We thank C. H. Twohy for his detailed review and valuable comments. The manuscript has been modified according to the suggestions proposed by the reviewer. The remainder is devoted to the specific response item-by-item on the reviewer's comments.
Notice that to match with the version, the CDPs were renamed; the ex-CDP1 is now CDP2.

RC=Reviewer Comments
AR=Author response
TC=Text Changes

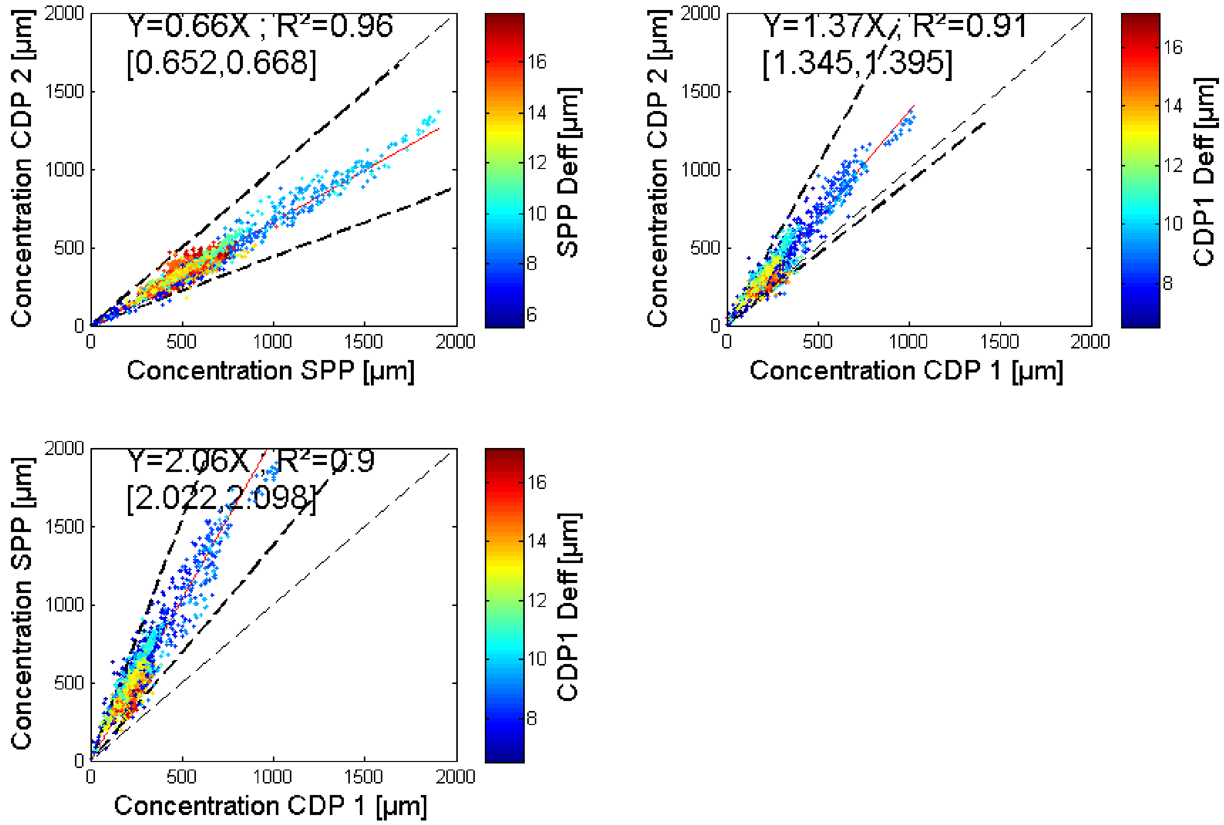
RC: General Comments:
The authors have presented an ambitious campaign to compare a number of instruments, some of them fairly new, for measuring cloud properties in-situ. Since older intercomparison studies focused on older instruments, this is worth publishing and suitable for AMT if the manuscript can be improved. I am somewhat disheartened that the results don't show better agreement between the instruments and that we are still resorting to simply normalizing results to a single instrument to produce a unified data set. If the paper is better organized and more clearly assesses uncertainty, as suggested below, hopefully these results will lead to a better path forward.

