Dear Editor Wiebke Frey,

We are very grateful for the suggestions provided by the editor and each of the external reviewers. I sincerely appreciate all constructive comments to improve the paper quality. Please see our detailed responses to comments with the changes made in the manuscript. We also attached the marked-up manuscript version at the end.

We revised the manuscript to provide more insights about the comparison, and the main changes are:

1) we included more details in the aerosol chemical composition comparison section 3.3.3 and added two figures (Fig 11 and 12);

2) we expanded the discussion in the trace gas and aerosol number concentration section 3.2 ;

3) we modified cloud probe comparison section to include details on the data processing (Line 587-596 and Line 628-636);

4) we ameliorated the comparison results with a new linear regression approach (Table 4). More details are provided in the responses to specific comments.

Thank you very much for your support and help during the process.

Best regards,

Fan and all co-authors.

Reply to Anonymous Referee #1:

We appreciate the referee efforts in reviewing our manuscript. However, we regret that the referee does not consider the manuscript worthy of publication. For reasons that we explain in the following, we do not share this opinion.

1. Many of the authors of this manuscript have participated in twenty-plus aircraft measurement campaigns. Based on their experiences, a merely technical document comparing instruments to each other for each campaign is not a common practice.
2. Thus, we feel that such comparison studies should appear in the referenced literature to educate the community on how to examine the measurements obtained from more than one aircraft, including pointing out potential issues. This valuable information is very important to share with a broader audience.
3. Also, publications of instrument intercomparisons from field or laboratory campaigns are common in the refereed literature, in AMT and other journals. Here are a few examples:

   - Brock et al. (2019): "Aerosol size distributions during the Atmospheric Tomography Mission (ATom): methods, uncertainties, and data products", https://www.atmos-meas-tech.net/12/3081/2019/amt-12-3081-2019-discussion.html

   - Meyer et al. (2015): Two decades of water vapo measurements with the FISH fluorescence hygrometer: A review., ACP, 15, 8521–8538, https://doi.org/10.5194/acp-15-8521-2015, 2015.

   - Fahey et al (2014): The AquaVIT-1 intercomparison of atmospheric water vapor measurement techniques, AMT, 7, 3177–3213, https://doi.org/doi:10.5194/amtd-7-3177-2014.

   - Rollins et al. (2014): Evaluation of UT/LS hygrometer accuracy by intercomparison during the NASA MACPEX mission, JGR, 119, 1915–1935, https://doi.org/doi:10.1002/2013JD020817, 2014.

RC2: 'referee report', Anonymous Referee #4

Experimental studies of the atmosphere using aircraft are extremely important and multi-aircraft experiments are often performed to expand the range of measurements or the spatial or temporal scales. Sometimes comparisons between these different aircraft platforms are performed and are often very instructive for those involved since they improve the measurements and identify any issues with the data processing or instruments. However, they are rarely published. It is therefore good to see that the authors are trying to provide this for the major study carried out above the Amazon region. These exercises are often very important and allow the data sets from both aircraft to be combined and integrated together. This is very useful and the paper achieves this aim by providing statistical comparisons between the platforms. In this sense it provides a useful contribution to the ACRIDICON-CHIVA experiment.

Response: We thank the reviewer for all the valuable comments and appreciate the suggestions the reviewer made. Below, we have added our responses to the comments submitted.

Unfortunately, it does little more and this is an opportunity missed. It would have been very good to see a more insightful discussion of the instrument performance, pre and post flight calibration details and what happens if these are not carried out. What, if any ground comparisons were carried out and how useful these were to the overall performance of the instruments? Were data analysis approaches compared and what did these yield? A more detailed discussion of these topics would provide some real insight and information for others carrying out similar work, whether in a single aircraft project or when multiple aircraft are being used together. I would strongly recommend that this is carried out in a revised manuscript and that some of the sections are removed such as the extrapolation of the size distribution to smaller sizes and the radiation sections.

Response: Thank you very much for your suggestions. We changed the manuscript to provide more insights about the instruments' performance: 1) we included more details in the aerosol chemical composition comparison section 3.3.3 and added two figures (Fig 11 and 12); 2) we expanded the discussion in the trace gas and aerosol number concentration section 3.2 ; 3) we modified cloud probe comparison section to include details on the data processing (Line 587-596

and Line 628-636); 4) we ameliorated the comparison results with an additional linear regression approach (Table 4). More details are provided in the responses to specific comments.

Comments Line 193-207: Given that there are some discrepancies in the AMS measurements it would be very useful to have more information on the inlets and sample tubing for the two instruments, particularly the pressure controlled inlet systems. Were the instruments calibrated before and after each flight or if not when were the calibrations performed? Were all the instrument parameters (ionization efficiency/air beam, flow rates etc) varying in a consistent way throughout the experiment? How was the CE determined?

Response: The constant pressure inlets used by both G1 and HALO AMS were very similar to the design by Bahreini et al. 2008 (the reference had been included in section 2.1.3). Both AMS instruments were calibrated before and after the field deployment and also once a week during the field campaign (line 210-211). More details about the AMS measurements are given in separate AMS papers from the respective groups (Schulz et al., 2018; Shilling et al., 2018). A short summary is now included in the supplemental material. For example, the CE of the G1 AMS was determined by comparing AMS data to UHSAS and FIMS data. We confirmed the CE=0.5 by comparing mass loadings observed at the T3 site to the G1 data. For HALO AMS, CE of 0.5 was applied, as recommended by Middlebrook et al. (2012) for low nitrate conditions.

Line 414-420: No comment is made about the two sets of points at the start of the comparison which show enhancements in aerosol number in both aircraft at separate times, presumably one shortly after another. This gives rise to an increase in the uncertainty statistics but not the regression since the values are relatively low. It might also be good to discuss the breadth of points in the CPC regression since it could almost be argued that the pairs of points fall around two different regression lines.

Response: We included further discussion of the CPC difference in lines 443 – 457 with additional plots in Figure 6.

Lines 434-439: If you can demonstrate that the aerosol sources are systematically different in the two profiles from the G1 and HALO then I don't see any justification for including the plot in the

paper since there is no information to be gained. I suggest a clearer and more detailed explanation of why the aerosol sources in the two measured profiles are different and then a statement stating that this is the reason for not including the comparison, or if this cannot be satisfactorily demonstrated the statement of causality should be removed.

Response: Based on the aerosol number concentration, chemical composition, and CO concentration data, we believe that the G1 and HALO were sampling different air masses at altitudes between 2000 and 3000 m. Thus, we excluded these data points on replotted figure 8 and revised the corresponding discussion in section 3.3.1.

Line 471: I would recommend the removal of section 3.3.3. This is already a long paper and contains considerable amounts of detailed information. This section doesn't really show any comparison as such, it simply says that extrapolating a particle number size distribution below 100 nm based on optical particle size distribution information alone will underestimate the particle concentration if there is a small aerosol mode. In deep convection such particles can be activated and so extrapolations are to be treated with caution in environments where this occurs. A comment to this effect in the previous section is important as a caution but reducing the text would certainly help also.

Response: We moved this section to the supplemental material. The original objective was to emphasize the importance of expanding size distribution measurements below 50 nm range on an airborne platform using advanced instrumentation (e.g., FIMS). We learned from many modelers that they typically extrapolate UHSAS size distributions in their models due to scarcity of actual data below ~50 nm. Thus, we compared the measurement from FIMS to the UHSAS based extrapolations.

Line 519 and following: Despite Section 3.3.4 being titled aerosol composition there is no comment about the chemical composition only a focus on the transmission of one of the AMS inlets. The implication from what is written is that the aerosol is predominately organic. Some discussion of the composition and any difference between the two instruments discussed. This is particularly the case if the inorganic components are above the detection limit since one could

then test the effectiveness of the ion balance to derive ammonium concentrations. It would also be good to include some comment on the Collection Efficiency that is used and how this was calculated.

Response: More discussion of the chemical composition was added to the manuscript in section 3.3.3 on page 19-20.

The quality of the English, particularly through the cloud section is rather poor. This needs to be significantly improved before publication.

The revised manuscript has been reviewed by many native speaking co-authors and the cloud section is edited by a professional editor.

Lines 618-619: It is always difficult to compare cloud probes between aircraft due to the spatial and temporal distances between the two aircraft. Nevertheless, this section does fall short of any detailed insight at all. It is stated that "The difference between the G1 CDP and FCDP may be due to the data post-processing". The implication is that this wasn't checked out in detail. There is clearly no information here that can be used by a reader that would be remotely useful. I suggest that much more detailed analysis is provided for this to be useful. Why weren't the corrections for coincidence and shattering applied in a consistent manner?

Response: Thank you very much for your suggestions. We have modified the text accordingly. "The difference between the G1 CDP and FCDP is mainly due to the data post-processing. The G1 CDP used an earlier version of the data acquisition system from Science Engineering Associates, with limited capability to store the particle-by-particle (PBP) data for further processing. The CDP had an 800-μm-diameter pinhole placed in front of the sizing detector to minimize the coincidence up to 1850 cm^{-3}. On the other hand, FCDP was equipped with new electronics and PBP data was locally stored on a flash drive onboard the Linux machine. For the G1 flights, a constant probe-dependent adjustment factor was applied to FCDP to adjust the coincidence further. The G1 CDP and FCDP operated with redesigned probe tips to minimize the shattering effect. An additional algorithm was applied to the FCDP data to eliminate particles with short interarrival times."

Line 632 and following: This section says almost nothing at all and could be removed.

Response: This radiation comparison section is mainly to illustrate the challenges of comparing two radiation instruments deployed on two aircraft, including many factors which affect the accuracy of the measurements. And we also confirmed the effects of the difference in spectral sensitivity of the radiometers using the NCAR tropospheric ultraviolet and visible (TUV) radiation model.

Line 652: Uncertainty Assessment: This section is extremely qualitative and non specific. As written it serves very little purpose. Instead I would recommend a much more detailed examination of uncertainties embedded with each of the sections and for this to be made more quantitative.

Response: We have modified the uncertainty assessment section contents (page 23, section 4) and revised Table 4 to be quantitative. The information about the sources for the discrepancy between the two measurements can be useful for users to understand data uncertainty and for future field campaign planning.

Minor comments: (I stopped writing the minor corrections after a while since the latter part of the paper needs a significant revamp if it is to remain).

Line 85: uncertainty ranges

Response: corrected

Line 101: issues

Response: corrected

Line 1100: delta

Response: corrected

Line 155: section 2.1.3: were the CPCs from the G1 and the HALO run side by side on the ground for a period? If so it would be good to report this. When were the instruments calibrated relative to the field experiment? This isn't said explicitly.

Response: Unfortunately, we never got a chance to run CPCs from the G1 and HALO side by side on the ground. The aircraft were parked far apart, and such a comparison would have required un-mounting and relocation of one CPC, which was not practical during tightly scheduled field campaign. All CPCs were calibrated before and after the field campaign and checked at least once a week during the deployment.

Line 188-189: which have a refractive index

Response: corrected

Line 178-192: when were the UHSAS instruments calibrated relative to the flight periods?

Response: The UHSAS was calibrated before and after the field deployment and checked with PSL's once a week during the deployment.

Line 227: should read in present tense "are discussed"

Response: corrected

Line 234: needs to be rewritten "working independently and electronics produce shadowgraph"

Response: changed to "working independently. The 2DS electronics produce shadowgraph…"

Line 263: not sure about the use of the word "proven"

Response: changed to "examined".

Lines 247-268: How were the sample volumes of the HALO probes established? This is stated for the G1 but not HALO.

Response: line 257-258, "The sample area of the CCP- CDP was determined to be 0.27±0.025 mm^2 with an uncertainty of less than 10% (Klingebiel et al., 2015)."

Lines 276-278: It is not clear how this is actually achieved.

Response: Dr. Long have provided more details about the tilt correction in his paper (Long et la., 2010). We modified the sentence to make it clear, "Additionally, the angular offset between the

actual orientation of each radiometer's detector and the level position from the navigation data has also been determined and corrected after the installation for each deployment."

Line 304: stacked pattern

Response: corrected.

Lines 308-309: "Due to the different aircraft speeds, the flight distance between two aircraft flight paths continued increasing from 15 min to 1 hour" I do not dispute that the distances between the flight paths continued increasing but since the G1 took off first and the HALO is the faster aircraft I cannot see how the increase in time between the aircraft is due to the different aircraft speeds.

Response: We have revised the sentence: "Due to the different aircraft speeds, the time difference between two aircraft visiting the same part of the flight path varied, increasing up to 1 hour at the end of the path, as shown in Figure 3."...

Line 321: present

Response: corrected.

Line 323: intervals

Response: corrected

Line 337: "The linear regression achieved a slope was near 1" should be "The linear regression achieved a slope that was near 1"

Response: corrected

Line 340-342: This is a good way of presenting the uncertainty though I am surprised that you didn't use the orthogonal distance that would also represent the variability in x.

Response: Thank you very much for your good suggestions. We have included the orthogonal regression in new Table 4.

Line 356: when (the) G1 flew

Response: corrected

Line 418: change “rest” for “remaining”

Response: changed to “rest of the 10-15%”

Line 387-393: I am unsure why the regression statistics are presented including the point with high CO measured by the G1 but not by the HALO in Fig 5b. By all means present the data point but it does seem a little strange to include it in the reporting of the agreement.

Response: We replotted the CO comparison and modified the discussion in lines 397-419.

Line 439: sources

Response: corrected.

Line 470: “has a reduced spatial resolution”

Response: corrected.

Line 718: strophic?

Response: changed to “lower troposphere”

Interactive comment on "Comparison of Aircraft Measurements during GoAmazon2014/5 and ACRIDICON-CHUVA" by Fan Mei et al.

Anonymous Referee #5 Received and published: 26 August 2019

Mei et al. provide a comparison of datasets from two research aircraft obtained during a coordinated comparison effort. Comparisons between calibrated instruments are quite useful for evaluating whether the estimated uncertainties for each instrument do accurately represent the data quality, which is of course paramount to the usefulness of the data. Ideally, analysis of such comparisons could be used to better understand estimated uncertainties and possibly reduce those uncertainties. This paper takes on a significant effort because the authors compare all of the possible measured parameters between these two aircraft (> 10 parameters). In general, I feel that the paper would be more useful if the scope were somewhat smaller with more significant analysis and discussion of the differences between a subset of the measurements. There are a few useful recommendations for measurements going forward, but also some of the disagreements between measurements which might be considered significant are not explored enough to understand if the measurements can be reconciled. At the same time, I don't think it is reasonable to ask the authors to change what they see as the purpose of the paper, but suggest that in the future such comparisons may better serve the community by going more in-depth on a smaller group of the measurements.

We thank the reviewer for the thoughtful comments and suggestions. We also agree that the community may benefit from in-depth comparison of a smaller group of the measurements in the future. Our responses to the specific comments are described below.

I have some suggestions and changes that I would like to see the authors address. These are listed below.

Table 3: Recommend instead of highlighting only slope and R2 that the systematic differences in measurements are calculated (two measurements could be perfectly correlated with a slope of 1 yet have a huge offset and differ on average by a large fraction). Can you also include something about the expected agreement based on the uncertainties of each instrument?

Response: The fitting slope and R2 in Table 3 assume the correlation between two measurements: the G1 measurement is equal to HALO measurement. Thus, the listed slope is from y=slope* x equation without the offset, and the R2 is based on the equation 1. We modified the line 336 "assuming that two measurements from the G1 and HALO have the 1:1 relationship." We also tried the orthogonal regression to relate the measurements from the G1 and HALO, which shows the similar results as Table 3. We added another table (new Table 4) to further illustrate the uncertainties of each instrument.

Line 365: Was the fact that the G1 sensor data point bad here known before the comparison and would it have been thrown out? If so, recommend removing this point from the figure as it does represent what is thought to be good data.

Response: We replotted Figure 4(c) without the bad data points, then changed the sentence (line 366). The initial data quality control did not exclude the G1 sensor data as questionable data i.e. chilled mirror sensor wetted by cloud droplets.

Section 3.2: Ozone: Table 3 shows a minimum ozone value of 0.5 for G1. Is this correct or a typo? It seems there is a slope and offset between the ozone instruments. Difference between the means is about 17%, which I think exceeds what is expected (~ 5% each instrument). I doubt the explanation that sampling losses in the tubing could account for the difference as O3 is not too difficult to sample. Please state clearly whether the differences observed between the O3 instruments exceeds what is expected for the sensors themselves, and what evidence there is to suggest sampling loss is to blame. Possibly, a leak of cabin air into the sample line affected one of the instruments.

Response: We agree that the sampling loss is not the main reason causing the measurements difference. We have edited the lines 386-390 "As mentioned in section 2.1.2, each instrument has a 2 ppb accuracy (or 5%) on the ground based on a direct photometric measurement measuring the ratio between a sample and ozone free cell. The in-flight calibration showed that the variation of each instrument could raise to 5-7% (or 2-3.5 ppb). Thus, the difference between the averaged ozone concentrations – 4.1 ppb is close to the instrument uncertainty."

CO: Recommend removing the outlier CO point if you have good reason to believe it was not coincident. At the same time, I don't see how the explanation on 389-391 about "different operation principles" has anything to do with lack of coincidence between the measurements. Please clarify if the disagreement is because of bad coincidence or if you think the instruments really do not measure the same thing.

Response: We removed the outlier CO points from the altitude between 2000-3000 m, which we believe the G1 and HALO are sampling different air mass. We agree that the "different operation principles" should not cause the significant difference in the measurements. We modified the manuscript lines 397-419.

Line 418: Kind of weak discussion here about CPC difference. Seems like HALO is systematically lower. It would be useful to understand something about the difference rather than just state that it can be attributed to the typical uncertainties and other unknown factors. The comparison between UHSAS does not support it being an issue with the isokinetic inlets.

Response: We included further discussion of the CPC difference in between line 443 – 457 with additional figures in Figure 6.

Figure 9: Why does HALO UHSAS look so much noisier?

Response: The airborne version of the UHSAS does have an issue with maintaining constant volumetric sheath flows (discussed by Kupc et al. 2017), which directly affects concentration since the sample flow is not directly measured but calculated as the difference between total and sheath flows.

