

The revised manuscript by Bau et al. deals with an interlaboratory comparison of optical particle size spectrometers. For this study, an impressive number of partners and tools were made available. The information provided makes the change to the heading worthwhile. Overall, the manuscript is well written, although there are still a few questions that need answering.

ABSTRACT

L33 “... instead of establishing or developing a calibration procedure...” - Am I right in thinking that this means a method is possible even without calibration, one that allows for comparability? Or that calibration isn't necessary? It could be interpreted either way.

MAIN

L101 - Is this a sort of reliability test, then? Would a service or a calibration be in order now?

L113 - Can you explain to what extent this constitutes a reference dataset, given that no calibration data for the instruments was available (factors for ambient refractive index and more)?

L120f - You can generate a summary OPSS number-size graph for each partner measurement, showing the OPSS comparison unit (that one that stayed always the same device) measurement for each comparison at that location, so that you can assess any fluctuations between measurements.

L168 “monodisperse” -In principle, yes, but in the later graphs larger particles can also be seen, which are likely agglomerates; in terms of total number, they probably account for as much as the main mode. Please also note that the atomisation method likely produces many agglomerates

L169 “optical properties...” – which one? Refractive index?

L187f “good laboratory practice” - Please explain what this means in this context. (Perhaps the length of the tubes, no right angles, etc.?)

L212 – “total number concentration” - Some of the instruments exhibit higher error rates the closer they get to their maximum count. Depending on the algorithm, only the large particles may then be counted, or larger bins may be counted even though none exist. In addition, miscounts can occur at low bins if very large particles (roughly three times the wavelength) are present in the system. Has this been accounted for in the evaluations?

Figure 4 - Taking the Grimm 1.108 as an example, the error appears to be highest here; however, the test aerosol is also relatively small and at the lower limit of the smallest size channel, which should be noted

Fig. 5 and 6 - Depending on the manufacturer, it is apparent that instrument types cluster closely in the comparison of GSD and mode; this could be due to the calibration that manufacturers use for the respective instrument; at Palas, the different devices are built and calibrated at different locations or procedures, or isn't it?

Fig.7 -Without or using a different control instrument, this would look very different indeed; what does that imply?

CONCLUSION

L431- "need of strengthening robustness in..." Please clarify what you might still missing in regard of ISO standard 21501 that is meant for the standardization for OPSS measurements

Please answer following questions:

How could this study be replicated without having the control instrument, since every measurement and provided data is relative to this one, or how does the dataset help if no absolute figures are given regarding number concentration density?