**Final author response**

Journal: AMT
Title: Field intercomparison of prevailing sonic anemometers
Author(s): Matthias Mauder and Matthias J. Zeeman MS No.: amt-2017-284 MS Type: Research article

We are grateful for the very valuable and constructive comments by the reviewer. The comments are in black font, responses are presented in blue font and changes in the manuscript in red font.

**Response to RC1 (by John Frank)**

John Frank, USDA Forest Service In this manuscript seven sonic anemometers from six non-orthogonal designs and four manufactures are compared during a two and a half-week study at the TERENO/ICOS site in southern Germany. Half-hourly mean wind velocity, mean and standard deviation of temperature, standard deviation of vertical wind velocity, friction velocity, and buoyancy flux were compared between the seven anemometers. In general, all anemometers were reasonably similar, with the largest discrepancy being temperature measurements with the Gill anemometers.

The topic of this study is timely, with a growing interest in the accuracy in sonic anemometer measurements. The work presented here is clear, convincing, and thorough. I believe there is one main comment and a few minor issues that need to be addressed before it is acceptable for publication in Atmospheric Measurement Techniques.

I have one main comment that should be addressed. The discussion states "it is also possible that the flow distortion errors were very small for our experimental set-up because the angles of attack are close to being surface-parallel" with support from Figure 9 showing that most angles were within ±6° with a majority falling within an even narrower window. I am concerned that there is a high likelihood that each of these non-orthogonal anemometers have erroneous and unpredictable measurements of σw for winds with very small angles-of-attack. This is described in Appendix 2 of Frank et al. (2016a) where it is demonstrated for the CSAT3 that as the angle of-attack approaches 0° (i.e., near surface-parallel) that systematic w measurement uncertainties approach ±∞. This finding could be extended to any of the non-orthogonal anemometers included in this study, and I have included an Appendix at the end in this review to expand upon this topic. The discussion also states "The other interpretation that all anemometers are afflicted with the same bias appears less likely, since it is difficult to imagine that several instruments measure the same quantity equally wrong, despite the obvious differences in sensor geometry and internal data processing."

The figure and the text passages that this reviewer refers to are not part of the current manuscript published in AMTD. Based on the reviewer's comments in the Quick Review, we have double-checked our results and we found that there was a mistake in the calculation of the angles of attack. The spread in angles of attack during our experiment was actually much larger as previously thought, i.e. a standard deviation of 15° rather than ±6°. This standard deviation of 15° is at the upper end of values reported for previous intercomparison experiments. Therefore, we have also changed the line of arguments in the discussion in accordance with the new results. We now really believe that the different instruments all show the same biases despite their differences in geometry and internal corrections. Since there is strong evidence for a bias in $\sigma_w$ of the CSAT3 from other studies (Horst et al. 2015, Frank et al. 2016, Huq et al. 2017), this seems to be the only logical conclusion.

Nevertheless, we highly appreciate the reviewer's extensive comments on the problem of non-orthogonal anemometers, which he provided in the Appendix of RC1.

I suggest that "the other interpretation" might actually be true, that these systematic w measurement uncertainties that approach ±∞ could explain why these "instruments measure the same quantity equally wrong".

We now fully agree that these instruments measure the same quantity equally wrong. We have informed the editor about the changes that we have made after the Quick Review stage, but unfortunately, those notes did not reach the reviewer in time.

A key component of this experiment is the inclusion of the CSAT3, which does not apply any transducer shadowing correction, which has been modeled (Huq et al. 2017) and observed (Horst et al. 2015) to have transducer shadowing errors that when transformed into orthogonal coordinates lead to unpredictable measurements for near surfaceparallel winds (Frank et al. 2016a). It is nearly impossible to know exactly what happens for near surface-parallel winds, but a simple evaluation of the Kaimal correction (Kaimal et al. 1990) applied to the CSAT3 yields a range of ± ~4° where systematic measurement errors approach ±∞ as shown in Figure 3f of Frank et al. (2016a). Because most angles were within ±6° in this study, I consider it extremely likely that the CSAT3 data has such errors. At the same time, all other anemometers in this study are probably more susceptible to this problem because their transducers are all tilted closer to the horizontal plane (45° for the Gill, R. M. Young and Metek versus 60° for the CSAT3).