Line 488 / Fig. 11: I don't see the value of this comparison. It is stated in the text that the UHSAS < 50 nm is not a measurement, but rather an extrapolation of the distribution down to sizes the UHSAS cannot measure, and that this extrapolation could easily be invalid during e.g. a nucleation event. Therefore, I don't understand when the extrapolated UHSAS data would ever be of use for scientific analysis. The fact that the extrapolated UHSAS distribution deviates from the FIMS measurements sometimes does not even require the UHSAS instrument to determine this. One could just extrapolate the FIMS data using the UHSAS sensitivity range and look at the difference between the FIMS measurements.

Response: We moved this section to the supplemental material. The original objective was to emphasize the importance of expanding size distribution measurements to below the 50 nm range on an airborne platform using advanced instrumentation (e. g. FIMS). We learned from many modelers that they typically use extrapolations of the UHSAS size distribution in their models due to scarcity of real data. Thus, we compared the real measurement from FIMS to the UHSAS based extrapolation.

Line 512: What is referred to here had been done for decades on other aircraft and has been referred to as NMASS. Recommend citing the relevant papers for that here and earlier in the paper (i.e. lines 478 – 487). Most recently: Williamson et al., AMT 11, 3491-3509, 2018.

Response: Thank you very much for your suggestions. The section is modified in the supplemental material.

Line 519/section 3.3.4: There is no actual discussion of the chemical composition, just the mass/volume. Recommend removing the reference to chemical composition here and earlier in the paper (e.g. abstract and introduction).

Response: More discussion of the chemical composition was added to the manuscript in section 3.3.3, pages 19-20.

L 650: How about a calculation with TUV to test whether the different sensitivity ranges can account for the 10%? ———-

Response: We used the tropospheric ultraviolet and visible (TUV) radiation model from NCAR website (https://www2.acom.ucar.edu/modeling/tropospheric-ultraviolet-and-visible-tuv-radiation-model) and estimated the weighted irradiance at 15:42:00 on Sep 9 2014. Note that the modeling output is limited to the range between 315 to 900 nm. It is different from the irradiance spectral range (400-2700 nm) in the G1 or the 300-1800 nm from HALO. The difference between the two aircraft measurements was 24.1 W/m^2 at that time, and the modeling suggested the irradiance difference between 315-400 nm was 13 W/m^2. Although we can't estimate the difference between 1800-2700 nm with TUV. We have shown that the difference in spectral range of the instruments is the main contribution to the difference in the comparison.

Editorial type notes: Line 101: issures -> issues

Response: corrected.

Line 146: change comma to period

Response: corrected.

Line 304: ‘paten’ -> ‘pattern’?

Response: corrected.

Line 326: ‘Tables’ -> ‘Table’

Response: corrected.

Line 418: ‘ rest of the 10-15. . .”

Response: corrected.

Comparison of Aircraft Measurements during GoAmazon2014/5 and ACRIDICON-CHUVA

Fan Mei[1], Jian Wang[2], Jennifer M. Comstock[1], Ralf Weigel[13], Martina Krämer[14, 13], Christoph Mahnke[13, 8], John E. Shilling[1], Johannes Schneider[8], Charles N. Long[7], Manfred Wendisch[5], Luiz A. T. Machado[3], Beat Schmid[1], Trismono Krisna[5], Mikhail Pekour[1], John Hubbe[1], Andreas Giez[6], Bernadett Weinzierl[6], Martin Zoeger[6], Christiane Schulz[8], Mira L. Pöhlker[8], Hans Schlager[6], Micael A. Cecchini[9], Meinrat O. Andreae[8,10], Scot T. Martin[4], Suzane, S. de Sá[4], Jiwen Fan[1], Jason Tomlinson[1], Stephen Springston[2], Ulrich Pöschl[8], Paulo Artaxo[11], Christopher Pöhlker[8], Thomas Klimach[8] , Andreas Minikin[12], Armin Afchine[14], Stephan Borrmann[13,8]

1. Pacific Northwest National Laboratory, Richland, WA, United States.
2. Brookhaven National Laboratory, Upton, NY, United States.
3. National Institute for Space Research (INPE), São Paulo, Brazil
4. Harvard University, Cambridge, MA, United States
5. University of Leipzig, Leipzig, Germany
6. Deutsches Zentrum für Luft- und Raumfahrt (DLR), Oberpfaffenhofen, Germany
7. NOAA ESRL GMD/CIRES, Boulder, CO, United States
8. Max Planck Institute for Chemistry, Mainz, Germany
9. University of São Paulo (USP), São Paulo, Brazil
10. Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA
11. Instituto de Física, Universidade de São Paulo, São Paulo, Brazil
12. DLR Oberpfaffenhofen, Flight Experiments Facility, Wessling, Germany
13. Institute for Physics of the Atmosphere, Johannes Gutenberg University, Mainz, Germany
14. Research Centre Jülich, Institute for Energy and Climate Research 7: Stratosphere (IEK-7), Jülich, Germany

Correspondence to: Fan Mei (fan.mei@pnnl.gov)

Abstract. The indirect effect of atmospheric aerosol particles on the Earth's radiation balance remains one of the most uncertain components affecting climate change throughout the industrial period. ~~This issue~~The large uncertainty is part~~ial~~ly ~~a result of the~~ due to the incomplete understanding of aerosol-cloud interactions. One objective of the GoAmazon2014/5 and ACRIDICON-CHUVA projects was to ~~improve the understanding~~ understand ~~of~~ the influence of the emissions ~~of~~ from the tropical megacity of Manaus (Brazil) on the surrounding atmospheric

Formatted: Font: 10 pt

Formatted: Font: 10 pt

environment of the rainforest and to investigate its role in the life cycle of convective clouds. During one of the intensive observation periods (IOPs) in the dry season from September 1 to October 10, 2014, comprehensive ~~instrument suites collected data~~ measurements of trace gases and aerosol properties were carried out at ~~from~~ several ground sites. In a coordinated way, the advanced suites of sophisticated in situ instruments were deployed ~~in situ both~~ aboard both ~~from~~ the U.S. Department of Energy Gulfstream-1 (G1) aircraft and the German High Altitude and Long-Range Research Aircraft (HALO) during three coordinated flights on September 9, 21, and October 1. Here we report on the comparison of measurements collected by the two aircraft during these three flights. Such comparisons are ~~difficult to obtain, but they are~~ challenging but essential for assessing the data quality from the individual platforms and quantifying their uncertainty sources. Similar instruments mounted on the G1 and HALO collected vertical profile measurements of aerosol particles number concentrations and size distribution, cloud condensation nuclei concentrations, ozone~~,~~ and carbon monoxide ~~concentration~~ mixing ratios, cloud droplet size distributions, and downward solar irradiance. We find that the above measurements from the two aircraft agreed within ~~the range given by~~ the measurement uncertainties. The relative fraction of the~~A~~ aerosol chemical composition measured by instruments on HALO agreed with the corresponding G1 data, although the total mass loadings only ~~collected~~ have a good agreement at high altitudes ~~only~~. Furthermore, possible causes of the discrepancies between ~~the data sets collected by~~ measurements on the G1 and HALO ~~instrumentation~~ are ~~addressed~~ examined in this paper. Based on these results, criteria for meaningful aircraft measurement comparisons are discussed.

1. Introduction

Dominated by biogenic sources, the Amazon basin is one of the few remaining continental regions where atmospheric conditions realistically represent those of the pristine or pre-industrial era (Andreae et al., 2015). As a natural ~~atmospheric~~ climatic "chamber", the area around the urban region of Manaus in central Amazonia is an ideal location for studying the atmosphere under natural conditions as well as under conditions influenced by human activities and biomass burning events (Andreae et al., 2015; Artaxo et al., 2013; Davidson et al., 2012; Keller et al., 2009; Kuhn et al., 2010; Martin et al., 2016b; Pöhlker et al., 2018; Poschl et al., 2010; Salati and Vose, 1984).

The Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5) campaign was conducted in 2014 and 2015 (Martin et al., 2017; Martin et al., 2016b). The primary objective of GoAmazon2014/5 was to improve the quantitative understanding of the effects of anthropogenic influences on atmospheric chemistry and aerosol-cloud interactions in the tropical rainforest area. During the dry season in 2014, the ACRIDICON (Aerosol, Cloud, Precipitation, and Radiation Interactions and Dynamics of Convective Cloud Systems)-CHUVA (Cloud Processes of the Main Precipitation Systems in Brazil) campaign also took place to study tropical convective clouds and precipitation over Amazonia (Wendisch et al., 2016).

A feature of the GoAmazon 2014/5 field campaign was the design of the ground sites' location, which uses principles of Lagrangian sampling to align the sites with the Manaus pollution plume (Figure 1: Source location – Manaus (T1 site), and downwind location – Manacapuru (T3 site)). The ground sites were overflown with the low-altitude U.S. Department of Energy (DOE) Gulfstream-1 (G1) aircraft and the German High Altitude and Long Range Research Aircraft (HALO). These two aircraft are among the most advanced in atmospheric research, deploying suites of sophisticated and well-calibrated instruments (Schmid et al., 2014; Wendisch et al., 2016). The pollution plume from Manaus was intensively sampled during the G1 and HALO flights and also by the DOE Atmospheric Radiation Measurement (ARM) program Mobile Aerosol Observing System and ARM Mobile Facility located at one of the downwind surface sites (T3 site- 70 km west of Manaus). The routine ground measurements with coordinated and intensive observations from both aircraft provided an extensive data set of multi-dimensional observations in the region, which serves i) to improve the scientific understanding of the influence of the emissions of the tropical megacity of Manaus (Brazil) on the surrounding atmospheric environment of the rainforest and ii) to understand the life cycle of deep convective clouds and study open questions related to their influence on the atmospheric energy budget and hydrological cycle.

As more and more data sets are merged to link the ground-based measurements with aircraft observations, and as more studies focus on the spatial variation and temporal evolution of the atmospheric properties, it is critical to quantify the uncertainty~~ies~~ ranges when combining the data collected from the different platforms. Due to the challenges of airborne operations, especially when two aircraft are involved in data collection in the same area, direct comparison studies are rare. However, this type of ~~study~~ research is critical for further combining the datasets between the

ground sites and aircraft. Thus, the main objectives of the study herein are to demonstrate how to achieve ~~—~~meaningful comparisons between two moving platforms, ~~—~~to conduct detailed comparisons between data collected by two aircraft, to identify the potential measurement issues, to quantify reasonable uncertainty ranges of the extensive collection of measurements, and to evaluate the measurement sensitivities to the temporal and spatial variance. The comparisons and the related uncertainty estimations quantify the current measurement limits, which provide realistic measurement ranges to climate models as initial conditions to evaluate their output.

The combined GoAmazon2014/5 and ACRIDICON-CHUVA field campaigns not only provide critical measurements of aerosol and cloud properties in an under-sampled geographic region but also ~~provide~~ offer a unique opportunity to understand and quantify the quality of these measurements using ~~closely~~ carefully orchestrated comparison flights. The comparisons between the measurements from similar instruments on the two research aircraft can be used to identify potential measurement issu~~r~~es and quantify the uncertainty range of the field measurements, which include primary meteorological variables (Section 3.1), trace gases concentrations (Section 3.2), aerosol particle properties (number concentration, size distribution, chemical composition, and microphysical properties) (Section 3.3), cloud properties (Section 3.4), and downward solar irradiance (Section 3.5). We evaluate the consistency between the measurements aboard the two aircraft for a nearly full set of gas, aerosol particle, and cloud variables. Results from this comparison study provide the foundation not only for assessing and interpreting the observations from multiple platforms (from the ground to low altitude, and then to high altitude) but also for providing high-quality data to improve the understanding of the accuracy of the measurements related to the effects of human activities in Manaus on local air quality, terrestrial ecosystems in rainforest, and tropical weather.

2. Measurements

2.1 Instruments

The ARM Aerial Facility deployed several in situ instruments on the G1 to measure atmospheric state parameters, trace gases concentrations, aerosol particle properties, and cloud characteristics (Martin et al., 2016b; Schmid et al., 2014). The instruments installed on HALO covered measurements of meteorological, chemical, microphysical, and radiation parameters. Details of measurements aboard HALO are discussed in the ACRIDICON-CHUVA campaign

overview paper (Wendisch et al., 2016). The measurements compared between the G1 and HALO are listed in Table 1.

2.1.1 Atmospheric parameters

All G1 and HALO meteorological sensors were routinely calibrated to maintain measurement accuracy. The G1 primary meteorological data were provided at one-second time resolution based on the standard developed by the Inter-Agency Working Group for Airborne Data and Telemetry Systems (Webster and Freudinger, 2018). For static temperature measurement, the uncertainty given by the manufacturer (Emerson) is ±0.1 K, and the uncertainty of the field data is ±0.5 K. The static pressure had a measurement uncertainty of 0.5 hPa. The standard measurement uncertainties were ±2 K for the chilled mirror hygrometer and 0.5 ms^{-1} for wind speed.

On HALO, primary meteorological data were obtained from the Basic HALO Measurement and Sensor System (BAHAMAS) at one-second time resolution. The system acquired data from airflow and thermodynamic sensors and from the aircraft avionics and a high-precision inertial reference system to derive the basic meteorological parameters like pressure, temperature, the 3D wind vector, aircraft position, and attitude. Water vapor mixing ratio and further derived humidity quantities were measured by the Sophisticated Hygrometer for Atmospheric Research (SHARC) based on direct absorption measurement by a tunable diode laser (TDL) system. The absolute accuracy of the primary meteorological data was 0.5 K for air temperature, 0.3 hPa for air pressure, 0.4-0.6 ms^{-1} for wind, and 5% (±1 ppm) for water vapor mixing ratio. All sensors were routinely calibrated and traceable to national standards (Giez et al., 2017; Krautstrunk and Giez, 2012).

2.1.2 Gas phase

Constrained by data availability, this comparison of trace gas measurements is focused on carbon monoxide (CO) and ozone (O_3) concentrations. Those measurements were made aboard the G1 by a CO/N_2O/H_2O instrument (Los Gatos Integrated Cavity Output Spectroscopy instrument model 907-0015-0001), and an Ozone Analyzer (Thermo Scientific, Model 49i), respectively. The G1 CO analyzer was calibrated for response daily by NIST-traceable commercial standards before the flight. Due to the difference between laboratory and field conditions, the uncertainty of the CO measurements is about ±5% for one-second sampling periods. An ultra-fast carbon monoxide monitor (Aero Laser GmbH, AL5002) was deployed on HALO. The detection

of CO is based on a vacuum-ultraviolet-fluorimetry, employing the excitation of CO at 150 nm, and the precision is 2 ppb, and the accuracy is about 5%. The ozone analyzer measures ozone concentration based on the absorbance of ultraviolet light at a wavelength of 254 nm. The ozone analyzer (Thermo Scientific, Model 49c) in the HALO payload is very similar to the one on the G1 (Model 49i), with an accuracy greater than 2 ppb or about ±5% for four-second sampling periods. The G1 ozone monitor was calibrated at the New York State Department of Environmental Conservation testing laboratory at Albany.

2.1.3 Aerosol

Aerosol number concentration was measured by different condensation particle counters (CPCs) on the G1 (TSI, CPC 3010) and HALO (Grimm, CPC model 5.410). Although two CPCs were from different manufacturers, they were designed using the same principle, which is to detect particles by condensing butanol vapor on the particles to grow them to a large enough size that they can be counted optically. Both CPCs were routinely calibrated in the lab and reported the data at one-second time resolution. The HALO CPC operated at 0.6-1 L min^{-1}, with a nominal cutoff of 4 nm. Due to inlet losses, the effective cutoff diameter increases to 9.2 nm at 1000 hPa, and 11.2 nm at 500 hPa (Andreae et al., 2018; Petzold et al., 2011). The G1 CPC operated at 1 L min^{-1} volumetric flow rate and the nominal cut-off diameter D_{50} measured in the lab was ~10 nm. During a flight, the cut-off diameter may vary due to tubing losses, which contributes less than 10 % uncertainty to the comparison between two CPC concentrations.

Two instruments deployed on the G1 measured aerosol particle size distribution. a Fast Integrated Mobility Spectrometer (FIMS) inside of the G1 cabin measured the aerosol mobility size from 15 to 400 nm (Kulkarni and Wang, 2006a, b; Olfert et al., 2008; Wang, 2009). The ambient aerosol particles were charged after entering the FIMS inlet and then separated into different trajectories in an electric field based on their electrical mobility. The spatially separated particles grow into super-micrometer droplets in a condenser where supersaturation of the working fluid is generated by cooling. At the exit of the condenser, a high-speed charge-coupled device camera captures the image of an illuminated grown droplet at high resolution. In this study, we used the FIMS 1 Hz data for comparison. The size distribution data from FIMS were smoothed. Aside from the FIMS, The airborne version of the Ultra High Sensitivity Aerosol Spectrometer (UHSAS) was deployed on G1 and HALO. The G1 and HALO UHSAS were manufactured by

the same company, and both were mounted under the wing on a pylon. UHSAS is an optical-scattering, laser-based particle spectrometer system. The size resolution is around 5% of the particle size. The G1 UHSAS typically covered a size range of 60 nm to 1000 nm. HALO UHSAS covered 90 nm to 500 nm size range for the September 9 flight.

Based on operating principles, FIMS measures aerosol electrical mobility size and UHSAS measures the aerosol optical equivalent size. Thus, the difference in the averaged size distributions from those two types of instruments might be linked to differences in their underlying operating principles, such as the assumption in the optical properties of aerosol particles. The data processing in the G1 UHSAS assumed that the particle refractive index is similar to ammonium sulfate (1.55), which is larger than the average refractive index (1.41-0.013i) from a previous Amazon study (Guyon et al., 2003). The HALO UHSAS was calibrated with polystyrene latex spheres, which haves a refractive index about 1.572 for the UHSAS wavelength of 1054 nm. The uncertainty due to the refraction index can lead to up to 10% variation in UHSAS measured size (Kupc et al., 2018). Also, the assumption of spherical particles affects the accuracy of UHSAS sizing of ambient aerosols.

