Review of “Calibrating low-cost rain gauge sensors for their applications in IoT infrastructure to densify environmental monitoring network” by Krüger et al

Review by Rolf Hut

The authors calibrated a collection of off the shelf low-cost rain gauges to test if they are usable in scientific applications with the factory calibration. Given the amount of projects that aim to use Personal Weather Stations (PWS) to supplement professional networks, this is a valuable addition to the literature. I do have, however, some suggestions to in my opinion improve the paper (and its usability by the scientific community) before publication.

Thanks for time and effort to review our manuscript. We have replied inline in the text. Author comments are in black, reviewer comments are grayed out.

Overall comments

The manuscript as written hinges on two thoughts: on the one side a lab and field calibration of low-cost rain gauges and on the other side an overview of IoT hardware and cost needed to use low cost sensors in general and rain gauges in particular. This last part (IoT hardware) is worked out in far less detail compared to the first part (rain gauge calibration). In literature a large collection of articles reviewing state of the art development boards (including Arduino and Raspberry Pi) for use in environmental sensing in general and weather stations, is available. I would suggest that the authors focus on the rain gauge calibration and remove, or move ta an appendix, paragraphs 2.1 until and including 2.2.4. In the main text the authors can cite relevant literature on IoT hardware reviews. (a quick search on google scholar already resulted in these DOIs, there is much more: https://doi.org/10.3390/ijerph17113995, https://doi.org/10.1016/j.cosrev.2021.100364, https://doi.org/10.1016/j.procs.2014.07.059, 9734/AJRCOS/2021/v9i130215)

We agree, that the paragraph on the used IoT hardware is relatively short. Nevertheless, although many studies have already been published on using IoT capable developer boards, we wanted to include the used setup. Here the aim is to increase applicability of the rain gauges related findings and a better transferability for the final user. The source code will be added to appendix.

In calibrations of rain gauges the crucial question is always: “what do we use as ‘the truth’ and the authors have three reference devices available. They choose to use the Hellman gauge as reference without further justification. I would ask the authors to substantiate why the Hellman, compared to the other devices, should be considered “reference” (or “thruth”).

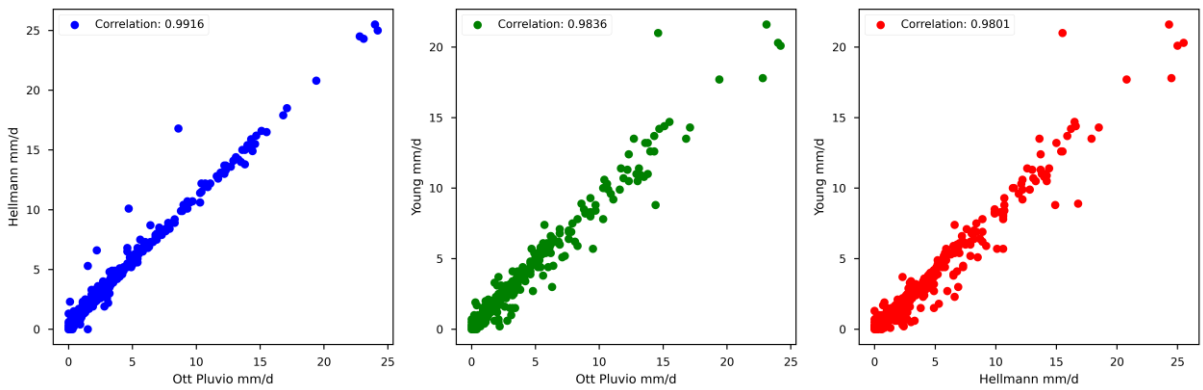
Thank you for the remark – indeed the decision here is not easy. First of all, the Hellmann device was used as reference because it’s considered as the

reference for the climatological measurements at this meteorological site since the 1950s. This fact is already stated in the manuscript, but not yet made clear as justification. Further, the used instrument principle requires no mechanical and electronic parts and thus the data quality should be stable, as the instrument is set up properly. For comparisons on timeframes shorter than 1 day, the Ott Pluvio will be used (see below). Will be clarified in the text.