Some important results include that rapid changes in airspeed do not affect some instruments (if the speed changes are accounted for), and that other instruments are strongly dependent on alignment. Conditions affecting other instruments–for example, mean inlet speeds (p. 22) and splashing (p. 16) are likely manifested differently with droplet size. However, limited comparisons based on droplet size are made (except for the alignment and SPP/FSSP comparison studies). It would be useful if this were done for the dataset as a whole. Likewise, droplet number concentration may affect the results due to coincidence, and some stratifying by number might also yield interesting results.

The structure of paper could be better organized and tightened up. There is some exact duplication of text, as noted below. In addition, some of the effects of operating conditions–low airspeed, location, etc, are discussed in different sections and perhaps could be better combined in one place. The authors should also consider whether other existing discussion, tables and plots can be condensed, especially since additional information will likely be added to satisfy the referees.

It is generally unclear which results are new in this manuscript and which just corroborate what others have already published. A better reporting of past intercomparison studies is needed in the Introduction, of published uncertainties in the Methods section, and of related results/characterizations in the Results section.

AR: The manuscript was reorganized according to reviewers' comments.
Section 2.2 was almost totally reorganized.
The size dependence is observed only for the FSSP and is investigate in the appendix.
Line 737: As the FSSP is aspirated with no flow straightener in front of it, turbulent flow and distortion of the size distributions can be expected. Anisokinetic sampling and errors in the sampling volume can explained the concentration overestimation. For the other instruments, the biases were constant during the campaign and independent of the wind speed and the droplet size (not show).
Here is an example with the wind tunnel instruments:



The coincidence phenomenon is not considered:
Line 849: Moreover, as the extinction comparisons of the SPCs with the PWD provide correlated linear regression, without saturation, the coincidence phenomenon was assumed negligible.

RC: Specific Comments:

Line 104: Given the uncertainties that still exist, perhaps "to obtain precise information" should be changed to "to attempt to obtain precise information".

AR: We agree, has been revised.

RC: Page 6: I think a sentence or two are needed before going forward to Section 2. The authors should explain that the Section 3 results show the differences between the measurements, and Section 4 tries to explain the reasons for differences that are outside the normal uncertainty windows cited. Otherwise, the Results section seems inadequately explained when one is reading it, when actually some answers are given in experiments described in the subsequent section. Some of my comments/questions on the Results section reflect this.
AR: A short summary of the plan has been added:
Line 195: Section 2 of this work presents the measurement site and the instrumentation used during the campaign. Section 3 addresses the comparison of the data recorded with the ground-based and the wind-tunnel instruments. Then, the proposed method to correct and standardize measurements is outlined. Main causes of potential biases and effects of the wind direction and speed are discussed at the end of the section. Section 4 is devoted to the conclusions.

RC: Instrumentation Section: Pre and post-calibration and maintenance procedures were not discussed. Please summarize.
AR: The date of the calibration for each instrument has been added to the manuscript.

RC: Lines 202-203: From which probe was this approximate range?

Line 240: The wind parameters were measured with a Vaisala sonic anemometer and a vane anemometer. Typically the weather conditions were dominated by westerly winds with speeds ranging from 1 to 22 m s^{-1}.

RC: Line 227 and throughout: "Accuracy" seems to be used interchangeably with "uncertainty" and "error"; for most of the statements, "uncertainty" would seem to be the most correct term.

AR: We agree that "uncertainty" is the most correct term. We employed the terms "accuracy" and "errors" as synonyms of "uncertainty" in order to avoid being overly repetitive.

RC: Pages 8-9 and page 12: Were the old-style or new aerodynamic probe tips used for the FSSP and SPP?

AR: The corresponding information has been added to the manuscript:
Line 350: It should be mentioned that no conical attachment (horn) was mounted on the instrument during this campaign. It means that the air suction into the FSSP inlet tube can generate curved streamlines leading to potential inertial concentration effect (Gerber et al., 1999).