The chemical composition of submicron non-refractory ($NR\text{-}PM_1$) organic and inorganic (sulfate, nitrate, ammonium) aerosol particles was measured using a high-resolution time-of-flight aerosol mass spectrometer (HR-ToF-AMS) aboard the G1 (DeCarlo et al., 2006; Jayne et al., 2000; Shilling et al., 2018; Shilling et al., 2013). Based on the standard deviation of observed aerosol mass loadings during filter measurements, the HR-ToF-AMS detection limits for the average time of thirteen seconds are approximately 0.13, 0.01, 0.02, 0.01 (3σ values) $\mu g\ m^{-3}$ for organic, sulfate, nitrate, and ammonium, respectively (DeCarlo et al., 2006). A Compact Time-of-flight Aerosol Mass Spectrometer (C-ToF-AMS) was operated aboard HALO to investigate the aerosol composition. Aerosol particles enter both the C-ToF-AMS and HR-ToF-AMS via constant pressure inlets controlling the volumetric flow into the instrument, although the designs of the inlets are somewhat different (Bahreini et al., 2008). The details about the C-ToF-AMS operation and data analysis are reported in Schulz's paper (Schulz et al., 2018). The overall accuracy has been reported as ~30 % for both AMS instruments (Alfarra et al., 2004; Middlebrook et al., 2012). Data presented in this section were converted to the same condition as the HALO AMS data, which is 995 hPa and 300 K. Both AMS instruments were calibrated before and after the field deployment and also once a week during the field campaign.

The number concentration of cloud condensation nuclei (CCN) was measured aboard both aircraft using the same type of CCN counter from Droplet Measurement Technologies (DMT, model 200). This CCN counter contains two continuous-flow, thermal-gradient diffusion chambers for measuring aerosols that can be activated at constant supersaturation. The supersaturation is created by taking advantage of the different diffusion rates between water vapor and heat. After the supersaturated water vapor condenses on the CCN in the sample air, droplets are formed, counted, and sized by an Optical Particle Counter (OPC). The sampling frequency is one second for both deployed CCN counters. Both CCN counters were calibrated using ammonium sulfate aerosol particles in the diameter range of 20-200 nm. The uncertainty of the effective water vapor supersaturation was ±5%. (Rose et al., 2008)

2.1.4 Clouds

Aircraft-based measurements are an essential method for in situ samplings of cloud properties (Brenguier et al., 2013; Wendisch and Brenguier, 2013). Over the last 50–60 years, hot-wire probes have been the most commonly used devices to estimate liquid water content (LWC) in the cloud from research aircraft. Since the 1970s, the most widely used technique for cloud droplet spectra measurements has been developed based on the light-scattering effect. This type of instrument provides the cloud droplet size distribution as the primary measurement. By integrating the cloud droplet size distribution, additional information, such as LWC can be derived from the high-order data product.

Three cloud probes from the G1 ~~were~~ are discussed in this manuscript. The Cloud Droplet Probe (CDP) is a compact, lightweight forward-scattering cloud particle spectrometer that measures cloud droplets in the 2 to 50 μm size range (Faber et al., 2018). Using ~~a~~ state-of-the-art electro-optics and electronics, Stratton Park Engineering (SPEC Inc.) developed a Fast Cloud Droplet Probe (FCDP), which also use forward-scattering to determine cloud droplet distributions and concentrations in the same range as CDP with up to 100 Hz sampling rate. The G1 also carried a two-dimensional stereo probe (2DS, SPEC Inc.), which has two 128-photodiode linear arrays working independently ~~and~~. The 2DS electronics produce shadowgraph images with 10 μm pixel resolution. Two orthogonal laser beams cross in the middle of the sample volume, with the sample cross section for each optical path of 0.8 cm^2. The manufacturer claims the maximum detection

size is up to 3000 μm for the 2DS. However, due to the counting statistic issue, the data used in this study is from 10–1000 μm only (Lawson et al., 2006). 2DS was upgraded with modified probe tips, and an arrival time algorithm was applied to the 2DS data processing. Both efforts effectively reduced the number of small (shattered) particles (Lawson, 2011). For G1 cloud probes, the laboratory calibrations of the sample area and droplet sizing were performed before the field deployment. During the deployment, biweekly calibrations with glass beads were performed with the size variation of less than 5%, which were consistent with the pre-campaign and after-campaign calibrations. Comparison between the LWC derived from cloud droplet spectra with hot-wire LWC measurement was made to estimate/eliminate the coincidence errors in cloud droplet concentration measurements (Lance et al., 2010; Wendisch et al., 1996)

On board of HALO, two cloud probes were operated and discussed in this manuscript, each consisting of a combination of two instruments: ~~the~~ Cloud Combination Probe (CCP) and a Cloud Aerosol Precipitation Spectrometer (CAPS, denoted as NIXE-CAPS; NIXE: Novel Ice Experiment). The CCP is a combination of a CDP (denoted as CCP-CDP) with a CIPgs (Cloud Imaging Probe with grey ~~-~~scale, DMT, denoted as CCP-CIPgs). NIXE-CAPS consists of a CAS-Dpol (Cloud and Aerosol Spectrometer, DMT, denoted as NIXE-CAS) and a CIPgs (denoted as NIXE-CIPgs). CIPgs is an optical array probe comparable to the 2DS operated on the G1. CIPgs obtains images of cloud elements using a 64-element photodiode array (15μm resolution) to generate two-dimensional images with nominal detection diameter size range from 15 to 960 μm (Klingebiel et al., 2015; Molleker et al., 2014).The CCP-CDP detects the forward-scattered laser light by cloud particles in the size range of 2.5 to 46 μm. The sample area of the CCP- CDP was determined to be 0.27±0.025 mm^2 with an uncertainty of less than 10% (Klingebiel et al., 2015). CAS-Dpol (or NIXE-CAS) is a light scattering probe comparable to the CDP but covers the size range of 0.6 to 50 μm in diameter, thus including the upper size range of the aerosol particle size spectrum (Luebke et al., 2016). Furthermore, CAS-Dpol measures the polarization state of the particles (Costa et al., 2017). Correspondingly to the G1 CDP, the performance of the CCP-CDP and NIXE-CAS were frequently ~~proven~~ examined by glass beads calibrations. Prior to or after each HALO flight, CCP-CIPgs and NIXE-CIPgs calibrations were performed by using a mainly transparent spinning disc that carries opaque spots of different but known size. The data of the CCP measured particle concentration on board of HALO are corrected to gain ambient conditions

using a thermodynamic approach developed by (Weigel et al., 2016). For NIXE-CAPS, the size distributions were provided where NIXE-CAS was merged with the NIXE-CIPgs at 20 μm.

2.1.5 Solar radiation

The G1 radiation suite included shortwave (SW, 400 - 2,700 nm) broadband total upward and downward irradiance measurements using Delta-T Devices model SPN-1 radiometers. The radiation data were corrected for aircraft tilt from the horizontal reference plane. A methodology has been developed (Long; et al., 2010) for using measurements of total and diffuse shortwave irradiance and corresponding aircraft navigation data (latitude, longitude, pitch, roll, heading) to calculate and apply a correction for platform tilt to the broadband hemispheric downward SW measurements. Additionally, whatever angular offset there may be between the actual orientation of each radiometer's detector and what the navigation data say is level has also been determined for the most accurate tilt correction.

The Spectral Modular Airborne Radiation measurement sysTem (SMART-Albedometer) was installed aboard HALO. Depending on the scientific objective and the configuration, the optical inlets determining the measured radiative quantities can be chosen. The SMART-Albedometer has been utilized to measure the spectral upward and downward irradiances; thereby, it is called as an albedometer, as well as to measure the spectral upward radiance. The SMART-Albedometer is designed initially to cover measurements in the solar spectral range between 300 and 2,200 nm (Krisna et al., 2018; Wendisch et al., 2001; Wendisch et al., 2016). However, due to decreasing sensitivity of the spectrometers at large wavelengths, the use of the wavelengths was restricted to 300 – 1,800 nm. The spectral resolution is defined by the full width at half maximum (FWHM), which is between 2 and 10 nm. In this case, the instruments were mounted on an active horizontal stabilization system for keeping the horizontal position of the optical inlets during aircraft movements (up to +/- 6 degrees from the horizontal plane).

2.2 Flight patterns

During the dry season IOP (September 1 – October 10, 2014), two types of coordinated flights were carried out: one flight in cloud-free condition (September 9) and two flights with clouds present (September 21 and October 1). In this study, we compare the measurements for both coordinated flight patterns. The discussion is mainly focused on the flights under cloud-free

conditions on September 9 and the flight with clouds present on September 21, as shown in figure 1. The other coordinated flight on October 1 is included in the supplemental document.

For the cloud-free coordinated flight, the G1 took off first and orbited around an area from the planned rendezvous point until HALO arrived in sight. It then coordinated with HALO and performed a wing-to-wing maneuver along straight legs around 500 m above sea level, as shown in Figure 2. The normal G1 average sampling speed is 100 m s^{-1}, and the normal HALO average sampling speed is 200 m s^{-1}. During the coordinated flight on September 9, both aircraft also adjusted their normal sampling speed by about 50 m s^{-1} so that they could fly side by side.

For the second type of coordinated flights, the G1 and HALO flew the stacked pattern at their own ~~normal~~ typical airspeed. On September 21, the G1 also took off from the airport first, followed by HALO 15 minutes later. Then, both aircraft flew above the T3 ground site and subsequently flew several flight legs stacked at different altitudes. The two aircraft were vertically separated by about 330 m and sampled below, inside, and above clouds. Due to the different aircraft speeds, the ~~flight distance~~time difference between two aircraft visiting the same part of the flight paths ~~varied, continued~~ increasing ~~from 15 min to~~ up to 1 hour at the end of the path, as shown in Figure 3. On October 1, the G1 focused on the cloud microphysical properties and contrasting polluted versus clean clouds. HALO devoted the flight to the cloud vertical evolution and life cycle and also probed the cloud processing of aerosol particles and trace gases. The G1 and HALO coordinated two flight legs between 950–1250 m above the T3 site under cloud-free conditions. Following that, HALO flew to the south of Amazonia, and the G1 continued sampling plume-influenced clouds above the T3 site, and then flew above the Rio Negro area.

In this study, to perform a meaningful comparison of in situ measurements, all the data from instruments were ~~time~~ time-synchronized with the aircraft (G1 or HALO) navigation system. For AMS and CPC data, the time shifting due to tubing length and instrument flow had been corrected. For the coordinated flight on Sep. 9, the data compared was from the same type of measurements with the same sampling rate. For the measurements with the different sampling rate, the data were binned to the same time interval for comparison. For the flight with the cloud present~~s~~ (Sep.21 and Oct. 1), the following criteria are used: 1) the data collected by the two aircraft must be less than 30 mins apart from each other; 2) the comparison data were~~as~~ binned to 200 m altitude intervals; ~~and~~ 3) the cloud flag was applied to the aerosol measurements, and the data affected by the cloud shattering are eliminated from the comparisons of aerosol measurements. Moreover, additional

comparison criteria are specified for individual measurements in the following section. Table~~s~~ 2 shows the total number of points used for the comparison.

3. Results

3.1 Comparison of the G1 and HALO measurements of atmospheric state parameters

The atmospheric state parameters comprise primary variables observed by the research aircraft. The measurements provide essential meteorological information not only for understanding the atmospheric conditions but also for providing the sampling conditions for other measurements, such as those of aerosol particles, trace gases, and cloud microphysical properties.

For cloud-free coordinated flights, the comparison focused on the near side-by-side flight leg at around 500 m, as shown in Figure 2. Table 3 shows the basic statistics of the data for primary atmospheric state parameters, assuming that two measurements from the G1 and HALO have a proportional relationship without any offset (Y=m0*X). In general, the atmospheric state parameters observed from both aircraft were in excellent agreement. The linear regression achieved a slope that was near 1 for four individual measurements. The regression is evaluated using the equation below ~~equation 1~~.

$$R^2 = 1 - \frac{SS_{regression}}{SS_{Total}} \quad (1)$$

Where the sum squared regression error is calculated by $SS_{regression} = \sum(y_i - y_{regression})^2$, and the sum squared total error is calculated by $SS_{Total} = \sum(y_i - \bar{y})^2$, y_i is the individual data point, $\bar{y}$ is the mean value, and $y_{regression}$ is the regression value. When the majority of the data points are in a narrow value range, using the mean is better than the regression line, and the R^2 will be negative (Neg in Table 3).

The difference between the average ambient temperatures on the two aircraft was 0.5 K, and the difference between the average dew point temperatures was about 1 K. For temperature and humidity, the G1 data were slightly higher than the HALO data. The main contributions to the observed differences include the error propagation in the derivation of the ambient temperature from the measured temperature, instrumental-measurement uncertainty, and the temporal and spatial variability. The average horizontal wind speed measured by HALO is 0.4 m s^{-1} higher than the average horizontal wind speed measured by the G1. The uncertainty source of wind estimation is mainly due to the error propagation from the indicated aircraft speed measurement and the

aircraft ground speed estimation from GPS. The static pressure distribution measured aboard HALO showed a smaller standard deviation (0.9 hPa) compared to the value of the G1 (1.5 hPa). ~~The standard deviation (std) was also 0.6 hPa narrower.~~ Part of the reason for this difference is a more substantial variation of the G1 altitude during level flight legs when the G1 flew at around 50 m/s higher than its normal airspeed. Thus, any biases caused by their near side-by-side airspeeds being different from their ~~normal~~ typical airspeeds would be undetected during these coordinated flights.

For the coordinated flights under cloudy conditions, we used the criteria from Section 2 to compare ambient conditions measured by the G1 and HALO aircraft. In addition to the ordinary linear regression, we also used the orthogonal regression to minimize the perpendicular distances from the data points to the fitted line. The ordinary linear regression assumes only the response (Y) variable contains measurement error but not the predictor (X), which remains unknown when we start the comparison between the measurements from the G1 and HALO. Thus, the additional orthogonal regression exams the assumption in the least square regression and makes sure the roles of the variables have little influence on the results. In Table 4, two equations were used for the orthogonal regression. One assumes that two measurements have a proportional relationship (Y=m1*X). The other one assumes a linear relationship, which can be described with the slope-intercept equation Y=m*X+b. Two regression results in Table 4 doesn't show a significant difference. The regression using the slope-intercept equation shows different level of improvement in each individual measurement and will be discussed in the corresponding sections.

Formatted: Font: (Default) Times New Roman

As shown in Figure 4, the linear regression slopes for ambient temperature, dew point temperature, and pressure were also close to 1 between the G1 and HALO measurements during the September 21 coordinated flight. The R^2 value is also close to 1. These results suggest that the G1 and HALO measurements achieved excellent agreements. Note that ~~T~~the dew point temperature from the G1 measurement was erroneous and removed from the comparison the data points between 2200–2700 m and above 3700 m (Figure 4(c)) because the G1 sensor was skewed by wetting in the cloud. The HALO dew point temperature was calculated from the total water mixing ratio measured by TDL, and that measurement in the cloud was more accurate than the measurement made by the chilled mirror hydrometer aboard the G1.

The lower value of the R^2 value in horizontal wind speed means the ratio of the regression error and total error in wind measurement is much higher than the temperature and pressure

measurements. The main contributions to this difference are the error propagation during the horizontal wind speed estimation and ~~also~~ the temporal and spatial variance between two aircraft sampling location. We observed differences between the two aircraft data of up to 2 m s^{-1}, caused by the increasing sampling distance as the two aircraft were climbing up. For example, the G1 flew a level leg above T3 around 2500 m between 16:20-16:30, while HALO stayed around 2500 m for a short period and kept climbing to a higher altitude. Due to strong vertical motion, turbulence, and different saturations (evaporation-condensation processes), the variances in the horizontal wind speed (Figure 4(d)) were also more significant compared to the variances of temperature and pressure measurements.

3.2 Comparison of trace gas measurements

For the cloud-free coordinated flight, ozone is the only trace gas measurement available on both aircraft. The linear regression slope shows that the HALO ozone concentration was about 8% higher than the G1 concentration. The difference between the averaged ozone concentrations was 4.1 ppb. As mentioned in section 2.1.2, each instrument has a 2 ppb accuracy (or 5%) on the ground based on a direct photometric measurement measuring the ratio between a sample and ozone-free cell. The in-flight calibration suggested that the accuracy of each instrument could raise to 5-7% (or 2-3.5 ppb). Thus, the difference between the averaged ozone concentrations – 4.1 ppb is within the instrument variation. The primary source of bias is probably the different ozone loss in the sampling and transfer lines.

Formatted: Indent: First line: 0"

The comparison made on September 21 flight in Figure 5 shows good agreement for the vertically averaged ozone measurements. Comparing the statistics data from September 9, the ozone measurement is not sensitive to the temporal and spatial changes. Although we do not have the comparison data on September 9, ~~T~~the G1~~-~~ and HALO CO measurements comparison shows~~s~~ a higher correlation than the ozone data comparison at different altitudes on September 21. ~~low R^2 value which is mainly from the systematic bias between the two instruments with different operation principles. T~~ Note that the data points with ~~larger~~ more substantial variance ~~between 2000-3000 m~~ in CO concentration ~~indicates the spatial variation contribution~~ were excluded~~,~~ ~~while~~ because the G1 and HALO were sampling different air masses between 2000-3000 m, as indicated in Figure S7. The CO plot in Figure (5b) shows the real atmospheric variability. Around 4000 ft,

the CO reading from the G1 and HALO has the minimum variation and is averaged around 85 ppb, which is at the atmospheric background level. At lower altitudes and higher CO concentrations, the local contribution is not well-mixed, and the inhomogeneity expresses as the more substantial variations observed in the plot.

3.3 Comparison of aerosol measurements

Aerosol particles exhibited ~~strong~~ substantial spatial variations, both vertically and horizontally, due to many aerosol sources and complex atmospheric ~~processing~~ processes in the Amazon basin, especially with the local anthropogenic sources ~~at~~ in Manaus. Thus, ~~any~~ spatially resolved measurement~~s is~~ are critical to characterizing the properties of the Amazonian aerosols. The cloud-free coordinated flights ~~provide us with suitable data~~ allow us to compare the G1 and HALO aerosol measurements and thus will ~~enable~~ facilitate further studies ~~combined with the ground measurements~~ that utilize the airborne measurements. The vertical profiles obtained using the G1 and HALO ~~platforms~~ in different aerosol regimes ~~in~~ of the Amazon basin have contributed to many studies (Fan et al., 2018; Martin et al., 2017; Wang et al., 2016).

