Review on AMTD manuscript by Eric Hintsa et al. "UAS Chromatograph for Atmospheric Trace Species (UCATS) – a versatile instrument for trace gas measurements on airborne platforms"

Overall appraisal and general comments:

UCATS is a fantastic airborne instrument that has produced valuable and high quality scientific data. Its development and numerous deployments over the past 15 years have been a great success story, and the manuscript by Eric Hintsa et al. is interesting and very well written. I absolutely enjoyed reading it.

However, when considering the manuscript as potential paper in Atmospheric Measurement Techniques (AMT), I am having trouble finding clear *purpose* and *added value* for the scientific community. I find the take home message that UCATS "successfully provided trace gas measurements on different types of aircraft for atmospheric science missions" a bit thin, and I would like to see something in terms of "new science" (which I think is there but not very visible): what's unique and novel about the UCATS instrument compared to similar airborne instruments or compared to what has already been presented in earlier publications? Are there any particularly innovative features or solutions to common problems that haven't been adequately solved before? In the conclusions, there is a sentence listing the "challenges encountered", but one or two paragraphs summarizing "innovative solutions found" would be better.

Currently, the manuscript comes across as a hybrid between a descriptive instrument paper, a review article on the work of the research group, and an instrument inter-comparison study. The title suggests that your intention is the former, and I suggest to make the instrument description more comprehensive, and to give more details on innovative aspects, i.e. problem solutions that have not been found or published previously. At the same time, try to shorten or compress the sometimes lengthy "history" bits, and to some extent also the instrument comparisons (concentrate on those analyses that provide valuable information on data quality and homogeneity). Maybe a little bit of restructuring would help to better see and find (i.e. if you search for it) some of the useful pieces of information, e.g. the integration of three different instruments (GC, O_3 and H_2O spectrometers) into one compact package, or how you dealt with moisture issues with both the GC part and the O_3 sensors. Here, I think dedicated subsections (with subheadings!) would be more valuable than the bits and pieces currently spread throughout the text.

In short, make the instrument description more comprehensive. Shift the focus to and provide more details on *innovations*, where appropriate put them into context of what other groups have done, and clearly list these highlights in the conclusions. This would tremendously increase the added value of your paper and I would comfortably recommend publication in AMT.

Specific comments:

Page 2, Line 5: Here in the abstract, you could be a bit more specific on the O_3 and H_2O instruments by adding the information that both are commercial sensors measuring by absorption spectroscopy.

Introduction (general):

While the second paragraph on the background and history of the HATS group and the UCATS instrument is certainly interesting, the information given is to some extent redundant with the more detailed and more technical development history given in Section 3. Maybe that part can be shortened here in the introduction with reference to Section 3, or alternatively, you give Section 3 a more technical/topical structure and focus and move the *time line view* to the introduction entirely.

Preceding this, a few lines could be added on why airborne gas chromatography is the preferred method to measure certain trace species and what the advantages and limitations are in general terms. Some references to at least the most important other airborne GC groups/instruments may also be appropriate.

Page 2, Line 32: The statement here is actually not correct, at least with respect to higher altitudes: to my knowledge, the ceiling altitude of the piloted ER-2 is at least equal to that of the unmanned Global Hawk.

Section 2 (general):

The statement in the introduction (page 2, line 55) suggests that the instrument description in Section 2 exclusively refers to the original design of the first version. Is this really the case, and does the block diagram in Figure 1 really reflect this first design? The caption and references made in the text are not always clear in that respect.

In any case, I think it would be most useful to attempt to describe the *latest/best version* in Section 2, in particular with respect to the modular layout. Please try to consider the following:

- It would be interesting to learn in what way the sub-instruments (up to three GC channels, O_3 and H_2O spectrometers) efficiently share the space of a common housing, a common power supply and a common PC architecture combining the data in real time. In other words: what are the advantages of packaging everything together in one instrument box rather than flying these instruments separately? How much weight/space/power-draw is saved?