I have two suggestions that can help address this issue. First, the authors could do a sensitivity analysis to quantify the potential impact of transducer shadowing on the CSAT3 σw and buoyancy flux measurements. I would suggest using both the piecewise (Kaimal et al. 1990) and sinusoidal (Wyngaard and Zhang 1985) corrections presented in the following appendix. While it is important to note that there is no consensus that these corrections are accurate (two studies have shown that they might account for about half of the shadowing (Frank et al. 2016b, Huq et al. 2017)), this will help to evaluate the statement that "the flow distortion errors were very small".

Again, this is a quotation from an earlier version of the manuscript, which is not part of the manuscript that is under review of this open discussion and published in AMTD. In contrast to the earlier version, we do not believe anymore that the flow distortion errors were very small, but rather that all tested anemometers are afflicted with a very similar flow distortion error. Therefore, we have even further strengthened the following statement accordingly:

A common significant systematic error of all tested instruments is quite possible, as suggested by Frank et al. (2016).

However, we find that this suggestion is extremely interesting for follow-up work, but we believe this is out of the scope of this manuscript. We acknowledge the need for an angle of attack dependent correction, but this should probably include data from multiple sites with different surface and vegetation properties.

Second, in conjunction with the histogram in Figure 9, an analysis of the relative contribution of winds from each angle of attack bin to the total σw and buoyancy flux measurements would be useful. Figure 9 currently shows that most data is in the bins -1° to 0° and 0° to 1°. How important are the measurements in these bins to the total σw and buoyancy flux measurements reported in figures

5 and 8? What is the contribution of winds that exist within -4° to 4° (i.e., a range over which the Kaimal correction as applied to the CSAT3 could conceivably result in unpredictable measurements)?

I very much look forward to the authors' critical evaluation of this topic as I believe it will be an immense benefit to the research community to thoroughly discuss non-orthogonal wind measurements.

We belief this is a misunderstanding, which stems from the fact the reviewer refers to an earlier version, which was altered during the Quick Review stage. Since the old Figure 9 was erroneous and is not included in the manuscript under discussion, we do not believe there is a need for further clarification here.

I have a minor comment about interpreting results based on offset/bias versus slope differences. My impression is that the authors focus more on offset/bias differences and less about slope. One example is on page 12, lines 4-5 "the fluctuations of sonic temperature agree much better". When I compared Tables 3 to 5, my attention immediately focused on the slopes, which are not much different. The average absolute difference from 1.00 (i.e., an extremely simple metric to summarize the group differences) was 4.0% for Table 3 (i.e., 1.05, 0.97, 1.01, 1.05, 1.06, 1.04 -> +5%, -3%, +1%, +5%, +6%, +4% -> (5+3+1+5+6+4)/6 = 4%) and 3.5% for Table 5. Similarly, on page 12, lines 12-14 it is commented the high slope of 1.06 "might be a direct consequence of the almost equally high regression slope of 1.05". This is not surprising, because ideally, slope errors in measuring Ts should commute to slope errors in σTs.

We agree that for flux measurements, an error in the slope is more severe that an error in the bias, since mean is always subtracted when calculating a covariance. The first sentence quoted in this comment refers to the Gill instruments. We think then this is a fair statement, because they have quite large deviations of the mean sonic temperatures, but its standard deviation appears comparable. We have modified this sentence for clarification:

Despite the large discrepancies of the mean sonic temperature measurements of the Gill instruments, the fluctuations of sonic temperature agree much better

We have also made the second quoted sentence stronger, because there is indeed a direct relation between the slope of the mean and the slope of the standard deviation:

The METEK.uSonic.omni stands out because it has the highest regression slope of 1.06, which is a direct consequence of the almost equally high regression slope of 1.05 for the mean sonic temperature measurement.