~~When comparing the measurements from the two aircraft, the inlet system is a critical item, especially for sampling aerosol particles (Wendisch et al., 2004). Inlet~~ The design and ~~characterization~~ performance of the aircraft inlets can ~~actively modify~~ strongly influence ~~the~~ measured aerosol particle number concentration, size distribution, and chemical composition (Wendisch et al., 2004)~~.~~. Therefore, they need to be taken into consideration when comparing the measurements aboard two aircraft. The G1 aerosol inlet is a fully automated isokinetic inlet. ~~Based on the m~~Manufacturer wind tunnel test and ~~peer-reviewed publications,~~earlier studies show that this inlet operates for aerosol particles with diameter up to 5 μm, with transmission efficiency around 50 % at 1.5 μm (Dolgos and Martins, 2014; Kleinman et al., 2007; Zaveri et al., 2010). The HALO sub-micrometer Aerosol Inlet (HASI) was explicitly designed for HALO. Based on the numerical flow modeling, optical particle counter measurements, and field study evaluation, HASI has a cut-off size of 3 μm, with transmission efficiency larger than 90 % at 1 μm (Andreae et al., 2018; Minikin et al., 2017).

3.3.1 Aerosol particle number concentration

For the cloud-free coordinated flight, the linear regression of CPC and UHSAS between the G1 and HALO measurements are also included in Table 3. ~~Also, t~~The total number concentration ~~of~~ measured by HALO CPC ~~data~~ was about 20 % lower than ~~the total number concentration from~~that by the G1 CPC, as shown in Figure 6 (a). The CPC measurement is critically influenced by the isokinetic inlet operation and performance. During the flights, the aircraft attitude, such as the pitch and roll angles will cause the isokinetic sampling under non-axial condition. The non-axial flow at the probe inlet may result in flow separation, turbulence, and particle deposition. Therefore, the quantitative particle measurements have a more substantial uncertainty. As shown in Figure 6 (b), we compared the CPC data by applying three different data quality criteria. The first criterion is the same criteria described in the previous section, and the linear regression is included in Table 3. The second criterion constrains the data under the isokinetic and iso-axial condition and the plot in Figure 6(b) shows the iso-axial criteria reduced the broadness of the scattered data, but no significant change to the linear regression. We further constrained the data with the averaging. Based on the average wind speed and distance between two aircraft, we averaged the data into 10 seconds interval and found that the regression R^2 increase to 0.9392. ~~-~~The typical uncertainty between two CPCs is 5-10% on the well-controlled environment.(Gunthe et al., 2009; Liu and Pui, 1974) ~~Many factors can contribute to the rest 10-15% difference.~~Although both CPCs from the G1 and HALO were characterized in the lab to be within 10% with its respective lab standard, we observed 20% variance during the flight. ~~-~~This result suggests the challenging condition of airborne condition can significantly increase the ~~ose~~ ~~include~~ systematic uncertainties of CPC measurements, such as systematic instrument drifts, different aerosol particle losses inside of the two CPCs, and different inlet transmission efficiencies in the two aircraft.

Formatted: Superscript

The CPC data in Figure 6 are color-coded with UTC time. The general trend is that the aerosol number concentration increased with ~~aircraft sampling~~ time through the Manaus plume between 15:30~~-~~ and 15:40. A similar trend was observed in aerosol particle number concentration (Figure 7) measured by the Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)-Airborne version (referred to as UHSAS). The total number concentration data given by UHSAS (Figure 7) is integrated over the overlapping size range (90 – 500 nm for the September 9 flight) for both the G1 and HALO UHSAS. The linear regression shows that the total aerosol particle number concentration from HALO UHSAS is about 16.5% higher than th~~e total aerosol concentration~~at from the G1 UHSAS. The discrepancy between the two UHSAS measurements is mainly due to

the error propagation in the sampling flow, the differential pressure transducer reading, the instrument stability, and calibration repeatability, ~~which is~~ consistent with the other UHSAS study (Kupc et al., 2018). In the airborne version of UHSAS, mechanical vibrations have a more significant impact on the pressure transducer reading than the case for the bench version of UHSAS.

For the coordinated flight on September 21, the G1 and HALO data are averaged to ~~a~~ 200 m vertical altitude intervals (~~, as shown in~~ Figure 8). ~~There was a good agreement in the CPC comparison, especially at altitudes of 200 m and above (<10 % variance). However, the linear regression slope significantly decreased, which was primarily related to the temporal and spatial differences in aerosol number concentration. Especially~~ The data points with an altitude between 2000 – 3000 m altitude were excluded from the comparisons, because the ~~difference between the~~ G1 and HALO ~~measurement is largely due to the~~sampled different ~~aerosol srouces~~airmass, as ~~shown in Figure S6(a)~~ evidenced from trace gas and aerosol chemical composition data (detailed in Section 3.2 and 3.3.3). The UHSAS size range was integrated from 100~~–~~ to 700 nm ~~for UHSAS~~ on September 21. The ~~change~~ variation of the size range was because the overlap of size distributions from both UHSAS instruments was changed. ~~The linear regression slope and the R^2 value slightly decreased in the UHSAS comparison as shown in Figure 8(b). And we can expect that the main contribution to the UHSAS measurement difference is from instrument systematic drift and the spatial/temporal variance of particle concentration in the ambient environment as shown in Figure S6(b). However, combining with~~ Both the CPC and UHSAS measurement comparisons show stronger variation at the low altitude, especially below 2000 m. Above 3500 m, the variations on the CPC and UHSAS measured concentration significantly smaller than the variation at the lower altitude~~,~~. This result is consistent with the observation from the trace gas measurement and ~~it indicates~~ confirms that ~~the smaller size~~ the variability of aerosol ~~particles (~~properties changes significantly with time and space.~~< 100 nm) have a more profound variance due to the temporal and spatial change.~~ It is noticeable that the discrepancy observed in the UHSAS measurements comparison is larger than that from the CPC comparison. That is because the aerosol flow control inside the UHSAS can't respond quickly enough to the rapid change of the altitude and caused significant uncertainty in the data.

3.3.2 Aerosol particle size distribution

For the cloud-free coordinated flight, the averaged aerosol size distributions measured by FIMS, G1 UHSAS, and HALO UHSAS during one flight leg ~~is~~ are compared in Figure 9. ~~Based on the comparison plot, at the size range less than~~For particle diameter below 90 nm, the G1 UHSAS overestimated the particle ~~counts~~concentration, which is due to the ~~error~~ uncertainty in ~~the~~ counting efficiency correction. The UHSAS detection efficiency is close to 100% for particles larger than 100 nm and concentrations below 3000 cm^{-3} but decreases considerably ~~both~~ for both smaller particles and ~~for~~ higher concentrations ~~considerably~~ (Cai et al., 2008). The aerosol counting efficiency correction ~~determined from~~ developed under the lab conditions does not ~~represent the real~~necessarily apply under the conditions ~~correction~~ during the flight ~~operation~~. Between 90 nm and 250 nm, FIMS agreed well with the G1 UHSAS, whereas HALO UHSAS wa~~i~~s about 30 % higher than the ~~other~~ two instruments. For the size range of 250–500 nm, FIMS had good agreement with HALO UHSAS~~, whereas FIMS is~~ and was about 30-50 % higher than the G1 UHSAS depending on the particle size. Because the UHSAS has a simplified "passive" inlet, the large size aerosol particle loss in the UHSAS inlet was expected to increase with the increasinge of the aircraft speed. Thus, the lower G1 UHSAS ~~counts~~ concentrations at a larger aerosol particle size are likely related to the particle loss correction.

For the September 21 flight, the vertical profiles of aerosol size distributions ~~vertical profiles we~~are averaged into 100 m altitude intervals (Figure 10). Overall, all size distribution measurements captured the mode near 100 nm between 800–1000 m~~, which is~~ at the top of the convective boundary layer, as indicated by the potential temperature (Figure 10(d)), which starts from a maximum near the ground and then becomes remarkably uniform across the convective boundary layer. ~~With the increase in altitude, we observe t~~The peak of the aerosol size distribution shifted from 100 nm to 150 nm with increasing altitude. Note that due to data availability, the aerosol size distribution data from the HALO UHSAS has a ~~less~~ reduced ~~spatial~~ vertical resolution.

~~3.3.3 A significant contribution of small aerosol particles~~

~~Comprehensive characterization of aerosol particles, especially small ones (<50 nm) during GoAmazon2014/5 has demonstrated that high concentrations of those small particles in the lower free troposphere are transported from the free troposphere into the boundary layer by a strong convective downdraft and sustain the population of particles in the pristine Amazon boundary layer. This important observation improved the current understanding of the aerosol~~

~~influence on cloud properties and climate under natural conditions (Fan et al., 2018; Wang et al., 2016). However, the aerosol particle size distribution measurement, especially for sizes less than 50 nm, is very rare due to the lack of high-frequency airborne measurements. The most common aerosol size spectrometer, UHSAS, covers aerosol particle sizes larger than 60 nm. The scanning mobility particle spectrometer cannot obtain size distribution in 1 Hz time resolution. The other approach is to estimate particle size distribution by extrapolating the UHSAS or Passive Cavity Aerosol Spectrometer Probe (size range 100 – 3000 nm) measured aerosol size distribution to smaller size ranges (down to 10 nm). The accuracy of the third approach is limited by the nature of the aerosol size distribution, and the aerosol particle concentration can be significantly underestimated if there is a dominant nucleation mode in the aerosol particle size distribution, such as during a new particle formation event.~~

~~As shown in Figure 11, we compared the integrated aerosol number concentrations between one wet season flight (on March 7), which was influenced by a long-range transport plume from Africa (Moran-Zuloaga et al., 2018) and one typical dry season flight (September 21). The agreement of the small aerosol number concentration between the FIMS measured size distribution and UHSAS/PCASP estimated size distribution is reasonably good for the dry season flight when the accumulation mode dominated the aerosol particle size distribution. During the wet season, there was a strong vertical gradient in the particle size spectrum above central Amazonia under clean conditions. Thus, we can observe an increase of underestimation of the small size particle concentration both for the size ranges less than 50 nm and less than 100 nm, as the filled markers move away from each other with the increase of altitude. Because of the negligible mass contribution to the total aerosol loading, those ultrafine aerosol particles (< 50 nm) are conventionally considered too small to affect cloud formation. However, the new observational evidence and numerical simulation of deep convective clouds outlined a new mechanism, which suggests an energetic anthropogenic invigoration of deep convective clouds by those ultrafine aerosol particles in previously pristine regions of the world (Fan et al., 2018). Two newly published studies (Fan et al., 2018; Wang et al., 2016) emphasize the importance of the airborne observation and suggest the ultrafine aerosol particles (<50 nm) measurement should be included as a baseline routine measurement in future airborne experiments.~~

~~For field studies without the deployment of FIMS, one option to assess the accuracy of UHSAS/PCASP estimated size distribution is to compare the total number concentration based on the integration of the UHSAS/PCASP estimated size distribution to the total number concentration from CPC. For field study focusing on the high concentration and variability of sub-50 or sub-100 nm aerosol particles, such as new particle formation events, it is highly recommended to request the deployment of FIMS. Due to the limited availability of FIMS, one option is to use several well-characterized CPCs, which operate at the different cut-off sizes, to measure the ambient aerosol simultaneously, and then use the data inversion technique to estimate the aerosol size distribution of sub-50 or sub-100 nm aerosol particles. Another reasonable substitute to the FIMS might be a Scanning Mobile Particle Sizer (SMPS), but it should be noted that on an airborne platform an SMPS does not nearly have the same time resolution as a FIMS. To better adapt the spatial change in aerosol concentration, a residence chamber similar to a system described in another study (Kotchenruther and Hobbs, 1998) should be deployed with SMPS.~~

~~3.3.4~~3.3.3 Aerosol particle chemical composition

~~Figure 12(a) shows vertical profiles of the total mass concentrations measured by the two AMS instruments on September 21. Above 2500 m altitude, the agreement between the two instruments is excellent (mean difference less than 5%). Between 2000 and 2500 m, the agreement is within the uncertainty range. Below 2000 m altitude, however, the aerosol particle mass concentrations measured by the AMS operated on HALO are lower than the concentrations measured by the AMS on the G1. The aerosol volume concentrations from G1 AMS was converted from the mass concentration from AMS, by assuming the organic compound density was 1.5 g cm-3 (Pöschl et al., 2010). The converted aerosol volume concentration agreed well with the volume concentration calculated based on UHSAS data below 2500 m, as shown in Figure 12(b). The agreement at lower altitudes suggests that the lower concentration in HALO AMS is due to the transmission efficiency issue in the constant pressure inlet used by HALO AMS. This inlet was a prototype, designed and built at MPIC Mainz, and works by changing the size of the critical orifice that regulates the flow into the aerodynamic lens. The design and transmission characteristics will be described in an upcoming publication (Molleker, S., in prep.). The AMS aboard the G-1 used a constant pressure inlet based on the design in Bahreini et al., 2008. Thus, we conclude that data above 2500 m altitude measured by the AMS aboard HALO in 2014 are valid, while data below 2500 m need to be corrected using correction factors derived from laboratory characterization before further study. After 2014, the HALO inlet design was improved to address the inlet transmission issues specific to this field campaign.~~

Figure 11 shows the vertical profiles of the aerosol mass concentrations measured by the two AMS on September 21. The upper panel shows the medians and interquartile ranges of the different species (organics, nitrate, sulfate, ammonium) and the total mass concentration for the G1 (circles) and HALO (triangles). The lower panel shows the difference between the medians of G1 and HALO. The error bars were calculated using error propagation from the error of the median (interquartile range divided by 2*sqrt(N)). The data were grouped into 400 m altitude bins. The total mass concentration is the highest in the lower altitudes between 100 m and 2000 m with a median value of 5 µg m^{-3} (G1-AMS). At altitudes between 2000 m and 3800 m, the aerosol mass concentration decreased to a median value of 1.2 µg m^{-3} (G1-AMS).

The most significant difference was observed at altitudes below 1800 m. The aerosol mass concentration measured by HALO-AMS is less than that measured by G1-AMS, likely due to particle losses in the constant pressure inlet (CPI) used on the HALO-AMS. Between 1800 m and 3000 m, the mass concentrations measured by the HALO-AMS exceed those measured by the G1-AMS. This is most likely because the G1 was sampling different air masses than the HALO as indicated by the differences in CO mixing ratios and the CPC concentrations for this altitude region (see Fig. 5 and 8). Above 3000 m altitude, both instruments agree within the uncertainty range.

Among individual species, the largest difference above 2000 m is observed for ammonium. The deployed G1-AMS is a high resolution mass spectrometer (HR-ToF), whereas the HALO-AMS has a lower resolution (C-ToF). The higher resolution of the G1-AMS allows for a better separation of interfering ions at m/z 15, 16, and 17 (NH+, NH2+, NH3+) and thereby a more reliable calculation of the ammonium mass concentration.

Overall the aerosol chemical composition is dominated by organics, as is evident from the vertical profiles of the relative fractions (Fig. 12). Both AMS show a dominant contribution of organics to the total mass concentration with values around 70 %. This contribution is constant at altitudes between 100 m and 3500 m and decreases to 50 % at 3800 m altitude. The inorganic fraction has the highest contribution from sulfate (20 %), followed by ammonium (7 %) and nitrate (2 – 4 %). For organics, ammonium, and sulfate both instruments give similar relative fractions, only for nitrate where a discrepancy is observed between 1000 and 3000 m. Although the absolute aerosol mass concentration measured by the HALO-AMS was affected by the constant pressure inlet below 1800 m altitude, the relative fractions of both instruments generally agree well. Similar

Formatted: Normal, Indent: First line: 0"

results were found for a second comparison flight on October 1, 2014 (see supplemental plots S13, S14).

~~3.3.5~~3.3.4 CCN number concentration

These measurements provide ~~valuable~~ key information about the aerosol's ability to form cloud droplets and thereby modify the microphysical properties of clouds. Numerous laboratory and field studies have improved the ~~current~~ understanding of the connections ~~between~~ among aerosol particle size, chemical composition, mixing states and CCN activation properties (Bhattu and Tripathi, 2015; Broekhuizen et al., 2006b; Chang et al., 2010; Duplissy et al., 2008; Lambe et al., 2011; Mei et al., 2013a; Mei et al., 2013b; Pöhlker et al., 2016; Thalman et al., 2017). In addition, based on the simplified chemical composition and internal mixing state assumption, various CCN closure studies have achieved success within ±20% uncertainty for ambient aerosols (Broekhuizen et al., 2006a; Mei et al., 2013b; Rissler et al., 2004; Wang et al., 2008).

According to earlier studies (Gunthe et al., 2009; Pöhlker et al., 2016; Roberts et al., 2001; Roberts et al., 2002; Thalman et al., 2017), the hygroscopicity (κ_{CCN}) of CCN in the Amazon basin is usually dominated by organic components (κ_{Org}). Long-term ground-based measurements at the Amazon Tall Tower Observatory also suggest~~ed that there were~~ low temporal variability and ~~no~~ lack of pronounced diurnal cycles in hygroscopicity only under natural rainforest background conditions (Pöhlker et al., 2018; Pöhlker et al., 2016).

Using FIMS and CCN data from both the G1 and HALO collected during the coordinated flight leg on September 9, the critical activation dry diameter (D_{50}) was determined by integrating FIMS size distribution to match the CCN total number concentration. Then, the effective particle hygroscopicity was derived from D_{50} ~~was combined with~~ and the CCN-operated supersaturation ~~to derive the effective particle hygroscopicity by applying~~using the κ-Köhler theory. The histogram plots based on the density of the estimated hygroscopicity (κ_{est}) from both aircraft were compared for the flight leg above T3. ~~For the G1 and HALO data, t~~The κ_{est} values derived from the G1 and HALO measurements during the flight leg above the T3 site ~~is~~ are 0.19~~86~~±0.0~~6~~7 and 0.1~~8~~9±0.08, ~~3 separately~~respectively. ~~That~~ Those values ~~is~~ agree very well with~~also slightly higher than~~ the overall mean ~~kappa~~ value of 0.17±0.06derived from long-term measurements ~~from~~ at the Amazon Tall Tower Observatory~~, which is 0.17±0.06~~ (Pöhlker et al., 2016; Thalman et al., 2017).

Formatted: Font: Symbol, Italic

An example of a comparison of the vertical profiles of the CCN concentrations at 0.5% supersaturation on September 21 is shown in Figure 13 as an example. The difference between the CCN measurements on the two aircraft is about 20% on average. However, tThe linear regression slope would increases to 0.9120 if we focused on the data above 2500 m. The main contributions to the difference include the difference in aerosol inlet structure, aerosol particle loss correction in the main aircraft inlet and the constant pressure inlet, the systematic inlet difference below 2500 m as shown in AMS data, and as well as the error propagation of CCN measurements.

