequentially coupled. Thus, the reflection of the directly illuminated first FPI is directed to the second FPI. Any incidence angle change in front of the Rayleigh spectrometer will act similarly on both FPIs. Considering that the initial condition was perpendicular incidence, both FPIs are affected similarly regarding a shift in the center frequency. Furthermore, as angular shifts of only a few µrad are expected to occur, large angles do not have to be considered. Considering these points, it is justified to consider the same offset for the center frequencies induced by small incidence angle changes. The related description has been added in **Sect 5.3 Page 15 Line 23-28**. The specific schematic of ALADIN Airborne Demonstrator (A2D) was shown in Fig. 1 in (Lux et al., 2018)

Lines 15-20: The text refers to the plots in Figure 8 and talks about range gates, but the figure shows altitude bins. Which terminology should be used?

R: revised. Replace "range gate" with "altitude bin" in the revised manuscript. Please see **Sect 5.2 Page 14 line 18-20:** "The responses of internal reference (red) and the 8th atmospheric altitude bin (blue dashed line, the corresponding height is around 5.7 km) from measurement and corresponding SRRCs are listed in Table 4 and shown in Fig. 6."

Line 20: "all available range gates ... are used to calculate the cost function..." – does this assume there is no aerosol present in this data set?

R: revised. "all available altitude bins" means all altitude bins of MRRC shown in Figure 7, from i=1 to N=17. The altitude bins affected by aerosol or cloud layers are hard to be flagged, unless there are auxiliary information such as CDL measurements. Therefore, these bins affected by Mie contamination are also taken into consideration in the calculation of $F(\Delta f_0)$.

The related description has been added in **Sect 5.3 Page 16 Line 6-9.**

13. Page 14 Lines 28-29: The sentence, "However, the temperature difference between MRRC and the actual wind measurement must..." is confusing. Perhaps the authors meant, "However, differences in the atmospheric temperature profile between when the MRRC was obtained and when the actual wind measurements were acquired are a known important source of wind bias, which are especially severe in cases of large temperature differences."
R: revised. Thanks for your suggestion, we didn't explain it accurately. It has been revised as "However, the atmospheric temperature affects the Rayleigh-Brillouin line shape, and has a direct effect on the instrument response calibration (Dabas et al., 2008). Differences in the atmospheric temperature profile of the time and location when the MRRC was obtained and the actual wind measurements are important sources of wind bias, which are especially severe in case of large temperature differences.", please see **Sect 1 Page 3 line 15-18.**

Lines 21-33 (and line 11 on page 15): This issue is the basis for all the work done in this paper, right? So this should be right up front in the beginning of the paper, to help the reader understand why the work is being done and described.
R: revised. Yes, this issue is the basis of this paper. The related description has been moved to **Sect 1 Page 3 Line 14-22**, as shown below:
"Currently, only measured Rayleigh response calibrations (MRRC) are used for the A2D (Marksteiner, 2013; Lux et al., 2018; Marksteiner et al., 2018). However, the atmospheric temperature affects the Rayleigh-Brillouin line shape, and has a direct effect on the instrument response calibration (Dabas et al., 2008). Differences in the atmospheric temperature profile of the time and location when the MRRC was obtained and the actual wind measurements are important sources of wind bias, which are especially severe in case of large temperature differences. This is the reason why it is mandatory to consider the atmospheric temperature in the Aeolus level 2B procedure to retrieve reliable winds (Dabas et al., 2008; Rennie et al., 2017). Furthermore, some experimental limitations, which will be introduced specifically in Sect. 2.1, need to be considered carefully to achieve a reliable MRRC. Overall, the atmospheric and instrumental variability coming along with a MRRC limits the reliability and repeatability of A2D instrument response calibrations."

14. Page 15 Line 11: This is the key point of the paper, but it is muddled a little due to grammar. Perhaps say "This is one of the limitations of the A2D MRRC approach which can be overcome using the SRRC approach"
R: revised. Please see **Sect 6 Page 18 line 15-16: "** ,…and this is one of the limitations of the A2D MRRC approach which can be overcome using the SRRC approach"

Line 15-17: Can you be more specific than saying the response calibration is affected directly? Perhaps say that the aerosol spectrum shifts the centroid of the atmospheric/filter transmission product, thereby biasing the wind speed estimates (or something like that)?
R: revised. It has been revised and updated in **Sect 1 page 5 line 23-27**:
"the spectrally narrowband Mie scattering which is not filtered out by the Fizeau interferometer will enter the FPIs and can be considered as Mie contamination of the Rayleigh signal. Because

of the different spectral widths of the particle and molecular backscatter signal, the sensitivities of the FPIs on them are different. The Mie contamination on Rayleigh channel is one of the sources for systematic errors because it modifies the instrument response calibration curve, which should be avoided to ensure the representativity of pure Rayleigh response."

