B. Witschas 13.01.2023

The very detailed reading and the corresponding mindful comments of anonymous reviewer #2 (Referee #2, received on 20 October 2022, shown in black) are highly appreciated, as they significantly contribute to improving the paper manuscript. Based on the given comments (and the one of Referee #1), we revised the paper manuscript extensively, especially chapter 4 (results). Our corresponding answers to the reviewer's comments are given below in blue, and the performed changes in the manuscript are highlighted in green.

Interactive comment on "Airborne coherent wind lidar measurements of the momentum flux profile from orographically induced gravity waves" by B. Witschas et al.

(Author response)

Reviewer statement:

The paper by B. Witschas and co-authors presents a method to measure gravity wave momentum fluxes from line-of-sight wind observations using an airborne 2-micron Doppler wind Lidar (DWL). The technique is applied to measurements gathered over a flight leg of the Falcon aircraft during the GW-LCYCLE II campaign in Scandinavia (2016). Lidar wind and momentum flux retrievals are compared with colocated in situ measurements by the HALO aircraft available for this case study. While the manuscript is well-written and the topic of interest to AMT readership, a major point of criticism I have is that a significant part of the material presented was already published by two of the authors of the present manuscript in Gisinger et al. (2020, article cited). Some figures are very similar to that previous study. In this context, I would expect the present manuscript to provide a thorough description of the instrument and its performance. On the contrary, although the method is sound and the results very impressive, the technical discussion remains superficial, in particular regarding the advantages of this new measurement mode with respect to other scanning modes and the uncertainties of the technique. I appreciate that earlier papers by the authors may already provide some of this information, but it would be necessary to repeat some of the details here. For those reasons, the paper should be reconsidered after major revisions.

Thanks a lot for this comment. Indeed, the primary goal of this paper is to address the measurement technique, corresponding algorithms, and reached accuracies, rather than discussing the scientific results as was done by Gisinger et al., 2020. Still, we decided to keep at least the MF-flux profile plot as an application example and differences between HALO in-situ and lidar are investigated in more detail (new Fig. 10). Furthermore, we added a whole paragraph about the LOS wind retrieval and added the analysis of the lidar ground return signals which indirectly demonstrated how accurate the measurements (and the flight attitude control loop for the scanner) work. Furthermore, to reduce repetitions with the work presented by Gisinger et al., 2022, we removed the figure with the wavelet analysis, as it does not provide additional information for the discussion. A detailed listing of the performed changes in the manuscript is given below.

Major comments :

1) I find that the paper lacks an a priori estimate of wind error and resulting momentum flux noise level. Granted, the agreement with in situ measurements is very good (for wind), but there may be sources of difference other than instrumental errors (e.g., slightly different timing inducing a phase shift). Would it not be possible to estimate the expected error, even roughly, and compare with the empirical estimate? This point should at least be discussed.

Thanks a lot for raising that point. As shown by the direct comparison of the Lidar and the in-situ data for u and w, the slight temporal and spatial occurrence of the respective measurements do not induce a

significant error. However, small deviations are obvious, especially in the range between 14°E and 15.8°E. We realized that, in this region, the HALO ground speed and with that, flight attitude changed significantly by a pitch angle change from 1.4° to 2.0°. Also, for the Falcon, the pitch angle changed in that region. To verify if this has affected our lidar measurement accuracy, we additionally analyzed the ground returns which should result in a LOS wind speed of 0 m/s. And indeed, no significant discrepancies from 0 m/s could be observed. On the other hand, the differences between Lidar and Halo in-situ data could be correlated with the observed changes in the HALO pitch angle. Thus, it is likely that the in-situ measurements are slightly affected by the changing flight attitude.
Regarding the ground return analysis, we added the following figure and the corresponding paragraph:

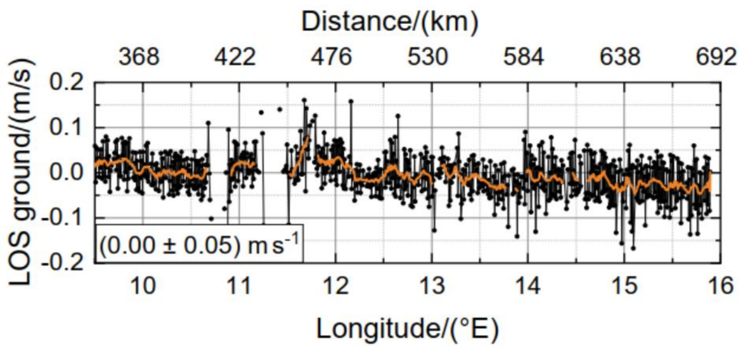

Figure 7. Line-of-sight wind speed retrieved from 2-μm DWL ground returns (black) and moving average of 10 successive data points (orange).