3.4 Comparison of cloud measurements

In situ cloud measurements help to capture the diversity of different cloud forms and their natural temporal and spatial variability. The G1 CDP and FCDP were deployed under the different wing pylons, and also on the different side of the aircraft. The G1 2DS was deployed on the same side of FCDP. The HALO cloud combination probe (CCP-CDP and CCP-CIPgs) and NIXE-CAPS (NIXE-CAS and NIXE-CIPgs) were deployed under the different wing pylons but on the same side of the aircraft. On September 21, 2014, based on the aircraft location and elevation information as shown in Figure 1(b) and Figure 3, two aircraft were sampling above T3 site and passing through the same cloud field at ~1600 m flight leg and ~1900 m flight leg as shown in Figure S8 and Figure S9. We used the cloud probes data from ~1900 m flight leg for the cloud droplet number concentration comparison. Two size ranges were considered: 3-20 μm from light scattering probes (CDP vs. FCDP on the G1, CCP-CDP vs. NIXE-CAS on HALO) and 2-960 μm from combined cloud probes.

3.4.1 Comparison of cloud droplet number concentration between 3-20 μm

The For underwing cloud probes, such as the CDP and the CAS, Lance (Lance, 2012) suggests an undercounting bias of measured particle number concentration by up to 44% due to coincidence as soon as the ambient cloud particle density rises to 1000 per cm^3. At identical cloud particle densities, an earlier study (Baumgardner et al., 1985) estimates the coincidence bias for underwing cloud probes to the range at 20%. Factually, the coincidence correction depends on the instruments' individual detection volume, the air's volume flow rate through the detector and the cloud particles' residence time within the detection volume (Hermann and Wiedensohler, 2001; Jaenicke, 1972). For this comparison, coincidence bias remained unconsidered for each of the

cloud probe measurements to avoid deviations that are caused by the application of different corrections.

The primary cloud layer was observed by both the G1 and HALO between 1000-2500 m above ground. Although the two aircraft have sampled along the same flight path, the instruments probably observed different sets of the cloud due to cloud movement with the prevailing wind or different cloud evolution stages. Thus, an initial comparison focuses on the redundant instruments on the same aircraft, that measured truly collocated and synchronous on board of HALO and of the G1, respectively. In Figure 14 (a), the data of the CCP-CDP and of the NIXE-CAS are juxtaposed sampled over about 13 minutes for particles detection size ranges which were considered as most equivalent. The comparison reveals two ranges of particle number concentrations at which densification of agreeing measurements become visible. At very low number concentrations (about 10^{-1} – 10 per cm³) the presence of inactivated (interstitial) aerosols in the clear air space between the very few cloud elements should be considered. Over specific ranges, however, the fine structure of varying cloud droplet number concentration may cause the regression's scattering, indicated by cloud particle measured by one instrument whilst respective antagonist seems to measure within almost clear air – and vice versa. At higher number concentrations, i.e., between 10^2 and 10^3 per cm³, the comparison of the highly resolved data constitutes increasing compactness with respect to the 1:1 line. The overall data scatter of this comparison, however, may indicate the highly variable structure within clouds as those investigated over the Amazon basin. The data of the G1 CDP and the FCDP are juxtaposed as the same as HALO cloud probes. However, the sampled cloud period was much shorter – about 3 minutes. Similar ~~a~~to~~s~~ the HALO cloud probes comparison, we observe two ranges of particle number concentrations at which ~~a~~ densification of agreeing measurements become visible, especially for the lower number concentrations~~, in~~ (Figure 14(b)). At higher number concentrations, only a few cloud elements were observed by the G1 cloud probes. That is because the G1 was about 7-23 minutes later to pass the same location as HALO, and experienced much fewer cloud elements.

~~(Baumgardner et al., 1985; Hermann and Wiedensohler, 2001; Jaenicke, 1972; Lance, 2012)~~

3.4.2 Comparison of cloud droplet size distribution between 2-960 μm from both aircraft

Comparing the cloud probes from the G1 and HALO, the size distributions from the HALO CCP and NIXE-CAPS probes are in remarkably good agreement between 2-960 μm, and both peaked around 10 μm, as shown in Figure 15. That is because the potential effects of cloud elements' shattering on the probe measurements were considered similarly for the HALO-deployed CCP and NIXE instruments. ~~although it seems that the size distributions never match better than for the cloud particle diameter size range below 10 μm,~~ On the G1, the CDP and FCDP had a more significant difference in the size range less than 8 μm, although both of them peaked between 10-20 μm. The difference between the G1 CDP and FCDP is mainly due to the data post-processing. The G1 CDP used an old data acquisition system from the Science engineering Associates, which limited its capability to store the particle-by-particle (PBP) data for further processing. CDP had placed an 800-μm-diameter pinhole in front of the sizing detector to minimize the coincidence up to 1850 cm^{-3}. On the other side, FCDP was equipped with new electronics and PBP data was locally stored in a flash drive onboard the Linux machine. For the G1 flights, a constant probe-dependent adjustment factor was applied to FCDP to adjust the coincidence further. The G1 CDP and FCDP operated with a redesigned probe tip to minimize the shattering effect. An additional algorithm was applied to the FCDP data to eliminate particles with short interarrival times. ~~the size distributions from two aircraft are in remarkably good agreement, considering the instance that the cloud detection on board the G1 occurred 7-23 minutes after the cloud probing on board of HALO, as shown in Figure 15. On HALO, the CCP and NIXE-CAPS probes agreed very well between 2-960 μm and both peaked around 10 μm. On the G1, although CDP and FCDP has a more significant difference in the size range less than 8 μm, both of them showed the peak of the size distribution was around 15 μm. The difference between the G1 CDP and FCDP may be due to the data post-processing. Additional coincidence correction and shattering correction were applied to FCDP, but not to CDP.~~

Formatted: Superscript

For cloud ~~elements~~ e droplets larger than 2~~1~~0 μm, the difference between the obtained cloud particle size distributions from two aircraft becomes substantial (up to two orders of magnitude) which ~~may be indicative~~indicated ~~for~~the observations of two different stages within the progressing development of a precipitation cloud. ~~which~~The precipitation cloud developing process is ~~particularly~~ evidently expressed in elevated number concentrations of larger cloud elements observed during the G1 measurement that happened later. We also observed that the

general cloud characteristic is similar at different altitude levels, as shown in Figure S10. The first two of three averaged periods were chosen during the flight leg of ~1600 m and the last average period is for the flight leg ~1900 m compared in Figure 15. Due to the averaging, the fine in-cloud structure gets suppressed. The small scale variabilities inside a cloud which are illustrated by the scattering of the highly resolved measurement data from the instrument comparison (cf. Figure 14) and the temporal evolution of in-cloud microphysics are not ascertainable and furthermore are beyond the scope of this study.

3.5 Comparison of radiation measurements

In this study, the downward irradiance measured by the SPN-1 unshaded center detector was compared with the integrated downward irradiance from the SMART-Albedometer between 300–1,800 nm wavelengths in Figure 16. Only measurements from flight legs, where the G1 and HALO flew near side-by-side and at the same altitude were taken into consideration for analysis. In Figure 16, the top panel shows the time series of SPN-1 measurements, and the bottom panel shows the time series of SMART-Albedometer measurements. The black dots represented all data, and the blue circles identified data when the navigation condition was within +/- 1 degree from the horizontal level. The large scatter in the data between 15:12-15:28 and 15:35-15:40 is mainly due to the different sensor trajectories during the maneuvering of the aircraft to get to the coordinated flight position. Because of the difference of each aircraft position from horizontal, the measured signal varied from the signal of the direct component of sunlight. Each sensor might look at different directions of the sky or different parts of the clouds. In addition, both aircraft flew under scattered clouds, and this uneven sunlight blocking is another contribution to the "drop-off" behavior in the time series plots of the downward irradiance.

Comparing the G1 and HALO measurements between 15:15-1:55 using the restricted navigation criteria in Figure S15~~7~~, we observed that the G1 SPN1 irradiance is slightly higher than the integrated irradiance from the SMART-Albedometer. We used the NCAR tropospheric ultraviolet and visible (TUV) radiation model estimated the weighted irradiance at 15:42:00 on Sep 9, 2014 and confirm that the spectral variation in the instruments is the main contribution to the difference in the comparison. ~~However, the difference in the averaged irradiance is less than 10 %. That result could be due to the difference in radiometer spectral ranges: 400–2700 nm (SPN1 radiometer) vs. 300–1800 nm (SMART).~~

4 Uncertainty assessment

As mentioned in the introduction, a low-flying G1 and a high-flying HALO cover the sampling area from the atmospheric boundary layer~~, low clouds~~ to the free troposphere, and the sampling period from the dry and wet seasons (Martin et al., 2016a). This spatial coverage provides the user community with abundant atmospheric-related data sets for their further studies, such as for remote sensing validation and modeling evaluation. However, one critical step to bridge the proper usage of the observation with further atmospheric science study is to understand the measurement uncertainty in this data set, especially the variation between the coexisting measurements due to the temporal and spatial difference.

For the majority of the measurements during this field study, three primary sources contribute to the measurement variation between the two aircraft: the temporal and spatial variations, the difference in the inlet characterization, and the limitation of the instrument capability. We used both ordinary least squares (OLS) linear regression and the orthogonal distance regression (ODR) to correlate the measurements from the G1 and HALO and confirmed that the slope and R^2 are very similar for the measurements made on September 9. The results from Table 2 confirmed that the G1 and HALO measurements should be in a linear relationship without an offset if there is no altitude variation. It also shows the minimum discrepancy between two aerosol instruments (CPC or UHSAS) could be around 20%, which will include the error caused by the difference in the inlet characterization and the limitation of the instrument capability. If we assume those two measurement variation sources are not affected by the altitude, then by comparing the linear regression data from Table 3 to Table 2, we can estimate the temporal and spatial variation between two aircraft in a stack flight pattern. Three linear regression approaches were assessed, and the results are listed in Table 3. If we assume that two measurements from the G1 and HALO should not have any offsets, the OLS and ODR regressions show similar results. For the meteorological parameters, this assumption is valid. In addition, good correlations also indicate that there is no significant temporal or spatial variation during the stack pattern flight. As expected, the wind speed and the aerosol measurements show that the correlations between the measurements from the G1 and HALO significantly improved with the offset assumption. This result suggests that the temporal and spatial variation in a half-hour will add an additional 20% variance to the measured aerosol properties. This will lead a considerable uncertainty when we combine the observation

data between the ground station and airborne platform. Thus, more routine and long-term airborne measurements should be used to evaluate or constrain atmospheric modeling work. ~~The difference in the inlet characterization and the instrument error are the same between the coordinated flights on September 9 and September 21. Thus, we can examine the sensitivity of each measurement to the spatial variation by comparing two flights. For the majority of the comparisons of the September 21 flight, there are no significant spatial and temporal variation between two aircraft measurements. However, we noticed the comparison uncertainty is more significant between 2000 – 3000 m altitude than the other altitudes in the aerosol and trace gas profile, especially for the aerosol particles smaller than 100 nm. This additional difference occurring between 2000-3000 m indicates the spatial variation contribution, while the G1 and HALO were sampling different air masses. The G1 had one flight leg around 2500 m above T3 site, while HALO continued climbing through 2000-3000 m range to reach the next flight leg around 4500 m. Thus, for the G1 measurements, the data show two modes in the histogram distribution. The large mode was typically from the data when the G1 passed through the pollution plume, and the small mode represented the background value. Because the flight path of HALO did not pass through the plume, their data shows only one mode in the histogram plots, as shown in S6 and S7 in the supplemental material.~~

~~For atmospheric meteorological variables, the overall uncertainty is relatively smaller (less than 1 %) comparing to the other airborne measurements. The main contribution for the three-dimensional wind measurement is more sensitive to the spatial variation than the ambient temperature and pressure due to the complex turbulence structure in the boundary layer (see Figure S5). The other measurement affected mainly by the spatial and temporal variation is the cloud measurements, which is consistent with a previous study (Andreae et al., 2004). The considerable variation in the comparison between 2-960 μm indicates the evolution of cloud droplet size distributions (DSDs) over time and space has a more significant influence on the large droplet size, and it serves as the major contributor for the DSDs comparison.~~

~~The inlet also significantly affects the aerosol and gas phase measurements. Inlet characterizations are inherently challenging. However, comparisons as shown here can be used to assess the performance of the inlets indirectly. In this study, reasonable agreement of the total number concentration of aerosol particles between two CPCs indicated the uncertainty caused by~~

~~the main aerosol inlet difference is less than 15%. In addition to the main aerosol inlet, the particle losses caused by AMS aerosol constant pressure inlet also affects the AMS comparison below 2500 m altitude. Based on a literature survey, this study, for the first time, compares the non-refractory particle mass concentration between two aircraft measurements. Although two AMS sampled different air masses during the majority of the campaign, the excellent agreement between the two measurements from the comparison flight linked the aerosol chemical composition from the wet to dry season and from the atmospheric boundary layer to the upper troposphere.~~

~~We also noticed that the CCN and UHSAS comparisons are associated with more substantial uncertainties because of the more complex instrument designs. The aerosol flow fluctuation, the CCN column temperature fluctuation, and the stabilization of the optical particle counter all contribute to the accurate estimation of the CCN concentration. In a similar sense, the aerosol flow fluctuation, the difference in the inlet efficiency at different platform speeds, the laser temperature fluctuation, and the signal-to-noise ratio at lower size range all contribute to the considerable uncertainty of the UHSAS concentration measurement. However, the CCN hygroscopicity estimation on the near side-by-side comparison on September 9, 2014, shows very encouraging agreement. Thus, the spatial variance and the instability of the CCN and UHSAS performance both led to the variance between two aircraft of up to 50% based on the comparison scenario on September 21, 2014. This remains the most significant variance we observed during these two aircraft measurement comparisons.~~

~~The summary of the major measurement uncertainty contributed by the spatial difference between the two aircraft is listed in Table 4.~~

5 Summary

In situ measurements made by well-characterized instruments installed on two research aircraft (the G1 and HALO) during the GoAmazon 2014/5 and ACRIDICON-CHUVA campaigns were compared. Overall, the analysis shows good agreement between the G1 and HALO measurements for a relatively broad range of atmospheric-related variables in a challenging ~~strophic~~ lower troposphere environment. Measured variables included atmospheric state parameters, aerosol particles, trace gases, clouds, and radiation properties. This study outlines the well-designed

coordinated flights for achieving a meaningful comparison between two moving platforms. The high data quality was ensured by the most sophisticated instruments aboard two aircraft used the most advanced techniques, assisted with the best-calibrated/characterized procedures. The comparisons and the related uncertainty estimations quantify the current measurement limits, which provide the guidance to the modeler to realistically quantify the modeling input value and evaluate the variation between the measurement and the model output. The comparison also identified the measurement issues, outlined the associated reasonable measurement ranges, and evaluated the measurement sensitivities to the temporal and spatial variance.

The comparisons presented here were mainly from two coordinated flights. The flight on September 9 was classified as a cloud-free flight. During this flight, the G1 and HALO flew nearly side-by-side within a "polluted" leg, which was above the T3 site and across the downwind pollution plume from Manaus, and a "background" leg, which was outbound from Manaus to the west and could be influenced by the regional biomass burning events during the dry season. Both legs were at 500 m altitude and showed linear regression slopes of ambient temperature and pressure, horizontal wind speed and dew point temperature near to 1 between the G1 and HALO measurements. These comparisons provide a solid foundation for further evaluation of aerosol, trace gas, cloud, and radiation properties. The total aerosol concentration from CPC and UHSAS were compared for the 500 m flight leg above the T3 site. The UHSAS measurement had a better agreement than the CPC measurement. That is because of the minor difference in the inlet structure and instrument design between two UHSAS aboard the two aircraft. The average size distribution from both UHSAS and one FIMS in the G1 suggests that UHSAS had an over-counting issue at the size range between 60-90 nm, which was probably due to electrical noise and small signal-to-noise ratio in that size range. Good agreement in the aerosol size distribution measurement provides a "sanity" check for AMS measurements. A CCN closure study suggested that FIMS provides valuable size coverage for better CCN number concentration estimation. Based on the κ-Köhler parameterization, $\kappa_{e\text{stff}}$ observed at 500 m above the T3 site is 0.198±0.089, which is similar to the overall mean kappa from long-term ATTO measurements - 0.17±0.06 (Pöhlker et al., 2016). This similarity suggests that there is no significant spatial variability along the downwind transect, although the freshly emitted aerosol particles may have much less hygroscopicity. The difference in the ozone measurement comparison is about 4.1 ppb, which suggests that the bias due to the sampling line loss inside of the G1 gas inlet. The irradiance from the SPN1 unshaded center

detector in the G1 was compared with the HALO integrated downward irradiance between 300–1800 nm and achieved a very encouraging agreement with a variance of less than 10%.

During the second type of the coordinated flights on September 21 (with cloudy conditions), HALO followed the G1 after take-off from Manus airport; then the two aircraft flew stacked legs relative to each other at different altitudes above the T3 site. For atmospheric state parameters, nearly linear correlations between the G1 and HALO were observed for ambient pressure, temperature, and dew point temperature measurements at an altitude range from ground to around 5000 m. ~~Cloud presence affected the measurements of dew point temperature in the G1, resulting in a large discrepancy in the dew point temperature measurement and the derived relative humidity between 2000–3000 m.~~ The horizontal wind had more variation than the rest of the meteorological properties, which is mainly due to the temporal and spatial variability. The aerosol number concentration comparison ~~had an excellent agreement (<15 %) for aerosol particles larger than 10 nm counted by the CPC below 2500 m~~ confirms that inhomogeneous aerosol distribution observed by the trace gas measurements. ~~While t~~The integrated aerosol number concentration from UHSAS showed consistent discrepancy at different altitudes~~,~~ ~~that suggests~~ This considerable uncertainty in the UHSAS measurements is caused by the significant ~~temporal and spatial variation of smaller aerosol particles (<100 nm)~~ aerosol flow variations due to the slow and unstable flow control. Although the aircraft-based UHSAS is a challenging instrument to operate, a reasonable size distribution profile comparison was made between both UHSAS and FIMS o~~i~~n the G1. Overall the chemical composition of the aerosol is dominated by organics. Around 70% organic contribution maintains constant up to 3500 m, then decreases to 50%. The most substantial difference among all the species is observed for ammonium due to the different resolution, and more reliable ammonium mass concentration can be achieved with high resolution mass spectrometer. Although the absolute aerosol mass concentration measured by the HALO AMS was affected below 1800 m altitude by the constant pressure inlet, the relative fractions of both instruments from the G1 and HALO agree well. ~~The aerosol concentration measured by AMS instruments showed a mean difference less than 5% above 2500 m during the flight on September 21, although due to the ongoing study, the correction factors allowing for correction of data below 2500 m are not available yet. The difference between CCN number concentration measured on the two aircraft was on average 20%, and these data show the same altitude behavior as the AMS data. The main contributions to this difference include the difference in aerosol inlet structure, the~~

~~aerosol loss correction in a constant pressure inlet, and the error propagation of CCN measurements. The ozone and CO vertical profile comparisons show a variation of less than 10%. The ozone measurement variation is mainly from the systematic bias between two instruments with different operation principles, especially at altitudes higher than 4000 m.~~

Cloud probe comparisons were made for the cloud droplet number concentration between 3–20 µm for the initial comparison between the redundant instruments on the same aircraft. Then the comparison of cloud droplet size distribution between 2-960 µm for a flight leg around 1900 m showed a remarkably good agreement. The major cloud appearance was captured by both aircraft, although the cloud elements observed were affected by the cloud movement with the prevailing wind and the different cloud evolution stages. Furthermore, the relatively short time delay of 7-23 minutes between the independent measurements may give a hint for the time scales in which the cloud droplet spectra develop within a convective cloud over the Amazon basin.