Line 25: "Indeed, the Mie contamination...." – this is another key point for the paper and justification for doing the SRRC. While a detailed discussion might not be within the scope of the paper, the paper would benefit greatly from some discussion as the topic of Mie contamination comes up several time.

R: revised. Thanks for your suggestion, we have updated related paragraphs, as shown in the reply to **Major comments #3**.

As the basis of this paper is the effect of atmospheric temperature and pressure on calibration response curve, Mie contamination correction is not our major topic of investigation in this paper, although it is another strength of the SRRC. The other reason why we didn't discuss Mie contamination deeply in this paper is that it needs ρ value as input, and it needs to be determined with the Mie channel signal.

15. Page 16 Line 17: remove the "are" from the beginning of the line.
    R: revised. Please see **Sect 7 Page 19 line 21.**

    Line 28: "overcame" should be "overcome" here.
    R: revised. Please see **Sect 7 Page 19 line 32.**

16. Page 17 Line 1: The sentence should probably read, "Overall, the SRRC allows correction for variability in atmospheric and temperature profiles, when known, ..."
    R: revised. It has been revised as "The SRRC allows to correct for variability in atmospheric temperature and pressure profiles, giving accurate wind retrieval especially in cases of large temperature differences between when the MRRC was obtained and when the actual wind measurements were acquired. Furthermore, SRRC is more accessible as the procedure doesn't need to meet the strict experimental requirement as MRRC's. It can also overcome the possible limitations induced by elevated ground altitudes, improving the accuracy of A2D wind measurements at lower altitudes. Overall, it can improve the reliability and repeatability limitations caused by atmospheric and instrumental variability and constraints during A2D MRRC process. Further studies based on A2D SRRC will be performed regarding the atmospheric temperature/pressure effect, Mie contamination correction and the particulate optical properties retrieval." Please see **Sect 7 Page 20 line 24-31 to page 21 line 1-2.**

17. Figures
    Figure2: Please also use the variable name (e.g. "Rx") with "Response" (in the caption and the axis labels)
    R: revised.

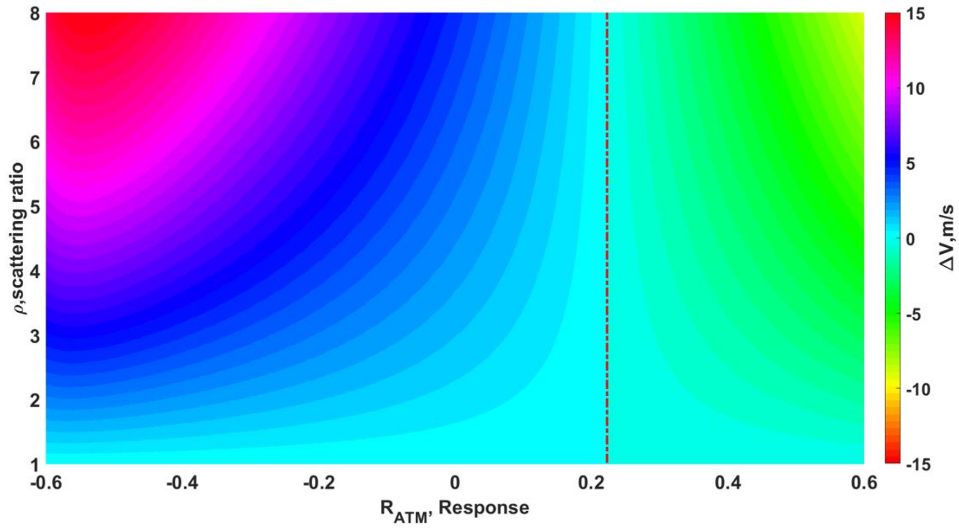

Figure 2: Simulation of LOS wind velocity errors ΔV generated by Mie contamination at T=223 K and P=301 hPa. The x-axis and y-axis represent the response value R_{ATM} and scattering ratio ρ, respectively. The red dashed-line corresponds to the response value with minimum ΔV at each scattering ratio.