In contrast to the in-situ measurements, airborne wind lidar measurements enable to estimate the accuracy of the performed wind observations by analyzing the returns from the non-moving ground which should yield values of 0 m/s by definition. For the discussed flight leg, ground returns are available from about 9.5° E to 15.8° E as shown in Fig. 7, black. A moving average of 10 successive data points is additionally indicated by the orange line. It can be seen that the LOS wind speeds are nicely varying around 0 m/s with peak-to-peak variations mostly below 0.1 m/s. The mean of this data set yields a value of 0.00 m/s with a standard deviation of 0.05 m/s, further demonstrating the high accuracy of the 2-µm DWL measurements during this flight. Furthermore, the flight attitude changes between 14° E to 15.5° E do not induce any remarkable systematic error, confirming that the discrepancy of 2-µm DWL and HALO in-situ data is unlikely to be caused by the lidar measurements.
In addition, we added an estimate of the MF uncertainty based on the random errors of u and w according to Smith et al., 2016. Here, it turns out that the error induced by the not defined flight leg length is larger than the one originating from the random error of the lidar measurements. In order to clarify this situation, we added the following paragraph to the manuscript:
The uncertainty of the derived momentum fluxes can be estimated as suggested by Smith et al. (2016). As the mean values are removed before calculating MFx, the corresponding uncertainty of MFx only arises from the random errors of u_{par} and w. To propagate these errors for the momentum flux, a transect with anti-correlated sinusoidal u′ par and w′ oscillations is considered with common amplitudes of 5.0 m s−1 and 1.0 m s−1, respectively. Furthermore, a mean air density of ρ = 0.5 kg m−3 for the altitude range between 7.0 km and 10.0 km is assumed. This enables to define a reference value of the leg-averaged momentum flux $MF_x = \rho\langle u'_{par}w'\rangle$ = (0.5)(0.5)(5.0)(1.0) = 1.25 Pa. In a worst-case situation, the random errors $\sigma u'_{par}$ = 0.5 m s−1 and σw′ = 0.17 m s−1 would translate to a corresponding error in the amplitude of the u'_{par} and w′ oscillation. This would lead to MF_x errors of 0.5/5 = 10% and 0.17/1 = 17%, respectively. Assuming that these errors are random and uncorrelated, the MF_x error reduces in proportion to the number of samples through the wave. Considering waves with a length of 8 km and larger, each wave is sampled with at least 10 points, reducing the relative error by a factor of F = 10−1/2 ≈ 0.32. Hence, the relative errors in MF_x are about 3.2% and 5.5%, where the true error probably lies in between these two values.

Furthermore, the authors do not explain the preliminary steps involved in LOS wind retrievals (e.g., estimating Doppler shift, subtracting the aircraft ground-relative speed). The potential impact of aircraft motions and associated uncertainty could also be discussed in more detail.

Thanks a lot for this comment. Originally, it was planned to just concentrate on the MF-mode retrieval, as the other processing steps were explained in detail in previous publications for instance in Chouza et al. 2016 and Witschas et al., 2017. But we agree that further information is helpful for the reader, especially in a measurement-technique-related paper such as this one. For this reason, we extended the instrument description with a condensed explanation of the LOS processing steps. The impact of the aircraft motions and other noise sources is furthermore characterized by means of the ground return analysis that was also added to the manuscript (Fig. 7 of the new manuscript and corresponding explanation). Section 3.1 was extended by the following paragraph:

…. harsh environment of an aircraft. To do so, the beat signal of the emitted laser pulse and the local oscillator (seed laser) is analyzed. When the beat frequency or the laser pulse built-up time exceeds a certain difference from their nominal values, the signal from the respective laser pulse is not considered for accumulation. Moreover, before accumulating the respective reference pulse spectra, they are frequency-shifted to a defined reference value to correct for pulse-to-pulse frequency variations and thus avoid spectral broadening in the accumulation process. A similar frequency shift is also applied to the atmospheric signal power spectra. Afterwards, the signals from the remaining laser pulses in a one-second time interval are spectrally averaged. The part of the detector raw signal containing the atmospheric return is divided into segments that lead to 100-m range gates in the vertical by considering the actual laser beam pointing angle, the aircraft altitude and attitude, and the reference pulse timing. After that, the power spectrum is calculated for each range gate and laser pulse, is frequency shifted according to the reference pulse frequency shift and, subsequently, accumulated. The detector signal at the end of the record (below the ground return) is used to analyze and correct the detector noise characteristics. The power spectrum of the noise signal is calculated for each single laser pulse and is additionally averaged for a one-second time interval. Consequently, the power spectrum for each single range gate is divided by the respective noise spectrum to correct for the system noise and the receiver frequency response (noise whitening). In the next step, the power spectra are corrected for the aircraft speed projected onto the LOS direction, which is derived from the ground speed measured by the GPS module and the actual laser beam pointing direction. The latter is calculated from the actual scanner position, the aircraft attitude (inertial reference system) as well as the lidar installation position with respect to the aircraft reference frame (Chouza et al., 2016). Finally, the LOS wind speed v_{LOS} is calculated according to $v_{LOS} = (\Delta f\, c)/(2\, f_0)$, where Δf is the frequency shift between the reference pulse and the atmospheric signal, f_0 is the laser frequency, c is the velocity of light, and $\lambda_0 = c/f_0 = 2022.54$ nm is the laser wavelength. To retrieve vertical or horizontal wind information from respective LOS wind measurements, further processing steps are needed, as discussed in Sect. 3.2.

2) The comparison between Lidar and in situ wind observations could be more thorough. For instance, what is the power spectral density of the differences in Fig. 6? Are there specific artifacts at given frequencies?

The analysis of the respective u' and w' power spectra did not reveal any significant differences between in-situ and lidar measurements. As mentioned above, we determined small deviations in the range between 14°E and 15.8°E and realized that the HALO aircraft pitch angle changed from 1.4° to 2.0° in this region. Thus, we could not determine given frequencies that may explain the differences between lidar and in-situ measurements but could allocate the region where the differences occur. Furthermore, the analysis of the 2-µm DWL ground returns gives further confidence that the lidar measurements are accurate, as the shown deviations in the range of only 0.05 m/s. Moreover, we have split our analysis in two different wave classes of short waves (smaller than 30 km) and long waves (larger than 30 km). By doing so, it can be shown that the discrepancy of the retrieved MF values is only true for the long wave class. To further address that issue and to visualize where the actual deviations between both data sets appear, we added a u’w’ plot for both data sets and both wave classes as shown below:

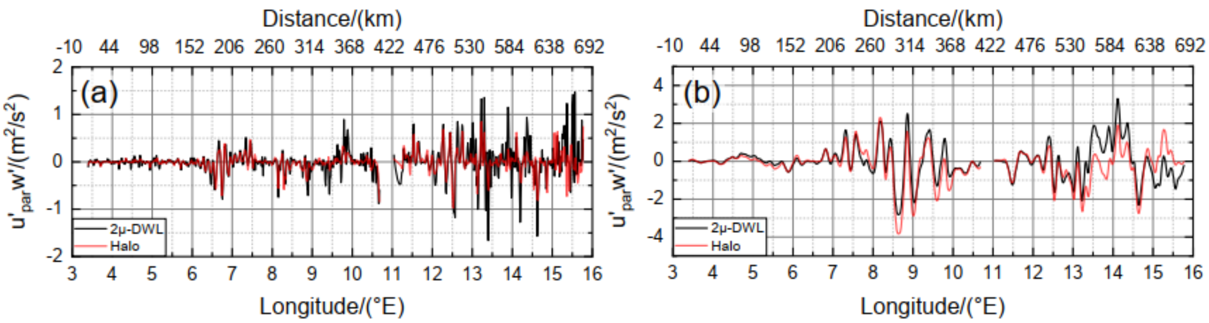

Figure 10. $u'_{\mathrm{par}}\, w'$ measured by the 2-μm DWL (black) and Halo (red) at 7.8 km for the short-wave case (a) and the long-wave case (b), respectively.

The following paragraph was also added to the manuscript:
Furthermore, for the long wave analysis it can be seen that the MFx ≈ −0.1 Pa derived from HALO in-situ measurements at 7.8 km altitude (magenta) differs significantly from the MFx ≈ −0.02 Pa derived from the 2-µm DWL (blue), which is surprising, as the respective horizontal and vertical wind speed measurements are in great accordance (see also Fig. 6). To investigate the root cause of this discrepancy, the u’ par w’ data for the 2-µm DWL (black) and Halo (red) are compared for the short-wave class and the long wave class as shown in Fig. 10 a, and b, respectively. For the short-wave case, both time series are undulating around zero and with similar amplitudes, thus, providing comparable leg-averaged momentum fluxes. For the long-wave case, however, the u’par w’ patterns look comparable but show distinct deviations between 8.5° E and 10° E, and between 13° E and 15.5° E, where the 2-µm DWL data tendentially provide more positive values compared to the Halo in-situ data. Though the root cause of these deviations is not unequivocal proven, it is likely that they are caused by the discrepancy between both data sets that appear between 14◦ E and 15.5◦ E, where the Halo ground speed and attitude are shown to change (see also Fig. 6). As the 2-µm DWL ground returns in this region do not show an enhanced systematic error, the 2-µm DWL results are considered to be reliable here (see also Fig. 7).