The above results provide additional information about the reasonableness of measurements for each atmospheric variable. This study confirms the high-quality spatial and temporal dataset with clearly identified uncertainty ranges had been collected from two aircraft and builds a good foundation for further studies on the remote sensing validation and the spatial and temporal evaluation of modeling representation of the atmospheric processing and evolution.

Acknowledgments: This study was supported by the U.S. DOE, Office of Science, Atmospheric System Research Program, and used data from Atmospheric Radiation Measurement ~~Climate Research~~ Aerial Facility, a DOE Office of Science User Facility. The Pacific Northwest National Laboratory (PNNL) is operated for DOE by Battelle under contract DE-AC05-76RL01830. This work was also supported by the Max Planck Society, the DFG (Deutsche Forschungsgemeinschaft, German Research Foundation) HALO Priority Program SPP 1294, the German Aerospace Center (DLR), the FAPESP (São Paulo Research Foundation) Grants 2009/15235-8 and 2013/05014-0, and a wide range of other institutional partners. The contributions from Micael A. Cecchini were funded by FAPESP grant number 2017/04654-6.

References:

Alfarra, M. R., Coe, H., Allan, J. D., Bower, K. N., Boudries, H., Canagaratna, M. R., Jimenez, J. L., Jayne, J. T., Garforth, A. A., and Li, S.-M.: Characterization of urban and rural organic

959 particulate in the lower Fraser valley using two aerodyne aerosol mass spectrometers, Atmos
960 Environ, 38, 5745-5758, 2004.
961 Andreae, M. O., Acevedo, O. C., Araujo, A., Artaxo, P., Barbosa, C. G. G., Barbosa, H. M. J.,
962 Brito, J., Carbone, S., Chi, X., Cintra, B. B. L., da Silva, N. F., Dias, N. L., Dias, C. Q., Ditas, F.,
963 Ditz, R., Godoi, A. F. L., Godoi, R. H. M., Heimann, M., Hoffmann, T., Kesselmeier, J.,
964 Konemann, T., Kruger, M. L., Lavric, J. V., Manzi, A. O., Lopes, A. P., Martins, D. L.,
965 Mikhailov, E. F., Moran-Zuloaga, D., Nelson, B. W., Nolscher, A. C., Nogueira, D. S., Piedade,
966 M. T. F., Pohlker, C., Poschl, U., Quesada, C. A., Rizzo, L. V., Ro, C. U., Ruckteschler, N., Sa,
967 L. D. A., Sa, M. D., Sales, C. B., dos Santos, R. M. N., Saturno, J., Schongart, J., Sorgel, M., de
968 Souza, C. M., de Souza, R. A. F., Su, H., Targhetta, N., Tota, J., Trebs, I., Trumbore, S., van
969 Eijck, A., Walter, D., Wang, Z., Weber, B., Williams, J., Winderlich, J., Wittmann, F., Wolff, S.,
970 and Yanez-Serrano, A. M.: The Amazon Tall Tower Observatory (ATTO): overview of pilot
971 measurements on ecosystem ecology, meteorology, trace gases, and aerosols, Atmos Chem Phys,
972 15, 10723-10776, 2015.
973 Andreae, M. O., Afchine, A., Albrecht, R., Holanda, B. A., Artaxo, P., Barbosa, H. M. J.,
974 Borrmann, S., Cecchini, M. A., Costa, A., Dollner, M., Fütterer, D., Järvinen, E., Jurkat, T.,
975 Klimach, T., Konemann, T., Knote, C., Krämer, M., Krisna, T., Machado, L. A. T., Mertes, S.,
976 Minikin, A., Pöhlker, C., Pöhlker, M. L., Pöschl, U., Rosenfeld, D., Sauer, D., Schlager, H.,
977 Schnaiter, M., Schneider, J., Schulz, C., Spanu, A., Sperling, V. B., Voigt, C., Walser, A., Wang,
978 J., Weinzierl, B., Wendisch, M., and Ziereis, H.: Aerosol characteristics and particle production
979 in the upper troposphere over the Amazon Basin, Atmos. Chem. Phys., 18, 921-961, 2018.
980 Andreae, M. O., Rosenfeld, D., Artaxo, P., Costa, A., Frank, G., Longo, K., and Silva-Dias, M.:
981 Smoking rain clouds over the Amazon, Science, 303, 1337-1342, 2004.
982 Artaxo, P., Rizzo, L. V., Brito, J. F., Barbosa, H. M. J., Arana, A., Sena, E. T., Cirino, G. G.,
983 Bastos, W., Martin, S. T., and Andreae, M. O.: Atmospheric aerosols in Amazonia and land use
984 change: from natural biogenic to biomass burning conditions, Faraday Discuss, 165, 203-235,
985 2013.
986 Bahreini, R., Dunlea, E. J., Matthew, B. M., Simons, C., Docherty, K. S., DeCarlo, P. F.,
987 Jimenez, J. L., Brock, C. A., and Middlebrook, A. M.: Design and operation of a pressure-
988 controlled inlet for airborne sampling with an aerodynamic aerosol lens, Aerosol Sci Tech, 42,
989 465-471, 2008.
990 Baumgardner, D., Strapp, W., and Dye, J. E.: Evaluation of the Forward Scattering Spectrometer
991 Probe. Part II: Corrections for Coincidence and Dead-Time Losses, 2, 626-632, 1985.
992 Bhattu, D. and Tripathi, S. N.: CCN closure study: Effects of aerosol chemical composition and
993 mixing state, J Geophys Res-Atmos, 120, 766-783, 2015.
994 Brenguier, J. L., Bachalo, W. D., Chuang, P. Y., Esposito, B. M., Fugal, J., Garrett, T., Gayet, J.
995 F., Gerber, H., Heymsfield, A., and Kokhanovsky, A.: In situ measurements of cloud and
996 precipitation particles, Airborne Measurements for Environmental Research: Methods and
997 Instruments, 2013. 225-301, 2013.
998 Broekhuizen, K., Chang, R.-W., Leaitch, W., Li, S.-M., and Abbatt, J.: Closure between
999 measured and modeled cloud condensation nuclei (CCN) using size-resolved aerosol
1000 compositions in downtown Toronto, Atmos Chem Phys, 6, 2513-2524, 2006a.
1001 Broekhuizen, K., Chang, R. Y. W., Leaitch, W. R., Li, S. M., and Abbatt, J. P. D.: Closure
1002 between measured and modeled cloud condensation nuclei (CCN) using size-resolved aerosol
1003 compositions in downtown Toronto, Atmos Chem Phys, 6, 2513-2524, 2006b.

Cai, Y., Montague, D. C., Mooiweer-Bryan, W., and Deshler, T.: Performance characteristics of the ultra high sensitivity aerosol spectrometer for particles between 55 and 800 nm: Laboratory and field studies, J Aerosol Sci, 39, 759-769, 2008.
Chang, R. Y. W., Slowik, J. G., Shantz, N. C., Vlasenko, A., Liggio, J., Sjostedt, S. J., Leaitch, W. R., and Abbatt, J. P. D.: The hygroscopicity parameter (kappa) of ambient organic aerosol at a field site subject to biogenic and anthropogenic influences: relationship to degree of aerosol oxidation, Atmos Chem Phys, 10, 5047-5064, 2010.
Costa, A., Meyer, J., Afchine, A., Luebke, A., Günther, G., Dorsey, J. R., Gallagher, M. W., Ehrlich, A., Wendisch, M., and Baumgardner, D.: Classification of Arctic, midlatitude and tropical clouds in the mixed-phase temperature regime, Atmos Chem Phys, 17, 12219-12238, 2017.
Davidson, E. A., de Araujo, A. C., Artaxo, P., Balch, J. K., Brown, I. F., Bustamante, M. M. C., Coe, M. T., DeFries, R. S., Keller, M., Longo, M., Munger, J. W., Schroeder, W., Soares, B. S., Souza, C. M., and Wofsy, S. C.: The Amazon basin in transition (vol 481, pg 321, 2012), Nature, 483, 232-232, 2012.
DeCarlo, P. F., Kimmel, J. R., Trimborn, A., Northway, M. J., Jayne, J. T., Aiken, A. C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K. S., Worsnop, D. R., and Jimenez, J. L.: Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer, Anal Chem, 78, 8281-8289, 2006.
Dolgos, G. and Martins, J. V.: Polarized Imaging Nephelometer for in situ airborne measurements of aerosol light scattering, Optics express, 22, 21972-21990, 2014.
Duplissy, J., Gysel, M., Alfarra, M. R., Dommen, J., Metzger, A., Prevot, A. S. H., Weingartner, E., Laaksonen, A., Raatikainen, T., Good, N., Turner, S. F., McFiggans, G., and Baltensperger, U.: Cloud forming potential of secondary organic aerosol under near atmospheric conditions, Geophys Res Lett, 35, 2008.
Faber, S., French, J. R., and Jackson, R.: Laboratory and in-flight evaluation of measurement uncertainties from a commercial Cloud Droplet Probe (CDP), Atmos Meas Tech, 11, 3645-3659, 2018.
Fan, J., Rosenfeld, D., Zhang, Y., Giangrande, S. E., Li, Z., Machado, L. A., Martin, S. T., Yang, Y., Wang, J., and Artaxo, P.: Substantial convection and precipitation enhancements by ultrafine aerosol particles, Science, 359, 411-418, 2018.
Giez, A., Mallaun, C., Zöger, M., Dörnbrack, A., and Schumann, U.: Static pressure from aircraft trailing-cone measurements and numerical weather-prediction analysis, J Aircraft, 54, 1728-1737, 2017.
Gunthe, S., King, S., Rose, D., Chen, Q., Roldin, P., Farmer, D., Jimenez, J., Artaxo, P., Andreae, M., and Martin, S.: Cloud condensation nuclei in pristine tropical rainforest air of Amazonia: size-resolved measurements and modeling of atmospheric aerosol composition and CCN activity, Atmos Chem Phys, 9, 7551-7575, 2009.
Guyon, P., Boucher, O., Graham, B., Beck, J., Mayol-Bracero, O. L., Roberts, G. C., Maenhaut, W., Artaxo, P., and Andreae, M. O.: Refractive index of aerosol particles over the Amazon tropical forest during LBA-EUSTACH 1999, J Aerosol Sci, 34, 883-907, 2003.
Hermann, M. and Wiedensohler, A. J. J. o. A. S.: Counting efficiency of condensation particle counters at low-pressures with illustrative data from the upper troposphere, 32, 975-991, 2001.
Jaenicke, R. J. J. o. A. S.: The optical particle counter: cross-sensitivity and coincidence, 3, 95-111, 1972.

Jayne, J. T., Leard, D. C., Zhang, X. F., Davidovits, P., Smith, K. A., Kolb, C. E., and Worsnop, D. R.: Development of an aerosol mass spectrometer for size and composition analysis of submicron particles, Aerosol Sci Tech, 33, 49-70, 2000.
Keller, M., Bustamante, M., Gash, J., and Dias, P. S. (Eds.): Amazonia and Global Change, American Geophysical Union, Washington, DC, 2009.
Kleinman, L. I., Daum, P. H., Lee, Y. N., Senum, G. I., Springston, S. R., Wang, J., Berkowitz, C., Hubbe, J., Zaveri, R. A., and Brechtel, F. J.: Aircraft observations of aerosol composition and ageing in New England and Mid-Atlantic States during the summer 2002 New England Air Quality Study field campaign, Journal of Geophysical Research: Atmospheres, 112, 2007.
Klingebiel, M., de Lozar, A., Molleker, S., Weigel, R., Roth, A., Schmidt, L., Meyer, J., Ehrlich, A., Neuber, R., Wendisch, M., and Borrmann, S.: Arctic low-level boundary layer clouds: in situ measurements and simulations of mono- and bimodal supercooled droplet size distributions at the top layer of liquid phase clouds, Atmos Chem Phys, 15, 617-631, 2015.
Kotchenruther, R. A. and Hobbs, P. V.: Humidification factors of aerosols from biomass burning in Brazil, J Geophys Res-Atmos, 103, 32081-32089, 1998.
Krautstrunk, M. and Giez, A.: The transition from FALCON to HALO era airborne atmospheric research. In: Atmospheric Physics, Springer, 2012.
Krisna, T. C., Wendisch, M., Ehrlich, A., Jäkel, E., Werner, F., Weigel, R., Borrmann, S., Mahnke, C., Pöschl, U., Andreae, M. O., Voigt, C., and Machado, L. A. T.: Comparing airborne and satellite retrievals of cloud optical thickness and particle effective radius using a spectral radiance ratio technique: two case studies for cirrus and deep convective clouds, Atmos. Chem. Phys., 18, 4439-4462, 2018.
Kuhn, U., Ganzeveld, L., Thielmann, A., Dindorf, T., Schebeske, G., Welling, M., Sciare, J., Roberts, G., Meixner, F. X., Kesselmeier, J., Lelieveld, J., Kolle, O., Ciccioli, P., Lloyd, J., Trentmann, J., Artaxo, P., and Andreae, M. O.: Impact of Manaus City on the Amazon Green Ocean atmosphere: ozone production, precursor sensitivity and aerosol load, Atmos Chem Phys, 10, 9251-9282, 2010.
Kulkarni, P. and Wang, J.: New fast integrated mobility spectrometer for real-time measurement of aerosol size distribution - I: Concept and theory, J Aerosol Sci, 37, 1303-1325, 2006a.
Kulkarni, P. and Wang, J.: New fast integrated mobility spectrometer for real-time measurement of aerosol size distribution: II. Design, calibration, and performance characterization, J Aerosol Sci, 37, 1326-1339, 2006b.
Kupc, A., Williamson, C., Wagner, N. L., Richardson, M., and Brock, C. A.: Modification, calibration, and performance of the Ultra-High Sensitivity Aerosol Spectrometer for particle size distribution and volatility measurements during the Atmospheric Tomography Mission (ATom) airborne campaign, Atmos. Meas. Tech., 11, 369-383, 2018.
Lambe, A. T., Onasch, T. B., Massoli, P., Croasdale, D. R., Wright, J. P., Ahern, A. T., Williams, L. R., Worsnop, D. R., Brune, W. H., and Davidovits, P.: Laboratory studies of the chemical composition and cloud condensation nuclei (CCN) activity of secondary organic aerosol (SOA) and oxidized primary organic aerosol (OPOA), Atmos Chem Phys, 11, 8913-8928, 2011.
Lance, S.: Coincidence Errors in a Cloud Droplet Probe (CDP) and a Cloud and Aerosol Spectrometer (CAS), and the Improved Performance of a Modified CDP, J Atmos Ocean Tech, 29, 1532-1541, 2012.

