

## Response to reviewer comment RC2

The original reviewer comments are in black font and our detailed responses use blue font.

In this paper, a field study that compares turbulence measurements from a CSAT3 sonic anemometer and a bistatic Doppler lidar is investigated. In general, measurements from the two instruments were very similar, notably the average horizontal wind velocity and the standard deviation of the vertical wind velocity ( $\langle w'w' \rangle^{1/2}$ , henceforth  $\sigma_w$ ). There was a small difference in the friction velocity  $u^*$ , with the CSAT3 being lower. When a transducer shadowing algorithm was used on the CSAT3 (i.e., referred to as H15 as described by Horst et al. (2015)), all three of these measurements increased in the CSAT3, such that there was now a small difference in the vertical wind velocity with the CSAT3 being higher. An analysis of the spectral densities of the inertial subrange showed that while theoretically the  $S_v/S_u$  and  $S_w/S_u$  ratios should all be  $4/3$ , only the  $S_v/S_u$  ratio of the Doppler lidar was near this value. When the H15 correction was applied, the  $S_w/S_u$  ratio for the CSAT3 actually decreased. The authors determine that the H15 flow-distortion correction cannot be recommended for standard applications based on the paper's results. The authors conclude that probe-induced flow distortion errors in the CSAT3 contribute little to underestimates in eddy covariance fluxes.

A big importance of this paper is that it introduces bistatic Doppler lidar measurements to sonic anemometer field studies. This is a massive step forward for micrometeorological discipline. At the same time, the results of this study appear to contradict several previous studies concerning the CSAT3. Ultimately, I do not believe this paper invalidates those studies, nor do I believe that those studies invalidate this one. Clearly, there is much more to be learned and understood about turbulence measurements, and fortunately, innovations such as the bistatic Doppler lidar help push the science forward. While I do have some major comments that should be addressed, I believe that ultimately this paper will be an extremely valuable contribution to the scientific community.

We are grateful for the insightful comment of RC2, which helped us to reconsider and reformulate some of our statements, and to provide additional information to improve clarity.

Major comments:

There are several inaccuracies in referring to Horst et al. (2015).

First, on page 2, line 17-19, it states the H15 increase in vertical fluxes was 3 to 5% due to the shadowing correction algorithm. Yet, I do not find this specific range listed in that citation. Instead, that paper refers to a range of 4-5% (these values appear in their abstract and elsewhere). There is a mention in their text (i.e. the top of page 385) that  $\sigma_w$  increased 3.5%. Is this where the lower value in the range 3-5% comes from? If so, this sentence should be revised to state the increase in  $\sigma_w$  (i.e., 3.5%) is different from the increase in the vertical fluxes (4-5%). A concern as this sentence is written, is that it has the appearance that H15 found a smaller minimum increase in the vertical fluxes than they reported (3% versus 4%), which has the effect of diminishing the importance of H15's findings.

Here are the two sentences from H15, which led us to summarize their results in this way with a range of 3 – 5%:

“Our simulations of transducer shadowing with the CSAT3 path geometry using the HATS dataset find that the attenuation of  $w_t$  equals 5 % independent of stability, and our Marshall sonic

intercomparison data suggest averaged over all wind directions an attenuation of 3–4 % averaged over 6 months of data.”

However, we agree with RC2 that it is better to quote the range given in the abstract, which is 4-5%, and we corrected this in the revised manuscript.

As an extension to this, Frank et al. (2016a) calculated the increase of vertical fluxes by applying shadowing correction to a CSAT3 for a more robust set of field sites and found this ranged between 4.5-6.8% (note, these calculations were based on the original Kaimal (1979) piecewise formulation for the shadowing correction and not the Wyngaard and Zhang (1985) sinusoidal formulation that is used in H15, which in Figure 11d in Frank et al. (2016a) is demonstrated to be  $\sim+0.6\%$  higher).

Thanks for this additional information.