- Optimized weights and dimensions should already be given in the context of the statement on “compact design for small payload spaces” (page 3, line 8), because people’s interpretation of what “compact” and “small payload space” actually mean may vary. Maybe mission specific information with respect to box size and weights, and whether gas bottles and O_3 and H_2O sensors were integrated or placed outside the box can be summarized in a table that can be referenced here and in other places where appropriate.

- Provide detailed information on what gas bottles are used and at what pressure. Do they need to be refilled for each flight?

<u>Page 5, line 1:</u> By “a series of …”, do you mean more than one sensor in one setup, or one sensor for each different instrument version?

<u>Page 5, lines 19 – 28:</u> I’d like to see more detailed information here, i.e.

- What is the overall weight of the entire package with everything that is needed to measure? How has this changed over time?

- Please provide more information on the power system: how much power is needed for the different parts (GCs, O_3 and H_2O sensors)? Do you use any converters, filters, etc.? Are there any EMC issues you had to deal with?

- Does Ampro Computer actively control the instruments or is it just used for data acquisition? Are the data from the GCs and the spectroscopic sensors synchronized and merged into one file? Ideally, the main features of the instrument software would be briefly described and the code be made available via a repository.

<u>Page 6, lines 25 – 31:</u> I find the description of what you did and how successful it was a bit sparse and hard to fully understand. Maybe this can be expanded?

<u>Page 7, lines 1 – 7:</u> I see the advantage of flying two sensors to increase time resolution from 10 s to 5 s, but I don’t how the addition of the second 2B would substantially remedy the *precision at 100 hPa issue* and bring precision to anywhere near 1 ppb. Please explain this in more detail and convincingly.

<u>Page 7, lines 25 – 31:</u> The described solution to remove moisture that can be applied at large pressure differentials could be interesting to other groups facing similar problems. Consider going into even more detail here, possibly including a drawing so the air flow can be more easily visualized. If you have results from controlled experiments with air of varying moisture, it would be great if you could show them to demonstrate how much moisture can be removed this way, and how it quantitatively solves the problems.

<u>Page 7, lines 33 – 48:</u> It would be good to see some results from your experiments to ensure the accuracy of the H_2O sensors. Also, it would be good to see how consistent the short path absorption / long path absorption / long path 2nd harmonic detection are in the mixing ratio ranges where they overlap.

<u>Page 7, lines 54/55:</u> As the next UCATS deployment is with the ER-2, and the caption to Table 1 suggest that another repackaging is currently being carried out, it might be an idea to add one or two sentences here.

<u>Page 8, lines 12 – 14, 45 – 46 and Figures 2 + 3:</u> I'm not convinced the use of the 1997 ACATS data qualifies as *comparison*. The improvement in precision from ACATS to UCATS seems trivial, and I'm not sure the correlation plots are the best way to demonstrate let alone quantify this improvement. Nevertheless, the demonstration of consistency between the instruments flown in different time periods may be useful for people working with these data to address long term records and trends. Such a consistency demonstration would, of course, become even more useful if more data from the many UCATS missions were included. According to Table 1, N_2O and SF_6 were always measured, and CH_4 was measured always but in the first Altair campaign. And from the text I read that stratospheric air was sampled at least for some flight sections in many of these campaigns. Besides these general considerations, it would be good to provide a more detailed rationalization including references for the tropospheric growth of N_2O, SF_6 and CH_4.

<u>Page 10, lines 8 – 21:</u> Determining precision using atmospheric measurements of "near-constant N_2O" is not ideal because you make assumptions about natural variability. You are trying to work around that by using concurrent data from other instruments (mainly the QCLS) to filter out this natural variability, but you're only replacing the natural variability assumption by an assumption on QCLS precision. Why did you not measure UCATS (including as much of the sampling system as possible) precision in a dedicated experiment repeatedly measuring a standard? To me, that would seem the most intuitive and straightforward method. Alternatively (or additionally) it could be interesting to show the variability over all in-flight CAL gas measurements. If you inject that every 9 minutes, you should have quite good statistics over each mission.