1093 Lance, S., Brock, C. A., Rogers, D., and Gordon, J. A.: Water droplet calibration of the Cloud
1094 Droplet Probe (CDP) and in-flight performance in liquid, ice and mixed-phase clouds during
1095 ARCPAC, Atmospheric Measurement Techniques, 3, 1683-1706, 2010.
1096 Lawson, R.: Effects of ice particles shattering on the 2D-S probe, Atmos Meas Tech, 4, 1361-
1097 1381, 2011.
1098 Lawson, R. P., O'Connor, D., Zmarzly, P., Weaver, K., Baker, B., Mo, Q., and Jonsson, H.: The
1099 2D-S (stereo) probe: Design and preliminary tests of a new airborne, high-speed, high-resolution
1100 particle imaging probe, J Atmos Ocean Tech, 23, 1462-1477, 2006.
1101 Liu, B. Y. and Pui, D. Y.: A submicron aerosol standard and the primary, absolute calibration of
1102 the condensation nuclei counter, Journal of Colloid and Interface Science, 47, 155-171, 1974.
1103 Long;, C. N., Bucholtz;, A., Jonsson;, H., Schmid;, B., Vogelmann;, A., and Wood, J.: A Method
1104 of Correcting for Tilt from Horizontal in Downwelling Shortwave Irradiance Measurements on
1105 Moving Platforms, The Open Atmospheric Science Journal, 4, 78-87, 2010.
1106 Luebke, A. E., Afchine, A., Costa, A., Grooss, J. U., Meyer, J., Rolf, C., Spelten, N., Avallone,
1107 L. M., Baumgardner, D., and Kramer, M.: The origin of midlatitude ice clouds and the resulting
1108 influence on their microphysical properties, Atmos Chem Phys, 16, 5793-5809, 2016.
1109 Martin, S., Artaxo, P., Machado, L., Manzi, A., Souza, R., Schumacher, C., Wang, J., Andreae,
1110 M., Barbosa, H., and Fan, J.: Introduction: observations and modeling of the Green Ocean
1111 Amazon (GoAmazon2014/5), Atmos Chem Phys, 16, 2016a.
1112 Martin, S. T., Artaxo, P., Machado, L., Manzi, A. O., Souza, R. A. F., Schumacher, C., Wang, J.,
1113 Biscaro, T., Brito, J., Calheiros, A., Jardine, K., Medeiros, A., Portela, B., de Sa, S. S., Adachi,
1114 K., Aiken, A. C., Albrecht, R., Alexander, L., Andreae, M. O., Barbosa, M. J., Buseck, P.,
1115 Chand, D., Comstock, J. M., Day, D. A., Dubey, M., Fan, J., Fast, J., Fisch, G., Fortner, E.,
1116 Giangrande, S., Gilles, M., Goldststein, A. H., Guenther, A., Hubbe, J., Jensen, M., Jimenez, J.
1117 L., Keutsch, F. N., Kim, S., Kuang, C., Laskskin, A., McKinney, K., Mei, F., Miller, M.,
1118 Nascimento, R., Pauliquevis, T., Pekour, M., Peres, J., Petaja, T., Pohlker, C., Poschl, U., Rizzo,
1119 L., Schmid, B., Shilling, J. E., Dias, M. A. S., Smith, J. N., Tomlmlinson, J. M., Tota, J., and
1120 Wendisch, M.: The Green Ocean Amazon Experiment (Goamazon2014/5) Observes Pollution
1121 Affecting Gases, Aerosols, Clouds, and Rainfall over the Rain Forest, B Am Meteorol Soc, 98,
1122 981-997, 2017.
1123 Martin, S. T., Artaxo, P., Machado, L. A. T., Manzi, A. O., Souza, R. A. F., Schumacher, C.,
1124 Wang, J., Andreae, M. O., Barbosa, H. M. J., Fan, J., Fisch, G., Goldstein, A. H., Guenther, A.,
1125 Jimenez, J. L., Poschl, U., Dias, M. A. S., Smith, J. N., and Wendisch, M.: Introduction:
1126 Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5), Atmos Chem
1127 Phys, 16, 4785-4797, 2016b.
1128 Mei, F., Hayes, P. L., Ortega, A., Taylor, J. W., Allan, J. D., Gilman, J., Kuster, W., de Gouw, J.,
1129 Jimenez, J. L., and Wang, J.: Droplet activation properties of organic aerosols observed at an
1130 urban site during CalNex-LA, J Geophys Res-Atmos, 118, 2903-2917, 2013a.
1131 Mei, F., Setyan, A., Zhang, Q., and Wang, J.: CCN activity of organic aerosols observed
1132 downwind of urban emissions during CARES, Atmos Chem Phys, 13, 12155-12169, 2013b.
1133 Middlebrook, A. M., Bahreini, R., Jimenez, J. L., and Canagaratna, M. R.: Evaluation of
1134 composition-dependent collection efficiencies for the aerodyne aerosol mass spectrometer using
1135 field data, Aerosol Sci Tech, 46, 258-271, 2012.
1136 Minikin, A., Sauer, D., Ibrahim, A., Franke, H., Röschenthaler, T., Fütterer, D. A., and Petzold,
1137 A.: The HALO Submicrometer Aerosol Inlet (HASI): Design concept and first characterization,

1st HALO symposium: Airborne Research with HALO: Achievements and Prospects, Oberpfaffenhofen, Deutschland,, 2017. 2017.
Molleker, S., Borrmann, S., Schlager, H., Luo, B., Frey, W., Klingebiel, M., Weigel, R., Ebert, M., Mitev, V., Matthey, R., Woiwode, W., Oelhaf, H., Dornbrack, A., Stratmann, G., Grooss, J. U., Gunther, G., Vogel, B., Muller, R., Kramer, M., Meyer, J., and Cairo, F.: Microphysical properties of synoptic-scale polar stratospheric clouds: in situ measurements of unexpectedly large HNO3-containing particles in the Arctic vortex, Atmos Chem Phys, 14, 10785-10801, 2014.
Moran-Zuloaga, D., Ditas, F., Walters, D., Saturno, J., Brito, J., Carbone, S., Chi, X. G., de Angelis, I. H., Baars, H., Godoi, R. H. M., Heese, B., Holanda, B. A., Lavric, J. V., Martin, S. T., Ming, J., Pohlker, M. L., Ruckteschler, N., Su, H., Wang, Y. Q., Wang, Q. Q., Wang, Z. B., Weber, B., Wolff, S., Artaxo, P., Poschl, U., Andreae, M. O., and Pohlker, C.: Long-term study on coarse mode aerosols in the Amazon rain forest with the frequent intrusion of Saharan dust plumes, Atmos Chem Phys, 18, 10055-10088, 2018.
Olfert, J. S., Kulkarni, P., and Wang, J.: Measuring aerosol size distributions with the fast integrated mobility spectrometer, J Aerosol Sci, 39, 940-956, 2008.
Petzold, A., Marsh, R., Johnson, M., Miller, M., Sevcenco, Y., Delhaye, D., Ibrahim, A., Williams, P., Bauer, H., Crayford, A., Bachalo, W. D., and Raper, D.: Evaluation of Methods for Measuring Particulate Matter Emissions from Gas Turbines, Environ Sci Technol, 45, 3562-3568, 2011.
Pöhlker, M. L., Ditas, F., Saturno, J., Klimach, T., de Angelis, I. H., Araujo, A. C., Brito, J., Carbone, S., Cheng, Y. F., Chi, X. G., Ditz, R., Gunthe, S. S., Holanda, B. A., Kandler, K., Kesselmeier, J., Konemann, T., Kruger, O. O., Lavric, J. V., Martin, S. T., Mikhailov, E., Moran-Zuloaga, D., Rizzo, L. V., Rose, D., Su, H., Thalman, R., Walter, D., Wang, J., Wolff, S., Barbosa, H. M. J., Artaxo, P., Andreae, M. O., Pöschl, U., and Pöhlker, C.: Long-term observations of cloud condensation nuclei over the Amazon rain forest - Part 2: Variability and characteristics of biomass burning, long-range transport, and pristine rain forest aerosols, Atmos Chem Phys, 18, 10289-10331, 2018.
Pöhlker, M. L., Pöhlker, C., Ditas, F., Klimach, T., de Angelis, I. H., Araujo, A., Brito, J., Carbone, S., Cheng, Y. F., Chi, X. G., Ditz, R., Gunthe, S. S., Kesselmeier, J., Konemann, T., Lavric, J. V., Martin, S. T., Mikhailov, E., Moran-Zuloaga, D., Rose, D., Saturno, J., Su, H., Thalman, R., Walter, D., Wang, J., Wolff, S., Barbosa, H. M. J., Artaxo, P., Andreae, M. O., and Poschl, U.: Long-term observations of cloud condensation nuclei in the Amazon rain forest - Part 1: Aerosol size distribution, hygroscopicity, and new model parametrizations for CCN prediction, Atmos Chem Phys, 16, 15709-15740, 2016.
Pöschl, U., Martin, S., Sinha, B., Chen, Q., Gunthe, S., Huffman, J., Borrmann, S., Farmer, D., Garland, R., and Helas, G.: Rainforest aerosols as biogenic nuclei of clouds and precipitation in the Amazon, Science, 329, 1513-1516, 2010.
Poschl, U., Martin, S. T., Sinha, B., Chen, Q., Gunthe, S. S., Huffman, J. A., Borrmann, S., Farmer, D. K., Garland, R. M., Helas, G., Jimenez, J. L., King, S. M., Manzi, A., Mikhailov, E., Pauliquevis, T., Petters, M. D., Prenni, A. J., Roldin, P., Rose, D., Schneider, J., Su, H., Zorn, S. R., Artaxo, P., and Andreae, M. O.: Rainforest aerosols as biogenic nuclei of clouds and precipitation in the Amazon, Science, 329, 1513-1516, 2010.
Rissler, J., Swietlicki, E., Zhou, J., Roberts, G., Andreae, M. O., Gatti, L., and Artaxo, P.: Physical properties of the sub-micrometer aerosol over the Amazon rain forest during the wet-to-

dry season transition-comparison of modeled and measured CCN concentrations, Atmos Chem Phys, 4, 2119-2143, 2004.
Roberts, G. C., Andreae, M. O., Zhou, J., and Artaxo, P.: Cloud condensation nuclei in the Amazon Basin:"Marine" conditions over a continent?, Geophys Res Lett, 28, 2807-2810, 2001.
Roberts, G. C., Artaxo, P., Zhou, J., Swietlicki, E., and Andreae, M. O.: Sensitivity of CCN spectra on chemical and physical properties of aerosol: A case study from the Amazon Basin, Journal of Geophysical Research: Atmospheres, 107, LBA 37-31-LBA 37-18, 2002.
Rose, D., Gunthe, S., Mikhailov, E., Frank, G., Dusek, U., Andreae, M. O., and Pöschl, U.: Calibration and measurement uncertainties of a continuous-flow cloud condensation nuclei counter (DMT-CCNC): CCN activation of ammonium sulfate and sodium chloride aerosol particles in theory and experiment, Atmos Chem Phys, 8, 1153-1179, 2008.
Salati, E. and Vose, P. B.: Amazon basin: a system in equilibrium, Science, 225, 129-138, 1984.
Schmid, B., Tomlinson, J. M., Hubbe, J. M., Comstock, J. M., Mei, F., Chand, D., Pekour, M. S., Kluzek, C. D., Andrews, E., Biraud, S. C., and Mcfarquhar, G. M.: The Doe Arm Aerial Facility, B Am Meteorol Soc, 95, 723-+, 2014.
Schulz, C., Schneider, J., Amorim Holanda, B., Appel, O., Costa, A., de Sá, S. S., Dreiling, V., Fütterer, D., Jurkat-Witschas, T., Klimach, T., Knote, C., Krämer, M., Martin, S. T., Mertes, S., Pöhlker, M. L., Sauer, D., Voigt, C., Walser, A., Weinzierl, B., Ziereis, H., Zöger, M., Andreae, M. O., Artaxo, P., Machado, L. A. T., Pöschl, U., Wendisch, M., and Borrmann, S.: Aircraft-based observations of isoprene-epoxydiol-derived secondary organic aerosol (IEPOX-SOA) in the tropical upper troposphere over the Amazon region, Atmos. Chem. Phys., 18, 14979-15001, 2018.
Shilling, J. E., Pekour, M. S., Fortner, E. C., Artaxo, P., Sá, S. d., Hubbe, J. M., Longo, K. M., Machado, L. A., Martin, S. T., and Springston, S. R.: Aircraft observations of the chemical composition and aging of aerosol in the Manaus urban plume during GoAmazon 2014/5, Atmos Chem Phys, 18, 10773-10797, 2018.
Shilling, J. E., Zaveri, R. A., Fast, J. D., Kleinman, L., Alexander, M., Canagaratna, M. R., Fortner, E., Hubbe, J. M., Jayne, J. T., and Sedlacek, A.: Enhanced SOA formation from mixed anthropogenic and biogenic emissions during the CARES campaign, Atmos Chem Phys, 13, 2091-2113, 2013.
Thalman, R., de Sa, S. S., Palm, B. B., Barbosa, H. M. J., Pohlker, M. L., Alexander, M. L., Brito, J., Carbone, S., Castillo, P., Day, D. A., Kuang, C. G., Manzi, A., Ng, N. L., Sedlacek, A. J., Souza, R., Springston, S., Watson, T., Pohlker, C., Poschl, U., Andreae, M. O., Artaxo, P., Jimenez, J. L., Martin, S. T., and Wang, J.: CCN activity and organic hygroscopicity of aerosols downwind of an urban region in central Amazonia: seasonal and diel variations and impact of anthropogenic emissions, Atmos Chem Phys, 17, 11779-11801, 2017.
Wang, J.: A fast integrated mobility spectrometer with wide dynamic size range: Theoretical analysis and numerical simulation, J Aerosol Sci, 40, 890-906, 2009.
Wang, J., Krejci, R., Giangrande, S., Kuang, C., Barbosa, H. M., Brito, J., Carbone, S., Chi, X., Comstock, J., Ditas, F., Lavric, J., Manninen, H. E., Mei, F., Moran-Zuloaga, D., Pohlker, C., Pohlker, M. L., Saturno, J., Schmid, B., Souza, R. A., Springston, S. R., Tomlinson, J. M., Toto, T., Walter, D., Wimmer, D., Smith, J. N., Kulmala, M., Machado, L. A., Artaxo, P., Andreae, M. O., Petaja, T., and Martin, S. T.: Amazon boundary layer aerosol concentration sustained by vertical transport during rainfall, Nature, 539, 416-419, 2016.

Wang, J., Lee, Y.-N., Daum, P. H., Jayne, J., and Alexander, M.: Effects of aerosol organics on cloud condensation nucleus (CCN) concentration and first indirect aerosol effect, Atmos Chem Phys, 8, 6325-6339, 2008.
Webster, C. and Freudinger, L.: Interagency Working Group for Airborne Data and Telecommunications, 2018.
Wendisch, M. and Brenguier, J.-L.: Airborne measurements for environmental research: methods and instruments, John Wiley & Sons, 2013.
Wendisch, M., Coe, H., Baumgardner, D., Brenguier, J.-L., Dreiling, V., Fiebig, M., Formenti, P., Hermann, M., Krämer, M., Levin, Z., Maser, R., Mathieu, E., Nacass, P., Noone, K., Osborne, S., Schneider, J., Schütz, L., Schwarzenböck, A., Stratmann, F., and Wilson, J. C.: Aircraft Particle Inlets: State-of-the-Art and Future Needs, B Am Meteorol Soc, 85, 89-92, 2004.
Wendisch, M., Keil, A., and Korolev, A. V.: FSSP characterization with monodisperse water droplets, Journal of Atmospheric and Oceanic Technology, 13, 1152-1165, 1996.
Wendisch, M., Müller, D., Schell, D., and Heintzenberg, J.: An airborne spectral albedometer with active horizontal stabilization, J Atmos Ocean Tech, 18, 1856-1866, 2001.
Wendisch, M., Poschl, U., Andreae, M. O., Machado, L. A. T., Albrecht, R., Schlager, H., Rosenfeld, D., Martin, S. T., Abdelmomonem, A., Afchine, A., Araujo, A. C., Artaxo, P., Aufmhoff, H., Barbosa, H. M. J., Borrmann, S., Braga, R., Buchholz, B., Cecchini, M. A., Costa, A., Curtius, J., Dollner, M., Dorf, M., Dreiling, V., Ebert, V., Ehrlich, A., Ewald, F., Fisch, G., Fix, A., Frank, F., Futterer, D., Heckl, C., Heidelberg, F., Huneke, T., Jakel, E., Jarvinen, E., Jurkat, T., Kanter, S., Kastner, U., Kenntner, M., Kesselmeier, J., Klimach, T., Knecht, M., Kohl, R., Kolling, T., Kramer, M., Kruger, M., Krisna, T. C., Lavric, J. V., Longo, K., Mahnke, C., Manzi, A. O., Mayer, B., Mertes, S., Minikin, A., Molleker, S., Munch, S., Nillius, B., Pfeilsticker, K., Pohlker, C., Roiger, A., Rose, D., Rosenowow, D., Sauer, D., Schnaiter, M., Schneider, J., Schulz, C., de Souza, R. A. F., Spanu, A., Stock, P., Vila, D., Voigt, C., Walser, A., Walter, D., Weigel, R., Weinzierl, B., Werner, F., Yamasoe, M. A., Ziereis, H., Zinner, T., and Zoger, M.: ACRIDICON-CHUVA CAMPAIGN Studying Tropical Deep Convective Clouds and Precipitation over Amazonia Using the New German Research Aircraft HALO, B Am Meteorol Soc, 97, 1885-1908, 2016.
Zaveri, R. A., Berkowitz, C. M., Brechtel, F. J., Gilles, M. K., Hubbe, J. M., Jayne, J. T., Kleinman, L. I., Laskin, A., Madronich, S., and Onasch, T. B.: Nighttime chemical evolution of aerosol and trace gases in a power plant plume: Implications for secondary organic nitrate and organosulfate aerosol formation, NO3 radical chemistry, and N2O5 heterogeneous hydrolysis, Journal of Geophysical Research: Atmospheres, 115, 2010.

1263

Table 1. List of compared measurements and corresponding instruments deployed aboard the G1 and HALO during GoAmazon2014/5. The acronyms are defined in a table at the end of this paper. Dp indicates the particle diameter. ΔDp refers to the size resolution.

Measurement Variables	Instruments deployed on the G1 (Martin et al., 2016; Schmid et al., 2014)	Instruments deployed on HALO (Wendisch et al., 2016)
Static Pressure	Rosemount (1201F1), 0-1400 hPa	Instrumented nose boom tray (DLR development), 0-1400 hPa
Static air temperature	Rosemount E102AL/510BF -50 to +50 °C	Total Air Temperature (TAT) inlet (Goodrich/Rosemount type 102) with an open wire resistance temperature sensor (PT100), -70 to +50 °C
Dewpoint temperature	Chilled mirror hygrometer 1011B -40 to +50 °C	Derived from the water-vapor mixing ratio, which is measured by a tunable diode laser (TDL) system (DLR development), 5-40000 ppmv
3-D wind	Aircraft Integrated Meteorological Measurement System 20 (AIMMS-20)	Instrumented nose boom tray (DLR development) with an air data probe (Goodrich/Rosemount) 858AJ and high-precision Inertial Reference System (IGI IMU-IIe)
Particle number concentration	CPC, cut off size (D_p) =10 nm	CPC, cut off size (D_p) =10 nm
Size distribution*	UHSAS-A, 60-1000 nm.	UHSAS-A, 60-1000 nm.
	FIMS, 20 nm – 500 nm	
Non-Refractory particle chemical composition	HR-ToF-AMS: Organics, Sulfate, Nitrate, Ammonium, Chloride, 60-1000 nm	C-ToF-AMS: Organics, Sulfate, Nitrate, Ammonium, Chloride, 60-1000 nm
CCN concentration	CCN-200, SS= 0.25, 0.5%	CCN-200, SS= 0.13-0.53%
Gas phase concentration	N2O/CO and Ozone Analyzer, CO, O_3 concentration, precision 2 ppb	N2O/CO and Ozone Analyzer, CO, O_3 concentration, precision 2 ppb
Cloud properties*	CDP, 2-50 μm, ΔD_p=1-2 μm	CCP-CDP, 2.5-46 mm, ΔD_p=1-2 μm
	FCDP, 2-50 μm, ΔD_p=1-2 μm	NIXE-CAS: 0.61 -52.5 μm
	2DS, 10-1000 μm	NIXE-CIPgs, 15-960 μm
		CCP-CIPgs: 15-960 μm
Radiation	SPN1 downward irradiance, 400-2700 nm	SMART Albedometer, downward spectral irradiance, 300-2200 nm

***for an individual flight, the size range may vary.**

Table 2. Summary of the total data points compared between the G1 and HALO instruments.