Second, the description of the reference measurement used by H15, i.e., the ATI K-probe, and its correction of 1.05 for the w measurement on page 10, lines 22-26 is incorrect. While I could not deduce the exact amount of correction applied to the K-probe data in H15, it is not possible that all w-measurements are multiplied by a fixed 1.05, or even an average value of 1.05. In Frank et al. (2016b), in which Applied Technologies (i.e. ATI) were co-authors, the specific correction for the ATI K-probe is given as a function of angle of attack. In that paper the average increase in w measurements at a Wyoming field site was  $\sim 2\%$  (i.e., as demonstrated by the increase in the  $\sigma_w$  relative difference from +1% to +3% from Tables 2 and 3). By stating that the ATI-K has a fixed w correction of 1.05, instead of a variable correction that averages  $\sim 2\%$ , the reader is misled to believe that the K-probe reference in H15 is fundamentally flawed, and by extension that the findings of H15 could be fundamentally flawed.

Thanks for this important information. The effect of the ATI-K probe’s flow distortion correction was not clearly stated in H15 and we had relied on third party information. We changed the corresponding sentence in the revised manuscript accordingly:

*“H15 used an ATI K-probe sonic anemometer as reference instrument, which they assumed to be more accurate because of its orthogonal transducer array. However, the measurements by this instrument are also corrected for flow-distortion effects by a variable factor of 1.02 on average for w-measurements, and this wind-tunnel based correction factor might not be applicable in the turbulent free atmosphere.”*

Please note that even H15 state in the last paragraph of their conclusion section that it is a shortcoming of their study to use another sonic anemometer as reference:

*“The principal shortcoming of our research is the dependence of the results on a comparison between sonics ... Thus the supporting evidence for our proposed correction is somewhat indirect and incomplete. We have assumed that vertical velocity measurements made with a dedicated vertical path, such as with the ATI-K sonic, are a valid reference standard for the CSAT3 measurements.”*

Moreover, in this study at hand, we followed the call to action of H15 at the end of their conclusion section:

*“The ideal evidence for our proposal would be a comparison of sonic anemometer measurements to a reference that is free from flow distortion. One promising technique is that employed by Dellwik et al. (2015), who made simultaneous velocity measurements with a CSAT3 and with a three-component Doppler lidar system“*

I find it troubling that in this paper the results of Huq et al. (2017) are both confirmed (i.e., page 12, line 15-17) and also condemned (page 15, lines 15-18).

Indeed, the results of this study indicate that the [qualitative](#) finding of an azimuth dependence of the CSAT3 error is confirmed (which has also already been found by Grare et al.), but [quantitative](#) magnitude of the underestimation for  $\sigma_w$  is obviously overestimated by Huq et al. (2017). We base this assessment on the assumption that the flow-distortion free measurements of the PTB lidar in real-world turbulence are more reliable as a reference than the numerical simulations with fluctuating but not fully turbulent inflow.

I disagree with the question of validity on the Frank et al. (2016b) experiment on page 2, lines 5-7, that rotated instruments would have half the resolution which could invalidate the findings. The CSAT3 manual does specify the resolution as 0.001 m/s resolution for u and v measurements and 0.0005 m/s for w measurements (i.e., a higher resolution w-measurement). In Frank et al. (2016b) the most significant finding for the 90° rotated CSAT3 anemometers is listed in table 6, which tests that a hypothesis supporting the need for transducer shadowing would cause a -5% change in  $\sigma_v$  while there would be no change in  $\sigma_w$ . The observations of a -11% change in  $\sigma_v$  and 0% change in  $\sigma_w$  were somewhat consistent with this hypothesis. One interpretation of these results regarding measurement resolution is that the important observation that that  $\sigma_v$  decreased with the 90° rotated CSAT3 anemometers was conducted with the original w-measurement path which has the higher resolution. While the authors of Frank et al. (2016b) have received criticism for their experimental design, they are unsure how issues relating to measurement resolution could invalidate their results.

[We agree with RC2 that this was not a good argument, and this statement has been removed in the revised version.](#)

The range of the results from Peña et al. (2019) appear to be misstated on page 14, line 14-16. While the values of  $F_v/F_u$  of 1.32 and 1.34 do appear in their Table 2 for the Riso and Norrekaer Enge site under CSAT3/no-correction and the value of  $F_w/F_u$  of 1.13 appears for the Riso under CSAT3/no-correction, the value they list for Norrekaer Enge site for CSAT3/no-correction is listed as 1.07 and not 1.06.