Other comments :
Please double check Eq. 2. I obtain the same result as Referee 1, different from yours.
Thanks a lot for reviewing Equations (1) and (2) so carefully as indeed two typos were present in Equation (2). Hence, Equation (2) was corrected according to:

$$u_{\mathrm{par}}(x,R) = \csc(\theta_{\mathrm{f_i}} - \theta_{\mathrm{b_i}})(v_{\mathrm{f_i}}\cos\theta_{\mathrm{b_i}} - v_{\mathrm{b_i}}\cos\theta_{\mathrm{f_i}})$$

$$w(x,R) = \csc(\theta_{\mathrm{f_i}} - \theta_{\mathrm{b_i}})(-v_{\mathrm{f_i}}\sin\theta_{\mathrm{b_i}} + v_{\mathrm{b_i}}\sin\theta_{\mathrm{f_i}})$$

Furthermore, it was verified that all the data processing was performed with the correct equations.

Line 148-149 : Could you elaborate a bit on the ‘scanner control loop on a 1-second basis’? What is the uncertainty in attitude and how does it translate in LOS wind uncertainty ? Are pilot oscillations of attitude present? If yes, at which frequency? Are they sufficiently resolved at 1 s?
Thanks a lot for the question. The flight attitude in general changes only slightly in a one-second time frame and it depends on the actual pointing direction in the attitude changes and how this will translate into a LOS error. The best indication of the accuracy of retrieved LOS winds for airborne wind lidars is the analysis of ground returns (over land) if available. During the discussed research flight we were lucky and had a certain number of usable ground returns (see figure below). These measurements demonstrate a

mean LOS wind speed for the ground returns of 0 m/s and a random error of 0.05 m/s. Furthermore, the peak-to-peak fluctuation is more or less always less than 0.1 m/s, which means that potential attitude oscillations within the one-second time interval lead to an error of less than 0.1 m/s. The flight legs are flown with auto-pilot.

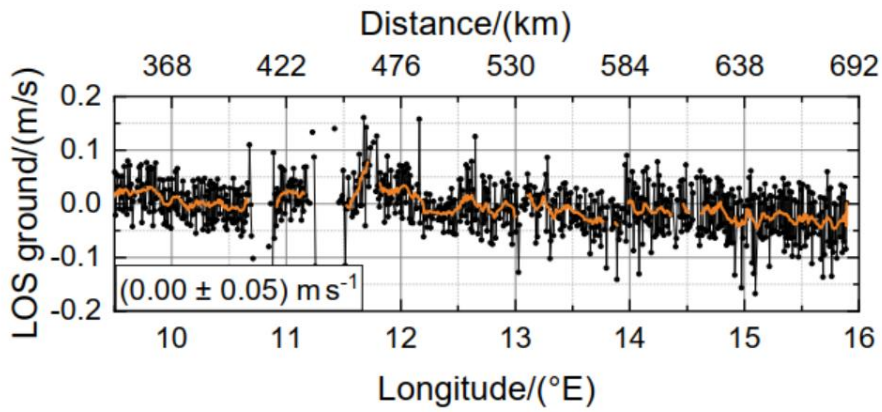

Figure 7. Line-of-sight wind speed retrieved from 2-μm DWL ground returns (black) and moving average of 10 successive data points (orange).

Line 154-155 : If I understand correctly, the 'wind vector mode' already enables computation of the momentum flux, do you confirm?

Yes, this is true, however, the horizontal resolution of about 8 km is a factor of ten lower. Furthermore, the representativeness error is expected to be worse for the wind-mode measurements. For further clarification, we modified the previous sentence accordingly. And in addition we added the following:

- … "However, it was also discussed that simultaneous measurements of the horizontal and the vertical wind speed with a horizontal resolution of a few hundred meters and a low representativeness error would be even more beneficial as such measurements would allow retrieving the vertical flux of horizontal momentum induced by GWs (Smith et al., 2008, 2016)."
- Commonly, the 2-μm DWL is used to either measure the three-dimensional wind vector (wind-mode) or rather to measure the vertical wind speed (vertical-wind-mode). When operating in wind-mode, the velocity-azimuth display (VAD) scan technique is applied (Browning and Wexler, 1968) by performing a conical scan around the vertical axis with an off-nadir angle of 20°. Typically, one scanner revolution with 21 line-of-sight (LOS) measurements separated by 18° in the azimuth direction takes about 42 s. By further considering the aircraft speed of about 200 m s-1, the horizontal resolution of wind-mode observations is about 8.4 km, depending on the actual ground speed of the aircraft. Hence, the vertical velocity wind field, which is shaped by scales of tens of kilometers and smaller in the presence of MWs (Smith and Kruse, 2017), is not well resolved in wind mode data.