Figure 3: Again, refer to the variable fc when discussing the cross point frequency.

R: revised.

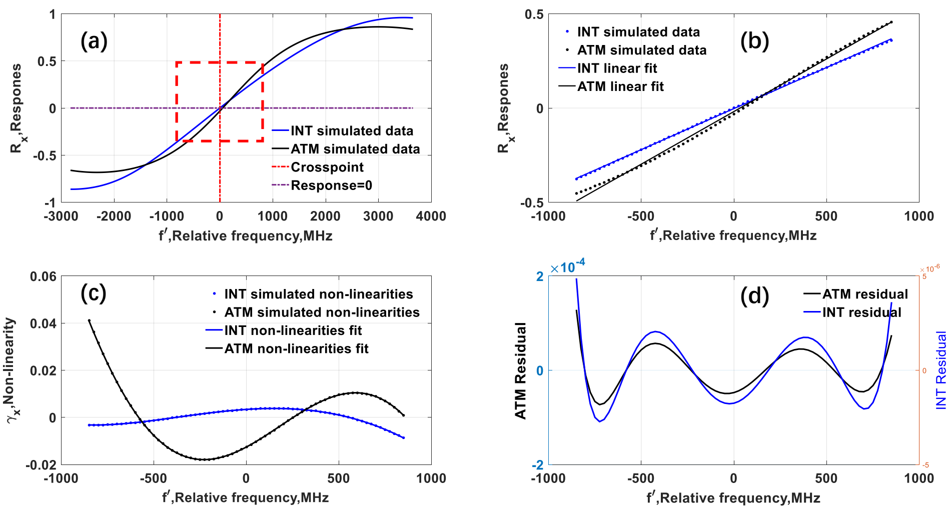

Figure 3: (a) The Simulated Rayleigh Response Calibration (SRRC) for internal reference (INT, blue line) and atmospheric return (ATM, black line), the frequency of the filter cross point is marked with a red dotted line, (b) INT (blue dots) and ATM (black dots) response and corresponding linear least squares fit (blue line for INT, black line for ATM) calibration with a frequency interval of ±850 MHz, where relative frequency is used instead of absolute frequencies, (c) the non-linearities of simulated (dots) and fitted (lines) response functions from INT (blue) and ATM (black). (d) response function residuals from INT (blue line) and ATM (black line).

Figure 5: Should the blue dashed curve be labeled “TB from INTA” (vs. TA) ?

R: revised.

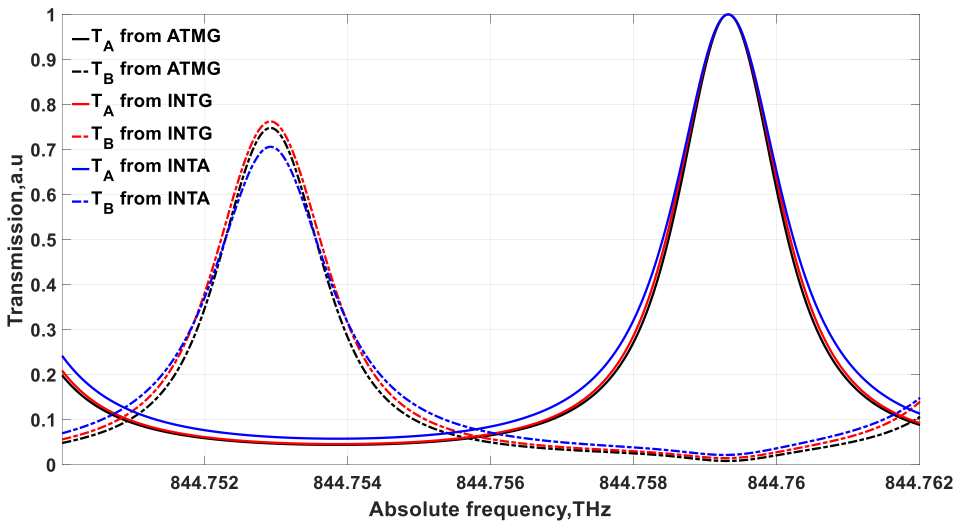

Figure 5: The fitted transmission functions of the FPIs from different campaigns, detection channels and illumination situations. The black, red and blue groups are obtained from ATM path measurement during BRAINS ground campaign (ATMG) in 2009, INT path measurement during NAWDEX from ground (INTG) in 2016 and INT path measurement during NAWDEX airborne measurement (INTA) in 2016, respectively.

