Remaining criticism:

The authors have responded well to most of the criticism, and errors have been corrected, which is good. I remain truly impressed by the PTB lidar, but I still see a few potential errors in the paper.

- Concerning stating uncertainties in the PTB lidar. All papers should address the main uncertainties in the presented data; this is standard scientific good practice. Given that the authors present somewhat inconsistent results (see below), it would be interesting for the readers to learn, whether there is a possible weak point. In example, the authors write in the review answer that an error as small as 0.0034deg in the azimuth is a huge problem. What is their pointing accuracy and how does it translate to wind speed errors?

- Abstract:
  "Analysis of the corresponding cospectra showed that the CSAT3B underestimates this quantity systematically by about 3 % on average as a result of too steep a drop-off in the inertial subrange."
  The authors still claim that the difference in momentum flux between the PTB lidar and the sonic anemometer is due to the clearly seen difference in the spectra between 0.1 and 5Hz in Figure 8. However, as stated in my original review, because of the double log-axis plot, this way of illustrating error could be very misleading. In order to let the reader be able to see the frequency dependent contribution to the flux, the authors should show the premultiplied cospectrum and let the scale on the y-axis be linear. The authors have stated why they want to keep with the log-scale in Fig. 8, but this is not a strong argument. The figure is used in the context of explaining the underestimation of the momentum flux. If they want to impress with their signal quality, they should show the u-spectrum, which is a much stronger achievement, given that they measure at a near-vertical path. *In any way, they must still document that the "missing" 3% in uw flux stems from the range between 0.1 and 5Hz.* I don't believe that it does, but that rather the sum of the flux contribution in this range is much, much smaller. I believe that the difference come from the energy containing range, and if this is the case, then the H15 shows a substantial improvement of 3% in the uw comparison contrary to what the authors state now.

  *"An analysis of the cospectra shows that the H15 correction somewhat distorts the expected −7/3 power-law behavior at very high frequencies in the inertial subrange (**Fehler! Verweisquelle konnte nicht gefunden werden.**). However, the values for $f > 1\ s^{-1}$ are very small anyways and represent mostly white noise, which appears as horizontal line in spectral space. It can also be seen that the H15 correction slightly increases the cospectral energy across the entire range of frequencies. However, the too steep drop-off of the CSAT3B ensemble cospectrum is not improved effectively."*

  Of course, H15 cannot be used for low-pass filtering corrections, it was never intended for this use and it is utterly strange that the authors use this argument for not using the H15. Low-pass filtering effects due to path-averaging should be compensated for with path-averaging corrections, please see next comment.

- Low-pass filtering correction/compensation for path-averaging:

  The Moore correction was not applied on the presented spectra but only on the standard deviations and u*. Therefore, it did not affect the calculation of the spectral ratio for this study. Moreover, the implementation of the Moore correction in the TK3 software is applied to all three velocity components and the effect on the resulting standard deviations is very small as a result of the large ratio between the measurement height and the path lengths of the two instruments (>250). Pena et al. also only applied the Horst and Oncley correction to the CSAT3 data for the lower measurement height (6.4 m) and not to ones with a larger measurement height (76 m)

  The authors apply an inaccurate low-pass filtering correction erroneously, and this is not acceptable in a study where effects as small as 1% are of importance. I asked the authors to change from the Moore correction, which is an inexact approximation for path-length averaging on the vertical velocity component, to the correction published by Horst and Oncley in 2006, which is a more exact correction determined explicitly for the CSAT sonic. The answer that this is of no importance and that the authors further have applied the approximate correction for all three velocity components is not reassuring, given that Horst and Oncley (2006) showed that the three velocity components are affected differently by the path-length correction.
  The authors should just do it right! Whereas the absolute size of the correction is likely negligible for large fluxes and higher wind speeds, it can be of importance for low-wind speed stable situations, and these situations were not included in the Pena et al 2019 paper. Hence, it is not obvious how large the correction is, based only on measurement height. In Pena et al (2019), the correction was applied for 6.5m and 16m but low-wind speed situations were not presented. They stated their results with and without path-averaging correction and it was clear what the effects were. This is quite different from the current study.
  The only alternative to applying the path-length correction by Horst and Oncley (2006) is to apply no correction for both statistics and spectra.

- I noted that the second reviewer asked the authors to refrain from being subjective in their judgement of the results, and found this a very good point. Yet, the authors still state in the abstract that the agreement is “very good”. The view of “very good” can however be challenged. In example, for wind resource assessment in the field of wind energy, systematic errors of 1-2% in the mean wind speed are viewed as problematic and the CSAT sonic shows such wind direction dependent errors if one trusts the PTB lidar; around 4% for 160 deg. and -2% for 270 deg. (Fig 6). The authors should just state their results and remove the occurrences in the abstract of “very good”.
  Further, unless the data are still affected by spikes, the result for sig_u must show the same directional dependency as that of the mean velocity. And if the sig_u shows a directional dependence, one wonders about the other two velocity components...

- lines 18-19, p. 15: Comment on citation to Hogstrom and Smedman (2004). If the authors read the whole Hogstrom and Smedman paper, it should be clear to them that the statement regarding transferability from wind tunnel to atmosphere is based on a speculation and not a proven result. Whereas wind tunnels for sure have their limitations, I believe that both Hogstrom and Smedman (2004) as well as the current authors are mistaken in their strong rejection of their usefulness. Whereas the difference in drag on cylinders in laminar and turbulent flows is indeed well-known, it should be stressed that wind tunnels are not entirely laminar, but flows in wind tunnel contain a large amount of very small scale turbulence. And it is the small-scale turbulence with similar or smaller length scale than the diameter of the cylinders that matters for the drag on the cylinder. We will for sure not settle this issue here, but since the statement is strong and of very high importance for people working with wind tunnels, I would recommend to cite carefully and correctly. However, I agree with the authors that conclusions stated in the abstract of papers should be citable, so technically, the authors have a right to cite as they do.