[Thanks, we corrected this typo and replaced the 1.06 by 1.07.](#)

While I admittedly am new to the concept of bistatic Doppler lidar, I believe that some caution should be used before it is accepted as an unbiased control or reference measurement. First, as illustrated by the measurement volume of 2mm in horizontal diameter versus 50 mm in vertical height, this instrument clearly treats the horizontal and vertical dimensions differently. Beyond the size of the measurement volume, I assume that there is a non-orthogonal to orthogonal conversion between the measurements along the three receiving unit axes that computes the vertical measurement differently from the horizontal measurements (i.e., similar to how the CSAT3 calculates orthogonal components as described on page 4, lines 6-8). I am also troubled by Figure 7, where the spectral for the PTB lidar w measurement is clearly differently than either the u or v in the region of the inertial subrange (i.e., it is concave down while the others are ramping up). Perhaps I am not alone in questioning the use of a nonorthogonal instrument that treats the vertical dimension differently to test another non-orthogonal instrument that treats the vertical dimension differently in order to determine if there are any errors with the vertical measurement. One improvement to help address this is to present the results of the other dimensions, i.e.,  $\sigma_u$ ,  $\sigma_w$ , etc. A second improvement that could only be achieved with a new field deployment would be to collect data with the Doppler lidar focused within the CSAT3 measurement volume as well as outside of it. I once saw Tom Horst give a talk that did this with another Doppler lidar and CSAT3 study, and I recall he

believed that there was a detectable difference when the lidar was focused within the path. Regardless, on page 2, line 30, it is stated that this study “eliminates the limitations” of previous studies that lacked an accurate standard. A more conservative statement is that this study seeks to improve on those limitations.

Thanks, we reformulated this statement in a more conservative way, saying now that this study seeks to overcome the limitations. We also considered measuring with the lidar within the measurement volume of the sonic anemometer, but we realized that measurements in the nearby undisturbed flow are what is needed to characterize an instrument, as it has also been done by Huq et al. (2017) in their numerical experiment and as it has been expressed in the last paragraph of H15.

Minor comments:

Page 2, line 31-32: It is stated that there is “uncertainty of the coordinate rotations” in previous studies that is improved upon in this study. But, on page 8, line 19-21 the double coordinate rotation is implemented in this study. Does that not mean this study is also influenced by the uncertainty of coordinate rotations?

We agree and we removed this statement. We also processed the data in natural coordinates and planar fit coordinates, and found more or less the same results. But we had to make a choice and decided to apply the double rotation method. By the way, the tilt angles are now presented as a function of wind direction in the Appendix of the revised version.

Page 5, line 13-14: It is stated that the bistatic PTB lidar is validated relative to a laser Doppler anemometer in a wind tunnel. I find this ironic since it is later stated on page 15, line 17-18 that the Huq et al. (2017) results might be exaggerated because of their relationship to the inaccuracies of wind tunnel calibrations as shown in Hogstrom and Smedman (2004).

It is absolutely no problem to validate a flow-distortion free remote-sensing instrument, such as a Doppler lidar, in a wind tunnel, since it has no wake effects that might be affected by the difference in Reynolds number between the wind tunnel (quasi-laminar) and the free atmosphere (highly turbulent).

Page 8, line 3: The word “for,” might be a typo.

Thanks, the word “for” has been removed.

Page 9, Equation 1: The “,” at the end of the equation might be a typo.

This comma introduces the following subclause starting with “where ...”. However we added an additional space between the comma and the equation to clarify that it is not part of the equation.

Page 11-12, last line/line1: The slope for u in Table 2 is actually closer to the 1:1 line than the slope for u in Table 1, so a more conservative interpretation is that the difference in u between the CSAT3 and PTB lidar does not change.

We agree that slope and intercept are similar after applying the H15 correction. Nevertheless, bias and RMSE are clearly increased. Hence, we modified this sentence in the revised version:

*“Moreover, the “corrected” mean wind velocity  $\bar{u}$  has a larger bias, 0.076 instead of 0.003 m s<sup>-1</sup>, and a larger RMSE, 0.107 instead of 0.082 m s<sup>-1</sup>, although intercept and slope are similar to before applying the H15 correction.”*

Page 12, line 1-2: While this may be the case, it is worth noting that these differences are also very small on an absolute scale.