RC: Lines 235-236: "Theoretical air speed" is not explained; this discussion should be combined with lines 314-315. It's not clear why the speed in the FSSP is 15 ms-1 and in the pump is 15 ms-1. Is this based on an area ratio between the pump and the FSSP? How much uncertainty is added due to this speed measurement? Update: finally found this info on p. 27. Please move it up to where it's first discussed.

AR: This explanation has been moved to the instrumentation section:
Line 345: The flow through the pump was monitored with a mass flow anemometer. The air speed was set to around 15 m s^{-1}. Theoretically, this flow leads to an air speed through the FSSP-100 inlet of 9 m s^{-1}; that value was employed for the data processing.

RC: Lines 317-327: Duplicate material from above.

RC: Line 370: "PVM1" has not been defined–apparently there were two PVMs operating at different sample rates? Please explain.

AR: There were 2 identical PVMs, arbitrarily named 1 and 2, but the PVM2 measurements have been removed. There is no sampling rate for the PVM.

RC: Lines 363-364: "The two CDPs were installed in the wind tunnel"–duplicate material from above.

RC: Table 1: Was the irregular coverage of various instruments due to problems, or were some instruments swapped out for others in the same location?
AR: This is due to problems with the instruments (absurd values, interference, electric problems …), there was no swap, only the SPP was moved on the roof but after the intercomparison campaign.

RC: Lines 410-411: This needs further explanation–was the instrument later found to be dysfunctional for some reason, or did it simply not agree?

AR: The PVM2 showed strong discrepancies with the other instruments during the campaign, the reason(s) is unknown. We choose to remove the PVM2 measurements as it doesn't show new results.

RC: Lines 424-425: A factor of two difference between two of the same model instruments (CDPs) mounted near each other is disturbing. This points to some systematic problem, possibly with effective sample volume as suggested by Lance et al (2010), which would need to be addressed through testing and conversations with the manufacturer. In the two years since this experiment, has any resolution been made?

AR: The two CDPs have a ratio of 1.35 (line 532) and this was not the same model, a sentence has been added in the section instruments:
Line 390 : Two types of CDP were used during the campaign: the first version (CDP1) with original tips and the second version (CDP2) with Korolev tips against possible shattering effects.

RC: Lines 450-451: Is there any indication that this problem was more prominent when larger cloud particles were present? Also "slashing" should be "splashing". Sentence on lines 484-486: Duplicate information discussed above.

AR: We have no indication as we have no measurement for droplet larger than 50 µm. This is just an assumption to explain the overestimation of the SPP.

RC: Line 500: "Such discrepancy" should be quantified–it looks like about 30% to me.

AR: The sentence has been deleted because the uncertainty of 30 % concerns the LWC in the study of Burnet and Brenguier (2002).

RC: Lines 509: "comparable" doesn't seem like the right word, since actually the slopes of 0.35 to 2.6 cited suggest the instrument results are not very comparable.
Line 512: "constant" would be better as "systematic", I think.

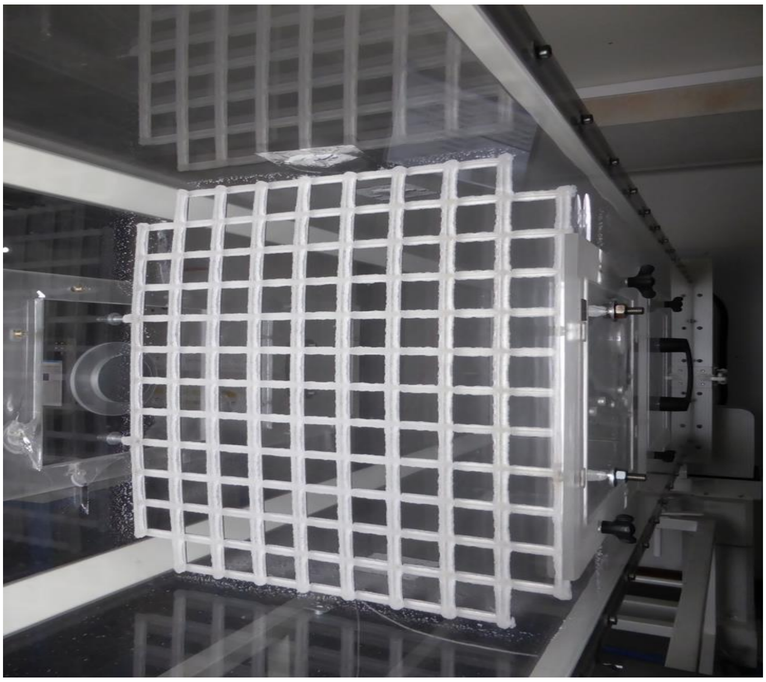
AR: We agree, has been revised.