<u>Page 10, lines 21 – 27:</u> Here you discuss accuracy, which I regard as the more relevant parameter when it comes to instrument comparisons. I find it a bit disturbing that the slope of the fit between QCLS and UCATS N_2O being so substantially different from 1.0 is not discussed in more detail. Looking at Figure 4b, there is a high bias of UCATS compared to QCLS at N_2O mixing ratios in the high 200s (stratospheric/mixed/polar vortex outflow?) of 2 – 4 ppb. Depending on the science you do with the data, this can be significant! Even if you haven't been able to fully resolve the reasons for this bias, a detailed discussion and possibly a

recommendation for users of the HIPPO dataset seems warranted. Interestingly, looking at the top right panel of Figure 5, the bias in the same mixing ratio range during ATom appears to be reversed, i.e. now UCATS is lower than QCLS.

Figures 4, 5 and 6: I think it would make sense to use the same symbol for UCATS data in all three comparison figures.

Figure 5: How useful is the lower right panel plotting UCATS/PANTHER/PFP SF6 against QCLS N_2O? In my opinion, a plot of UCATS vs. PANTHER SF_6 would be much more useful here.

Page 12, lines 25 – 26, Table A1: For people using UCATS data in scientific studies (whom I regard as at least one key audience of this paper), Table A1 is probably one of the most useful pieces of information in the paper. Thus, I first suggest moving this table from the appendix to the main body. And it could also be a good idea to use this table as the basis for the discussion of the instrument evolution in terms of the quality metrics *precision* and *accuracy*. The caption of Table A1 states that the NOAA airborne instruments use the same standards and scales as the surface network. It would be good to include more information on the laboratory and in-flight calibration procedures using these standards (currently, the CAL bottle is shown in Figure 1, and it is stated on page 5, line 21 that this calibration gas is injected every 9 minutes). Most importantly, show results from the calibrations as the first measures for precision and accuracy, and only then discuss comparison to other instruments. Finally, a simple number of 1 ppb for agreement of UCATS N_2O with other instruments for HIPPO is obviously not valid over the entire measurement range (cf. my previous comment). Please check all numbers in the table carefully and don't be too simplistic and/or two optimistic with them!

Page 13, line 15: It is not clear to me why Fig 3A is moved to the Appendix, other than for example Fig 7. Any figure that is discussed in detail in the paper should be shown within the main paper and not in an appendix or supplement.

Page 13, line 26: As an O_3 standard is not something you typically get in a bottle, please describe this NIST standard or at least provide an appropriate reference.

Page 15, lines 13 – 15: I don't buy the conclusion of moisture retained in the scrubber explaining the UCATS high bias, at least not based on the evidence presented. There is no doubt that the red line in the top panel of Fig 9 falls above the green line on the ascent sections, but to my eye, it does even more so on the descent sections, when the aircraft moves from drier to wetter air masses.

Page 19, lines 23 – 27: This list of important and valuable science papers using UCATS data could be merged into the introduction. Strictly, they are not summarizing the results of this particular paper.

Page 20, lines 8 – 10: This sounds as if at least the design of this layout is already available. As this new design seems to be a major or even some kind of *final* step in the instrument development (of course, such instruments never stop evolving), it would be good to actually include it in this paper (see my comment on Sections 2 and 3 above).

Appendix A (general):

The layout and content of this *Appendix* appear more like a typical AMT *Supplement*. I suggest moving any Figures and Tables that are really additional information to a supplement and drop the appendix. But as indicated in several comments above, a number of items in the current Appendix probably belong to the main paper and should be moved there.

And I'm not sure about the strategy to leave the description of the O_3 and H_2O sensors more or less in the main text but move any (or almost all) illustrative material to the Appendix. Either move the Figures also to the main text, or, alternatively, move the entire description of these commercial sensors into an Appendix (in this case, an Appendix would indeed be appropriate).

Figure A3: What's the peculiar group of points just below the 1:1 line at ~2500 ppb O_3? It does not really fall out of the range, but it visibly sticks out so at least a short sentence on this *feature* seems warranted.

Page 19, lines 6 – 9: If you include a supplement, then why not show a figure on this very nice comparison that you mention also?