We agree and we modified this sentence in the revised version by adding the word “slightly”:

*“as can be seen from Table 1,  $\bar{u}$  and  $\overline{w'w'^{1/2}}$  show slightly larger differences from the PTB lidar after applying the H15 correction”*

Page 12, lines 3-6: It is interesting that the 0.041 increase in the slope of  $\sigma_w$  is interpreted as “systematically too large” while the 0.034 increase of slope in  $u^*$  is determined to improve “slightly”. I would recommend a choice of words to emphasize that the increases in both slopes were fairly similar in size.

We agree, that these formulations might lead to a misunderstanding with respect to the effect of the H15 correction and we rephrased the corresponding sentences:

*“H15 reported that  $(w'w')^{1/2}$  is increased by 4-5 % though this correction. Our results are on the lower end of this range, as the regression slope is increased from 0.989 to 1.030 (Table 2). However, the slope is now clearly larger than unity and the regression intercept for  $(w'w')^{1/2}$  slightly more negative, so that the comparability is almost identical before and after the correction. The agreement of the  $u^*$  values improves slightly after applying the H15 correction, since the regression slope increases from 0.0973 to 1.007 and the correlation coefficient is marginally closer to unity than before (Table 1).”*

Page 14, Figure 7: I don’t understand specifically what the last sentence in the caption is describing in the figure.

The spectra are multiplied with  $f^{5/3}$ , and this factor is increasing rapidly towards higher frequencies. Therefore, deviations from the expected flat behavior appear larger at high frequencies than at low frequencies in the inertial subrange. We changed this sentence slightly in the revised version for clarification:

*“Note that the deviations from the expected behavior in the inertial subrange appear larger than at lower frequencies due to the premultiplication.”*

Page 14, line 14-16: I find it interesting that in a relative sense, the value of 1.26 is not that different from 1.32-1.34 while 1.16 is not that different from 1.13 and 1.06. But, in Peña et al. (2019), the difference between 1.32-1.34 and 1.07-1.13 was deemed to be evidence that there were flow distortion issues with the CSAT3 but here the difference between 1.26 and 1.16 is deemed to be evidence that there are minimal flow distortion issues with the CSAT3. It is also worth noting in Peña et al. (2019) that they present results that have the H15 correction without the path-averaging correction, but not results that have the path-averaging correction but without the H15 correction. In the case of the former, the  $F_w/F_u$  ratio actually decreases by 0.039 when the path averaging correction is applied. While I do appreciate this type of analysis, perhaps this all demonstrates that it is somewhat troublesome to interpret.

We would agree that the spectral ratios analysis alone makes it difficult to assess whether a sensor is affected by flow distortion or not, or whether a correction, either for path averaging or flow distortion effect, really improves the accuracy of the measurements. In our study, we have a flow-distortion free reference instrument, which measures in an even smaller volume than the sonic anemometer and at the same temporal resolution. Moreover, it can be traced back to SI standards. No measurement device is ideal, but because of these superior characteristics makes the PTB lidar a

very good reference instrument, and because of the good agreement with this reference instrument, we conclude that flow distortion effects of the CSAT3B are not as severe as expected and that the H15 correction does not effectively correct for the remaining flow distortion effects. Another problem is that even the flow-distortion free lidar data are not fully in agreement with the theoretical value of  $4/3$  for  $\sigma_w$ . We have slightly rephrased this statement in the revised version for clarification:

Hence, we suspect that this theoretical value was probably not fulfilled in reality for the ensemble spectrum, presumably because the turbulence was not quite isotropic under all atmospheric conditions during the measurement period:

*“However, these flow-distortion free data do not reach the theoretical value of  $4/3$ , neither for  $S_v/S_u$  and even less for  $S_w/S_u$ . Hence, we suspect that this theoretical value was probably not fulfilled in reality for the ensemble spectrum, presumably because the turbulence was not quite isotropic under all atmospheric conditions during the measurement period.”*

Page 15, Line 13-15: This is an incorrect statement. The main field studies of Horst et al. (2015) and Frank et al. (2016b) involved 5 simultaneously measured anemometers. If this statement is referring to the number of sonic anemometers that are simultaneously compared to each other, then the Bayesian statistical analysis in Frank et al. (2016b) simultaneously compares 13.