RC: Line 513: Please define "number calibration coefficient". Given the discussion later, this is apparently a factor that would correct the number concentration data for some physics that we don't yet understand.

AR: It was a misleading sentence. There was no calibration in the total number concentration. We meant the corrections (listed in section 2.2) concerned the sampling volume and the concentration absolute values. The sentence has been deleted.

RC: Lines 547-548: I don't think this has been addressed. Why is the airflow more accurately monitored in the wind tunnel than in the ambient air? Add in Methods section.

AR: The text below has been added to Section 2.2.
Line 456: The method of an icing grid (see, e.g., Irvine et al. 2001) was used for airflow uniformity measurement. The tests were performed at the maximal airspeed available in the wind tunnel. According to the preliminary results, the variations of the thicknesses and widths of the iced bands were lower than 5 %, i.e., of the order of the uncertainty of the method (see picture below, P. Personne and C. Verhaege personal communication). Thus, we can reject the hypothesis of the airflow heterogeneity as the cause of the differences between data.



Lines 570-572: Aren't the depth of fields for some instrument types more uncertain than for others? This should be addressed to help better understand the results.

AR: A priori, there is no difference in depth of field uncertainties between the SPCs, but these uncertainties are difficult to quantify. During our campaign, there was no DOF verification according to the manufacturer's values.
Line 320: Usually, all these errors are very difficult to quantify and extreme uncertainty can be very high. For example, Burnet and Brenguier (2002) reported that the DOF of the FSSP could be significantly different from the value given by the manufacturer; this difference may reach a factor 2.

RC: Line 599-600 and lines 952-952: Really? Discrepancies of 300- 500% are expected for instruments based on similar light-scattering principles during co-axial sampling? This is not something I would expect. The authors may be overstating the acceptance of uncertainty here, even for number concentration. See suggestion for quantifying uncertainty better in Table 2, below.
Table 2: This is somewhat simplistic, as some of the uncertainties are for optimal conditions and some are for non-optimal conditions. While difficult, a detailed assessment of known uncertainties is critical in evaluating your results, and would be an important summary for the community. I suggest that you should expand the table to cover the known range of uncertainties and when to expect them (perhaps as "optimal" and "non-optimal" values, with references). For example, the PVM is less accurate when the MVDs are higher; the CDP less accurate under higher droplet concentrations; the FSSP data degrades if not axial and isokinetic, etc. Then perhaps your later statement that a 300-500% difference between the different instruments "was expected" can be better evaluated, based on the known uncertainties and actual operating conditions. Or, perhaps it will show that there are still unknown sources of uncertainty.

AR: The table 2 has been revisited; we added two columns to display the uncertainties: one for normal and one for extreme conditions, with the references.
To explain the 300 – 500 % uncertainty, we added the following sentence into the manuscript:
Line 671: The bias between the instruments results from a systematic error originating from the inaccurate assessment of sampling volume. The single particle counters (SPCs) have uncertainties in optical parameters such as the DOF and in corrections like the activity. In

addition, the data of the ground-based FM and FSSP are affected by errors of the sampling speed assessment.

RC: As another referee suggested, expected uncertainty envelopes on the correlation plots would be useful to evaluate the comparisons.

AR: We agree, the dashed lines that show the envelope of uncertainty have been added to the figures.

RC: Lines 612-615: Unclear–why would a ground-based instrument be functionally different, as long as it is aligned and has the proper airflow moving through it? Or perhaps that is the point, that this is difficult to achieve for the FSSP.