Figure 6: The authors could clarify for the reader that the MRRC lines are repeated throughout the plots, e.g. say "(red and blue dashed-lines, respectively, same on every plot)"
R: revised.

Figure 9: clarify that (c) represents the retrieved LOS velocity
R: revised.

Figure 13: Can the authors say anything about the potential presence of vertical velocities and their impact on the comparison?
R: revised. Generally, the presence of vertical velocity has an effect on two main aspects:

1. During response calibration: for deriving the frequency dependency of the Rayleigh and Mie channel spectral response, a frequency scan of the laser transmitter is carried out, thus simulating well-defined Doppler shifts of the radiation backscattered from the atmosphere within the limits of the laser frequency stability. During the calibration, the contribution of (real) wind related to molecular or particular motion along the instruments' line-of-sight (LOS) has to be eliminated, i.e. the LOS wind speed v_{LOS} needs to be zero. In practice, this is accomplished by flying curves at a roll angle of the Falcon aircraft of 20°, resulting in approximate nadir pointing of the instrument and hence $v_{LOS} \approx 0$, while assuming that the vertical wind is negligible. Consequently, regions with expectable non-zero vertical winds, e.g. introduced by gravity waves or convection, are avoided during response calibration, otherwise, it will result in incorrect response calibration curve.
2. During wind measurement: the measured LOS velocity v_{LOS} is defined as the projection of horizontal wind vector on this direction without vertical velocity contribution. When vertical velocity is not negligible, v_{LOS} is the sum of the projection of horizontal wind vector and vertical velocity.

Can the authors provide error bars on the LOS velocity retrievals? Even CDL systems have errors.

R: revised. The error bars of LOS velocity derived from MRRC and SRRC can be seen in Figure 12 (b), (c), respectively. The CDL provides high performance with accuracy of <0.3 m/s and precision of <1 m/s, respectively (Chouza, F. et al., 2016), thus we prefer to plot no error bars to the CDL measurements. Please see **Sect 6 Page 18 Line 9-11**.

An estimation of the accuracy and the precision (also considering the representativeness error estimated by means of radiosonde comparisons) can be found in Chouza et al., 2016.

Reference: Chouza, F., Reitebuch, O., Jähn, M., Rahm, S., Weinzierl, B.: Vertical wind retrieved by airborne lidar and analysis of island induced gravity waves in combination with numerical models and in situ particle measurements, Atmos. Chem. Phys., 16, 4675–4692, 2016.

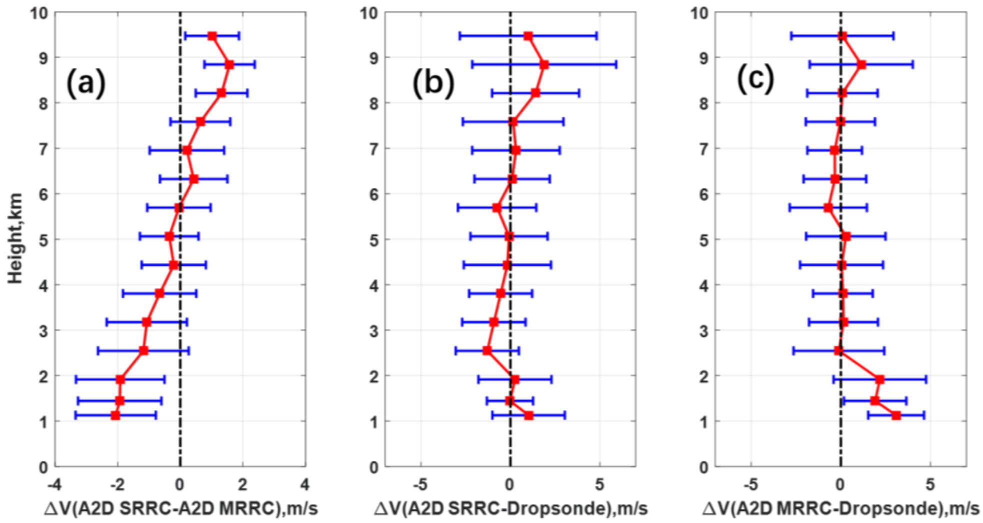

Figure 12: Comparison of profiles for LOS velocity (a) between A2D SRRC and MRRC (b) SRRC and dropsonde (c) MRRC and dropsonde.