Specific comments:

The introduction is excellent, and one of the better that I've read for sonic anemometer studies. What is the sampling rate of the sonic anemometers?

This information has been added to the manuscript.

The sampling rate was 20 Hz, except for the CSAT3_2, which was sampled at 60 Hz, and the Gill_HS, which was sampled at 10 Hz.

Figure 2: It would be good to mention either on the figure or in the caption that the model of anemometer corresponds to the same names listed in Table 1.

We agree and we have added that information in the caption of Figure 2.

they are presented from left to right in the same order as they are listed in **Fehler! Verweisquelle konnte nicht gefunden werden.**.

Page 7, line 15: Are the obstructed wind directions based on 30-minute mean direction or some other metric?

They are indeed based on 30-minute mean wind directions. We have added this information.

These quantities were filtered for rain (during the respective half hour or the half hour before as recorded by a Vaisala WXT520 sensor of the nearby TERENO station DEFen), obstructed wind directions $\varphi$ based on 30-minute averages (70° < $\varphi$ < 110°; 250° < $\varphi$ < 290°) and non-steady-state conditions, …

2.2 Data processing: The software R should be cited, with details provided by the R function "citation()".

This citation of the R software has been added.

Table 2 and Figure 3: There is a slight difference in nomenclature between "mean wind velocity" and "total wind velocity". Based on the 2-D rotation, these should be the same, but consistent labels would be good.

We agree and we have now consistently used the term "mean total wind velocity".

Table 2 and 4. The slope difference between the CSAT3 anemometers is actually quite large (1.02 versus 0.97 for U and 0.98 versus 1.03 for $\sigma w$). Do the authors believe that this is due more to the repeatability of measurements using the same anemometer design or to the 54 m separation distance?

We cannot really answer this question based on our results. Based on the experience from other intercomparison experiments, differences are on the order of 2 - 3% are certainly possible between the different locations but can also be possible due to instrumental differences. Moreover, we would like to stress that the biases and RMSDs are really small despite the slightly larger slopes.

Tables 2-7 and Figures 3-8: I could imagine all information from these tables could be moved to the blank space of the corresponding figures, thus saving space in the manuscript. I would also mention that something like R-squared values might be very useful to include as well.

The final layout will be done be the publishing company and we will then check the galley-proofs to avoid blank space. We decided against using R-squared values because we find this metric sort of redundant when RMSD and bias values are already provided. These two metrics also provide more specific information as to what degree the differences are systematic or random, while R-squared lumps both differences together.

Tables 3, 5, and 6: What was the metric used to identify "unusually large deviations from the etalon"? This should be mentioned in the methodology.

We agree that these criteria should be clarified and we added this information in the respective table captions:

Table 3: (slopes deviating more than 5% from unity and absolute differences of more than 1 K)

Table 5: (slopes deviating more than 5% from unity and absolute differences larger than 0.05 K)

Table 6: (slopes deviating more than 5% from unity)

Page 12, Line 14-15: What does "agreement between the two CSAT3 is except for a few outliers" mean?

Thanks, two words were missing here:

The agreement between the two CSAT3 is very good except for a few outliers

Page 13, line 4: "Yount" should be "Young".

This has been corrected.