AR: Yes, FM has a horn, not the FSSP. We thus insist in the text on the fact that this creates anisokinetic sampling for the FSSP, and we add the sentence:
Line 716: On the contrary, anisokinetic sampling of the FSSP leads to higher discrepancies when this instrument is compared to other ones.

RC: Lines 632-638: If the droplet speed inside the inlets is a major uncertainty, then it might be expected that the results depend on droplet size. This should be explored.

AR: This is the subject of the section discussion, it concerns only the FSSP:
Line 737: As the FSSP is aspirated with no flow straightener in front of it, turbulent flow and distortion of the size distributions can be expected. Anisokinetic sampling and errors in the sampling volume can explained the concentration overestimation. For the other instruments, the biases were constant during the campaign and independent of the wind speed and the droplet size (not show)

RC: Lines 663-669: Is this a new result? Or does the manufacturer already warn of this limitation?

AR: This is a new experimentally justified result that concerns only the PUY wind tunnel. The corresponding paragraph has been revised.
Line 788: For cloud measurements, it is thus recommend using the PUY wind tunnel with an air speed higher than 10 m s^{-1}.

RC: Lines 768-770: Again, is this a new result or is this outside the operating regime recommended by the manufacturer?

AR: This is new results about non-operating regime recommended by the manufacturer: the FSSP should be facing the wind, but if not, we show the effects on the size distribution.

RC: Lines 779-780: Some discussion of why this is the case, with depiction of the geometry of the two instruments would be useful.

AR: We agree. Some descriptions of the FM and the FSSP have been added in the section instrument. The most important point is the presence or not of a horn used to avoid turbulent flow.
Line 367: The design of the transport tubing (consisting of a contraction part and a wind tunnel) reduces mean flow problems during the sampling and make the FM-100 designed for ground-based studies.

RC: Line 782: Section 4 is titled Discussion, but is really a description of a specific set of experiments with only some of the instruments to try to explain results, and should be labeled accordingly (perhaps moved into section 3)?

AR: We agree, the ex-section 4 has been moved to the appendix as it is only devoted to the anisotropic sampling of the FSSP.

RC: Lines 862-864: It would be expected that particle speed would vary with particle size in changing flow conditions. Haven't computational fluid studies been done of these inlets that could help understand and possibly correct for this? Also, why not use the actual transit speed rather than the mean speed to calculate concentrations?

AR: The figure 13.a only show high dispersion due to turbulent flow, this dispersion is more important for small particles:

Line 1073: The results of the Figure 13.a show that droplets smaller than 20 µm have the range of the SPP transit time between 6 and 14 µs, whereas the range is between 7 and 10 µs for droplets larger than 20 µm. Small particles tend to be driven by streamlines and thus show more dispersion in SPP transit time than larger particles. This highlights anisotropy in the sampling suction, which is potentially turbulent. Indeed, in the non-isokinetic conditions and for high Reynolds number (about $2\cdot10^4$), turbulent flows are expected inside and near the FSSP inlet.

The FSSP does not provide the transit time or speed, so it cannot be used to compute the sampling speed. In the appendix, the SPP transit time is used only to try to understand its variations.

RC: Lines 873 to 882: Confusing–were these calculations done in the Gerber study or in this study?

AR: This is Gerber's results, the paragraph has been revised.

RC: Lines 890-892: "generates spurious droplet concentration" makes it sound like the instrument is creating droplets, which it is not. This sentence also seems out of place and a similar statement is made on lines 871-872. Suggest deleting it.

AR: As suggested, it has been deleted.

Lines 988-991: The concentration effect of the mean flow is also likely playing a role, and should be mentioned.

AR: the concentration effect has been added in the conclusion.

Line 1001: As this effect was more pronounced for small particles, the concentration effect of the mean flow and the presence of turbulent flow inside the FSSP inlet could be a plausible explanation of the discrepancies of the measurements based on particle counting.

RC: Fig. 1 caption: Define "LOAC" shown in the image or remove the label.

AR: The LOAC has been removed.

We are grateful to C. H. Twohy for suggested corrections. All of them were considered. The corresponding sentences were revised.