The number of sonic anemometers is not really relevant here, but we agree that it was of course more than two, and we rephrased this sentence accordingly:

*“However, these previous field intercomparisons only compared different sonic anemometers with each other, partially with different sensor geometries, but none of them can be considered as flow-distortion free as the bistatic Doppler lidar.”*

Page 17, line 11-12: I am not sure this is a good statement to end on, considering the spectral plot in Figure 7 shows strange behavior in the PTB lidar in the inertial subrange and the  $1.20 S_w/S_u$  ratio in Table 3 falls short of the theoretical  $1.33$  value.

If the PTB lidar does not fulfill the theoretical  $1.33$  value, this can have theoretically three reasons.

- a) Flow distortion, which can be ruled out because this is a remote sensing instrument
- b) path averaging, which is expected to be small due to the very small measurement volume, and which was additionally ruled out to be significant by the empirical determination of the cut-off frequency.
- c) The theory of isotropic turbulence does not fully apply to all 30-min intervals of this intercomparison experiment.

Since explanation a) and b) are ruled out, we believe that the deviations from the theoretical value of  $4/3$  are real. Note, that the deviations from the expected spectral behavior are enlarged in the inertial subrange in Figure 7 due to the premultiplication with  $f^{5/3}$ , as explained above. We modified this last paragraph slightly, just to express more precisely what we intend to say:

*“In summary, the agreement of all variables tested in this comparison experiment is at least as good as or better than that between two adjacent sonic anemometers (Mauder and Zeeman, 2018). This indicates that both instruments are very precise devices for measuring turbulence statistics, particularly for vertical scalar fluxes. Considering the findings of the intercomparison experiment of Mauder and Zeeman (2018), we conclude that the other sonic anemometers tested in that study are also suitable for general flux measurements within the range of comparability and bias described in*

*that study. However, our spectral analysis shows that the bistatic Doppler lidar developed by PTB is slightly more accurate, particularly for measurements of friction velocity or the momentum flux.*“

-John Frank

## References

- Frank, J.M., Massman, W.J. and Ewers, B.E., 2016a. A Bayesian model to correct underestimated 3-D wind speeds from sonic anemometers increases turbulent components of the surface energy balance. *Atmos. Meas. Tech.*, 9(12): 5933-5953.
- Frank, J.M., Massman, W.J., Swiatek, E., Zimmerman, H.A. and Ewers, B.E., 2016b. All sonic anemometers need to correct for transducer and structural shadowing in their velocity measurements. *Journal of Atmospheric and Oceanic Technology*, 33: 149-167.
- Hogstrom, U. and Smedman, A.S., 2004. Accuracy of sonic anemometers: Laminar wind-tunnel calibrations compared to atmospheric in situ calibrations against a reference instrument. *Boundary-Layer Meteorology*, 111(1): 33-54.
- Horst, T., Semmer, S. and Maclean, G., 2015. Correction of a non-orthogonal, three-component sonic anemometer for flow distortion by transducer shadowing. *Boundary-Layer Meteorology*, 155(3): 371-395.
- Huq, S., De Roo, F., Foken, T. and Mauder, M., 2017. Evaluation of probe-induced flow distortion of Campbell CSAT3 sonic anemometers by numerical simulation. *Boundary-Layer Meteorology*, 165(1): 9-28.
- Kaimal, J.C., 1979. Sonic anemometer measurement of atmospheric turbulence, *Proceedings of the Dynamic Flow Conference 1978 on Dynamic Measurements in Unsteady Flows. Proceedings of the Dynamic Flow Conference 1978, Skovlunde, Denmark, Skovlunde, Denmark*, pp. 551-565.
- Peña, A., Dellwik, E. and Mann, J., 2019. A method to assess the accuracy of sonic anemometer measurements. *Atmos. Meas. Tech.*, 12(1): 237-252.
- Wyngaard, J.C. and Zhang, S.-F., 1985. Transducer-shadow effects on turbulence spectra measured by sonic anemometers. *Journal of Atmospheric and Oceanic Technology*, 2(4): 548-558.