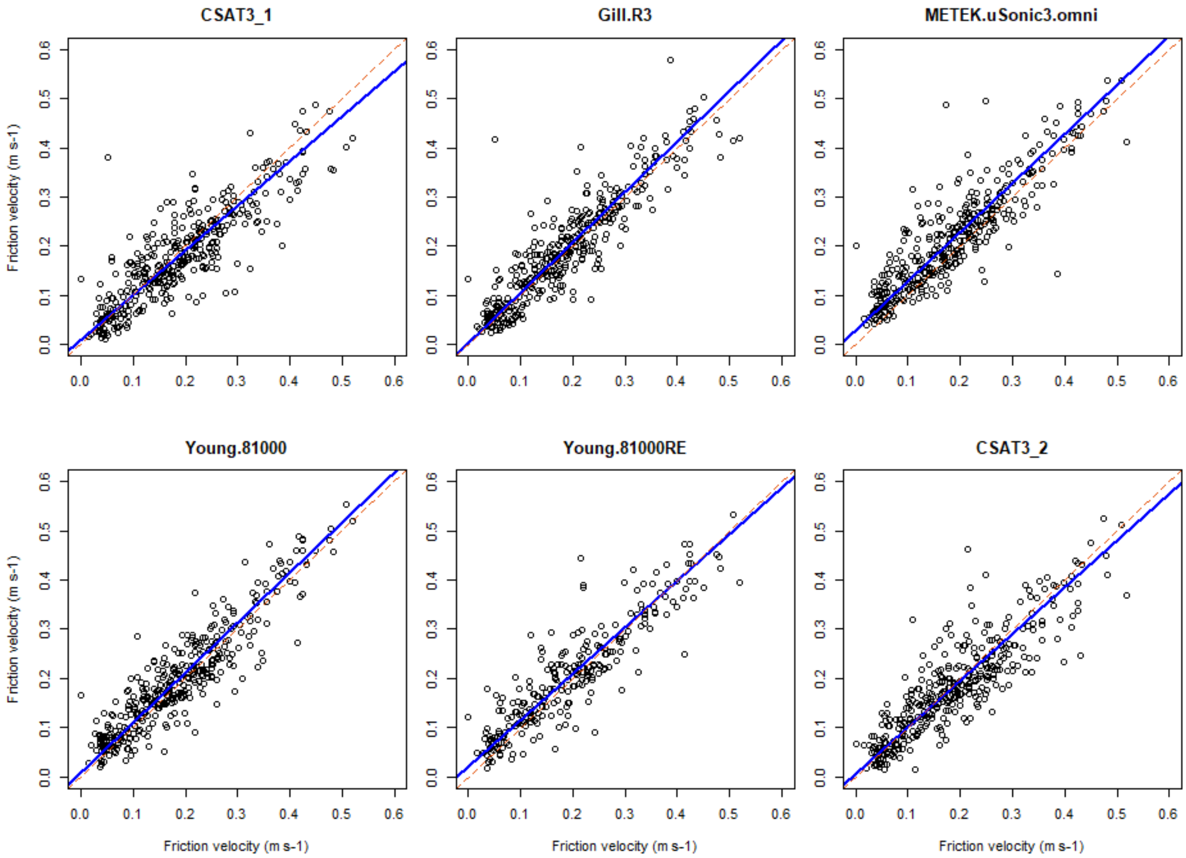
Page 13, line 11-12: It could be noted that CSAT3_2 is the second lowest, so both CSAT3 are fairly low.

We agree and we have added the following sentence:

Similarly, the CSAT_2 shows the second lowest regression slope, but its bias and RMSD is very similar to the other instruments.

Figure 7: One of the ylabels are missing.

Thanks, we have recompiled this figure including the ylabel:




Page 16, line 6: “error of due” should be “error due”.

Thanks, we have removed the word “of”

Figure 9: Are the magnitudes of the numbers of y-axis correct? The number of occurrences seems extremely low for instantaneous 10 Hz or 20 Hz data. Or is this half-hour average angle of-attack? The number of occurrences seems similar to the number of half-hours. If this is the case, then all text that refers to “small … angles of attack” (beginning with page 16, lines 25-26) must be revised to reflect the instantaneous angles experienced by the anemometers.

Indeed, the values were too low due to a mistake in the data analysis routine. We have therefore removed this figure already after the Quick Review because this whole line of argument saying that we had exceptionally low flow angles and therefore low systematic errors does not hold anymore. For completeness, we now provide the standard deviation of the flow angle, which is 15°, ranging at the upper end of values reported in the literature. Thanks to the reviewer’s remark in the Quick Review, we were already able to correct this mistake.

Page 18, line 11-14: The discussion of “type A” and “type B” should occur earlier in the paper than the conclusion.

We have added the following sentence to the discussion section:

Now, all tested instruments are within the limits that Mauder et al. (2006) classified as type A, i.e. sonic anemometers suitable for fundamental turbulence research.