Continuing on point 2: the authors only report on the difference between the different gauges, both within the groups of low-cost gauges and between the low cost gauges and the reference gauges. However, they do not quantify if these are significant in the light of uncertainties, either inherent in their way of measuring, or inherent in the nature of rainfall. Given the large amount of low-cost rain gauges they use, it should be straightforward to indicate if the values of the the reference gauges are significantly outside the distribution of low-cost gauges. If the authors have access to a long time series from the reference devices (which I assume they have), they could use triple co-location to estimate the uncertainties in the three reference devices. This would than allow for a two-way comparison between the reference devices and the low-cost gauges. There is a wealth of literature on (how to do) comparisons between rain gauges, including the statistics involved. I suggest starting at Lanza 2009 (https://doi.org/10.1016/j.atmosres.2009.06.012)

Thanks for the suggestion of the triple collocation method. We used a longer time series of the three reference gauges ranging from 2017 to 2019 consisting of daily observations to estimate the uncertainties. Inspection of the three scatterplots with all combinations of the reference gauges led to the assumption that the Ott Pluvio is the best performing as the Hellmann/Young Scatterplot had the lowest correlation (Stoffelen and Vogelzang, 2012). We then used an implementation provided by Jur Vogelzang[1] to estimate the error variances using the Ott Pluvio as reference system.

The following daily error standard deviations could be determined: std_{Pluvio}= 0.150mm/d, $std_{Hellmann}$ = 0.183mm/d, std_{Young} = 0.278mm/d



[1] https://github.com/knmiscat/triple_collocation

We then used a Monte Carlo simulation utilizing the daily datasets of the field campaign and the daily uncertainties of the rain gauges to generate a distribution of artificial datasets for each gauge. These have then been compared with the distribution of low-cost gauges utilizing a t-test, resulting in a rejection of the null hypothesis for all reference gauges. Thus, all references gauges are significantly outside of the distribution of low-cost gauges. These steps will be added to the manuscript.

In the lab calibration it is extremely important that the rain gauges are placed perfectly horizontal. I assume the authors made sure of this. I would suggest to add a few sentences on how this was done. Furthermore, it is important to know if all rain gauges were oriented exactly the same direction on the table. If the table was even slightly tilting, having all rain gauges in the same orientation would result in a bias towards a certain direction and could explain the left-right difference observed?

In preparation of the lab calibration the table was levelled utilizing the adjusting screws in the table legs. The rain gauges themselves have been levelled using the built-in bubble level. Further, slices of paper have been used to account for remaining unevenness on the table.

During the calibration, all gauges have been oriented in the exact same direction

More detailed explanation will be added to the manuscript as suggested.

The analyses done within the discussion is, in my point, central to the manuscript. I would suggest to move the results of the comparison of the field and lab experiment to the result, explain in the methods which (statistical) methods are used to compare the two datasets and in the discussion only reflect on the result, not present new ones.

Thank you for critically pointing that out. Will be restructured.

Specific comments

All figures need more detail in their captions to understand what is shown.

Will be improved.

In figure 1 I would add a vertical (red?) line at 0.20 mm to indicate where the factory calibration is.

Will be added.

Figure 6 could use the Hellman data as crosses or points. Especially on the one hourly data it is interesting to look at the uncertainties of the three reference devices (see above).

Will be added.

Overall I think this is a highly relevant paper given the focus on citizen science projects to use Personal Weather Stations to supplement professional networks. With the above suggestions implemented I would be happy to recommend publication in GI.

Rolf Hut

References:

Stoffelen, A. and Vogelzang, J.: Triple collocation, https://doi.org/10.13140/RG.2.2.30926.66888, 2012.