Line 160 : 'the wind field is constant for the time of intersecting fore and aft laser beam pairs' : specify the length of that delay in practice

Indeed, this information was missing. We added the information about the usual duration between intersecting beams as follows:

The assumption of a constant wind field is justifiable, considering that the duration between intersecting beams is about 36 s on ground and respectively less for altitudes closer to the aircraft for the actual measurement configuration (off-nadir angle of 20°, flight altitude of 9.8 km, aircraft speed of about 200 m/s, and LOS averaging time of 2 s).

Line 164 ‘ kept for 2 s ’ : Why is it necessary to have 2 s (1000 laser pulses) when 1 s is enough for the vertical wind retrieval ?

Principally, it was not necessary to keep each LOS viewing direction for 2 s. However, the aerosol load was rather poor during the GW-CYCYLE II campaign. Furthermore, the time to move the double-wedge scanner from +20° to -20° takes about 0.3 s. As we only had one flight to test the MF-mode, we want to make sure to have sufficient data coverage for the measurements and thus, decided to stare at each LOS for 2 s. For future applications it is foreseen to the average for 1 s, further reducing the horizontal resolution to about 400 m. For clarification, we added the following sentence:

Due to the rather clear air conditions during the flight, it was decided to keep each LOS pointing direction for 2 s, instead of the usually used 1 s time interval, to assure reasonable data coverage. … … For future applications, it is foreseen to reduce the time interval of each pointing direction to 1 s leading to a horizontal resolution of ≈ 400 m.

Line 234 : A comparison with a priori estimates of error would be valuable here

This is somehow difficult as for instance, the representativeness error of the MF-mode measurements is not easy to estimate. At the beginning of section 4.2, the systematic and random error of 2µ-DWL LOS measurements and wind-mode measurements are given (determined from dropsonde comparisons within the last two decades). Based on these results, one would expect that the MF-mode measurements yield a systematic error of only a few cm/s and a random error of up to 1 m/s for the horizontal wind speed. Thus, the obtained results are in the order of what is expected.

Figure 8 and associated discussion : I am not certain how to interpret the quantity shown. The main signature in this u'w' product is that of the high-frequency vertical wind oscillations in a lower frequency horizontal wind (Fig. 7 b), but this contribution cancels out over a period and does not contribute to the momentum flux. I would recommend showing wavelet co-spectra of u and w, as in the Gisinger paper (Fig. 7).

In order to reduce repetitions to the paper by Gisinger et al, 2020, we decided to skip the wavelet analysis plot. Also because it does not add further valuable information to the performed discussion. In Gisinger et al., more data is used to discuss the actual dynamical situation and the corresponding GW generation and propagation in more detail. Hence, we decided to keep the focus on the measurement method (we was significantly extended with respect to the original manuscript) but excluded the wavelet analysis or potential co-spectra analysis (as also shown in Gisinger et al, 2020).

Fig. 7 is very similar to Figure 9 of Gisinger et al. (2020).

In order to demonstrate the usefulness of the suggested MF-mode measurements, we decided to keep the figure, but now have the two u'w' plots (c) and (d) for the two different wave-classes. Hence, the plot is now indeed different to what is shown in Gisinger et al. Further, we acknowledge the work by Gisinger at two more places in the manuscript:

- Abstract: …, which are induced by interfacial waves as recently presented by Gisinger et al. 2020.
- Introduction: Whereas this paper concentrates on the description of the novel measurement technique and the careful characterization of related uncertainties based on in-situ measurements, the scientific results based on the retrieved leg-averaged momentum flux profile have partly been published by Gisinger et al. 2020 but are kept in this paper for completeness.

Fig. 9 is very similar to Figure 10 of Gisinger et al. (2020).

The figure was adapted and is now shown for two different wave classes. Moreover, the work by Gisinger is acknowledged making clear that the plots are just re-used for the demonstrations of the usefulness of the proposed MF-mode measurements.