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 (<w'w'>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 Sv/Su and Sw/Su ratios should all be 4/3, only the Sv/Su ratio of the Doppler lidar was near this value. When the H15 correction was applied, the Sw/Su 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.

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. 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 ~+0.6% higher). 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 ~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 ~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.

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).

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-measurment). 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.

The range of the results from Peña et al. (2019) appear to be misstated on page 14, line 14-16. While the values of Fv/Fu 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 Fw/Fu 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.

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 non-orthogonal 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.

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?
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).
Page 8, line 3: The word "for," might be a typo.
Page 9, Equation 1: The "," at the end of the equation might be a typo.
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.
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.
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.
Page 14, Figure 7: I don't understand specifically what the last sentence in the caption is describing in the figure.
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 Fw/Fu 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.
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.
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 Sw/Su ratio in Table 3 falls short of the theoretical 1.33 value.

-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.