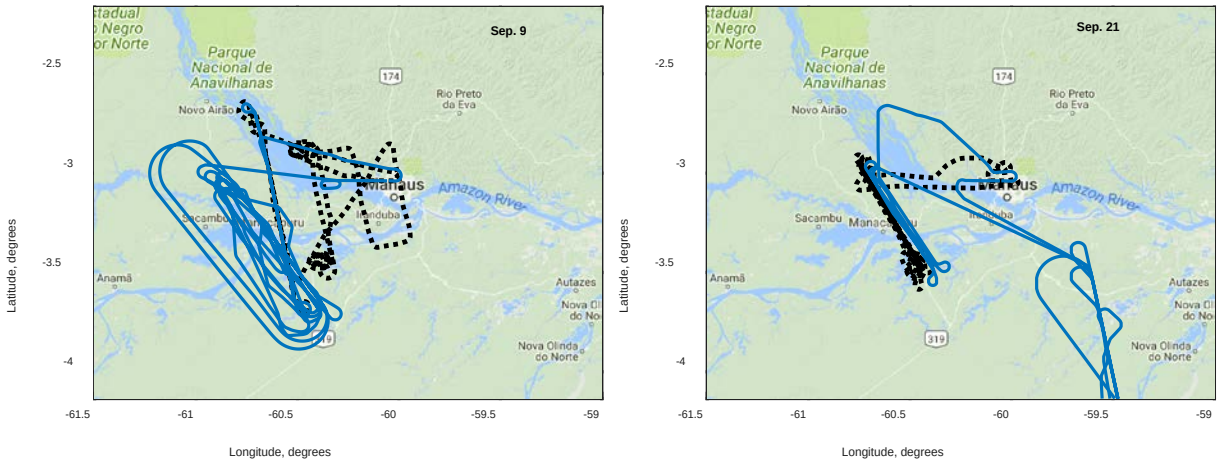
	SEP 9, 2014		SEP 21, 2014	
	G1	HALO	G1	HALO
Atmospheric parameters	2815	2815	7326	12065
Gas phase, CO	N/A	N/A	7326	12065
Gas phase, Ozone	2815	2815	7110	11766
CPC	2043	2043	8466	11646
UHSAS (FIMS)	2031	2031	5841 (9405)	828
AMS	N/A	N/A	587	818
CCNc	663	531	7982	4546
G1: CDP(FCDP) HALO: CCP-CDP (NIXE-CAS)	N/A	N/A	3627(4439)	2051(2260)
G1: 2DS HALO: CCP-CIPgs (NIXE-CIPgs)	N/A	N/A	2280	2261 (2260)
RAD	1355	1355	N/A	N/A

Table 3. Summary of basic statistics of data between in situ measurements on Sep 9.

Comparison of the coordinated flight on Sep. 9										
Variables	G1				HALO					
	min	max	mean	std	min	max	mean	std	slope	R^2
T, K	297.7	300.2	298.9	0.5	297.2	299.4	298.4	0.4	1.002	Neg.
P, hPa	955	965	960.1	1.5	958	964.9	961.8	0.9	0.998	Neg.
WSpd, m/s	0.3	8.9	3.4	1.2	0.3	7.7	3.8	1.1	0.998	Neg.
T_{dew}*, k*	293	296.5	295.0	0.5	292.9	294.9	294.0	0.3	0.996	Neg.
O_3*, ppb*	10.5	58.8	22.2	9.3	18.3	50.8	26.3	6.6	1.082	0.9401
CPC, cm^{-3}	696.0	3480.6	1591.3	568.7	687.4	2639.4	1313.8	473.5	0.819	0.8508
UHSAS, cm^{-3}	78.2	1118.	645.5	116.3	504.1	1622.2	756.3	138.6	1.165	0.8193
CCNc (κ)	0.010	0.347	0.1855	0.067	0.012	0.394	0.1890	0.083	0.8937	Neg.

Table 4 Summary of three statistics analysis of data between in situ measurements on Sep 21

Comparison of the coordinated flight on Sep. 21							
	m	offset	R^2	m0	R^2	m1	R^2
T, K	0.929	20.0	0.9992	0.999	0.9928	0.999	0.9928
P, hPa	1.001	0.929	0.9998	1.001	0.9998	1.001	0.9998
WSpd, m/s	0.885	1.0	0.7875	1.012	0.5076	1.023	0.5049
T_{dew}*, k*	0.989	3.8	0.9963	1.003	0.9904	1.003	0.9904
O_3*, ppb*	1.134	-1.5	0.9598	1.075	0.9369	1.101	0.9208
CO, ppb	0.922	5.4	0.9654	0.966	0.9254	0.967	0.9254
CPC, cm^{-3}	0.571	199.4	0.9482	0.635	0.8738	0.641	0.8735
UHSAS, cm^{-3}	1.126	178.0	0.8249	1.293	0.5070	1.384	0.4847
CCNc (κ)	0.766	55.3	0.8330	0.815	0.6544	0.829	0.6521



(a) (b)

Figure 1. Coordinated flight tracks for September 9 (a) and September 21 (b). The black dotted line is the flight track of the G1, and the blue line is the flight track of HALO.

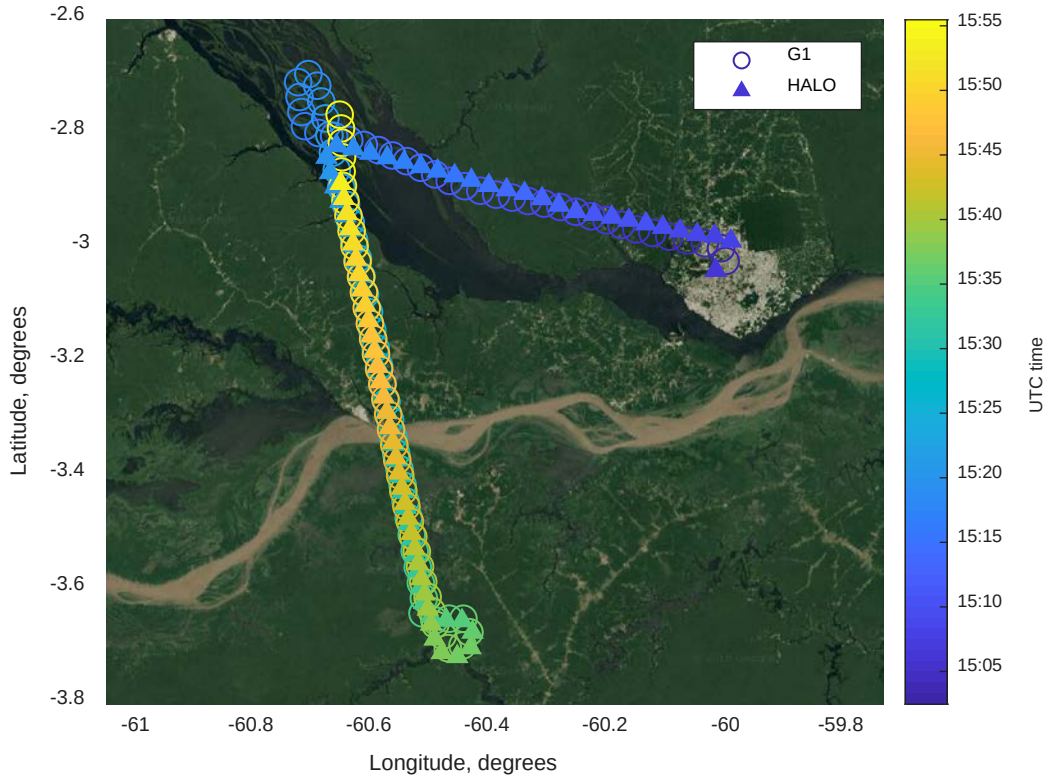
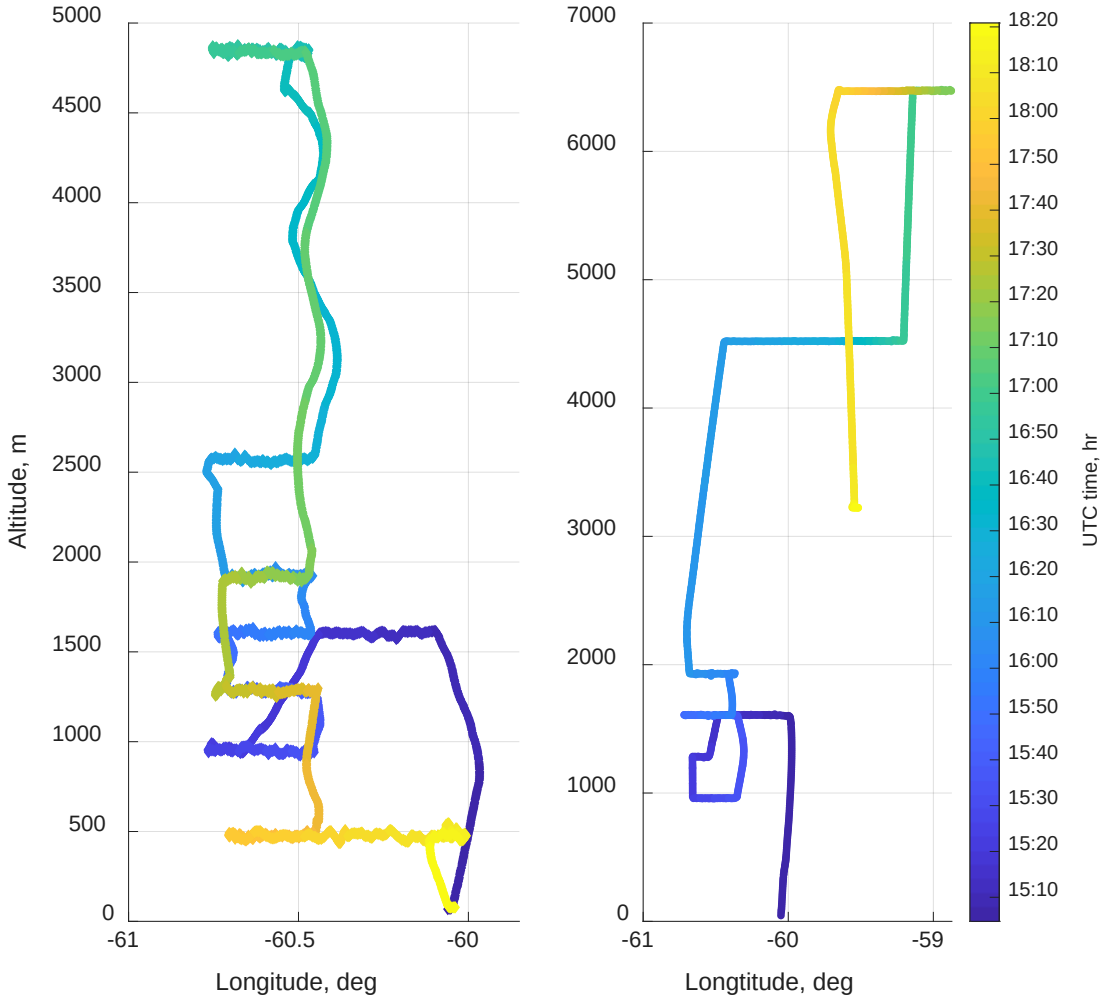


Figure 2. Time-colored flight track of the G1 (circle) and HALO (triangle) on September 9 during a cloud-free coordinated flight at 500 m above sea level (50 m apart as the closest distance).



(a) (b)

Figure 3. Time-colored flight profile of the G1 (a) and HALO (b) on September 21, during a coordinated flight.

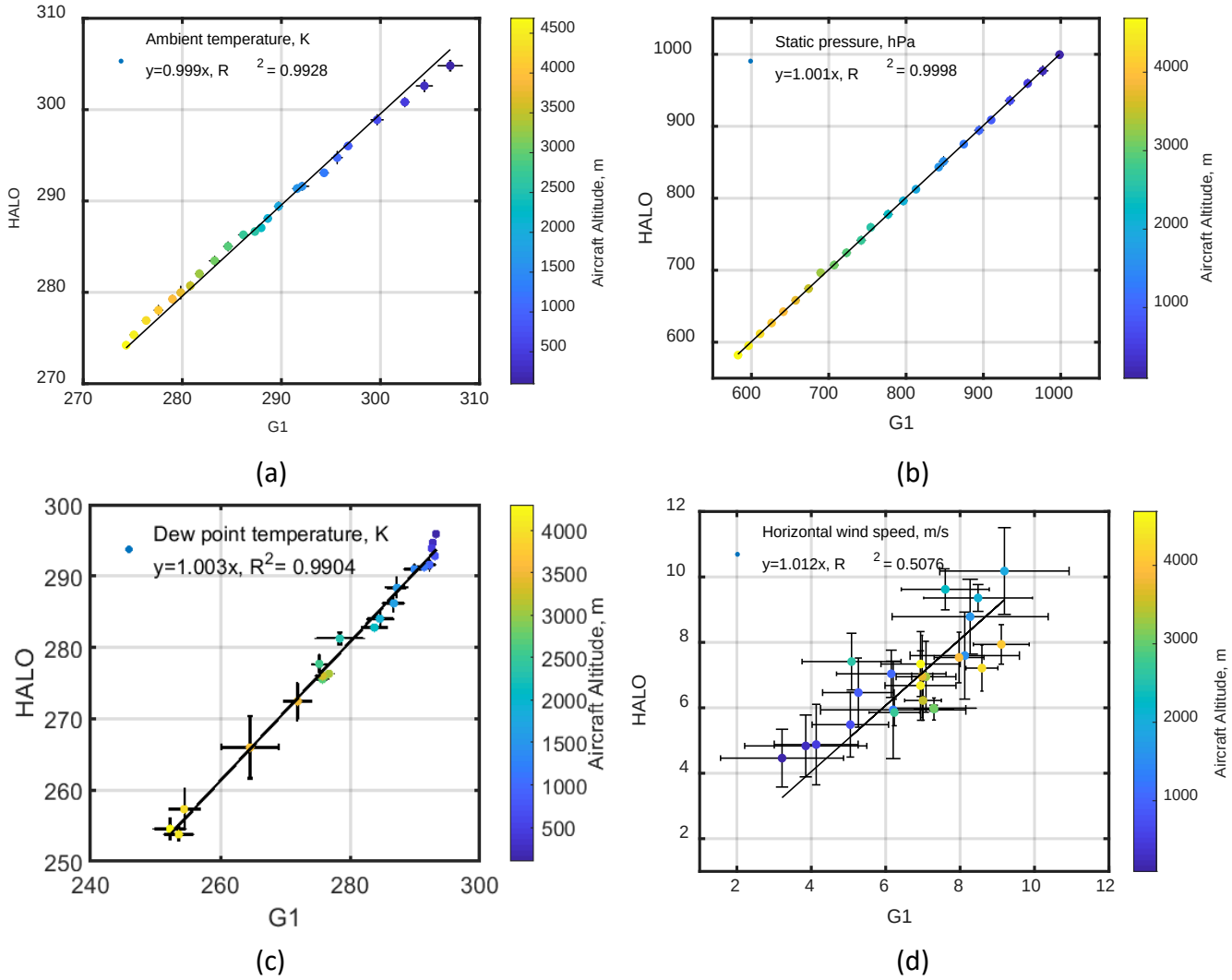


Figure 4. Aircraft altitude-colored plots of (a) ambient temperature, (b) static pressure, (c) dew point temperature, and (d) horizontal wind speed observed by the G1 and HALO on September 21.

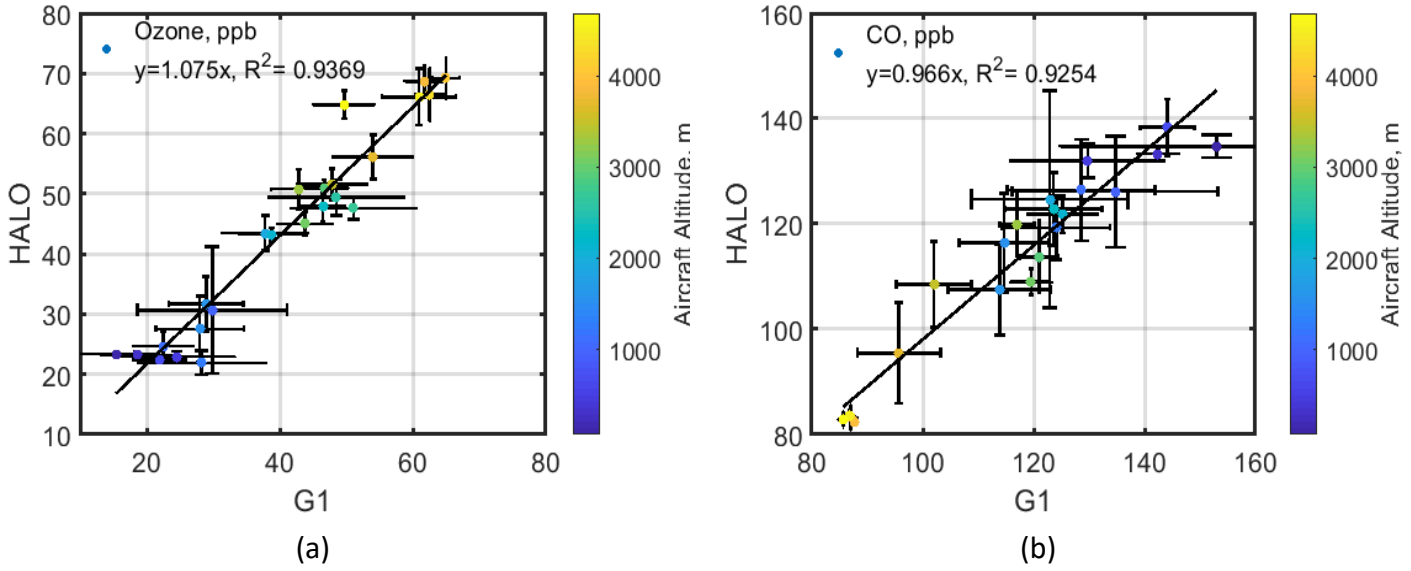


Figure 5. Aircraft altitude-colored plots of trace gas (a) Ozone, (b) CO, for the coordinated flight on September 21.

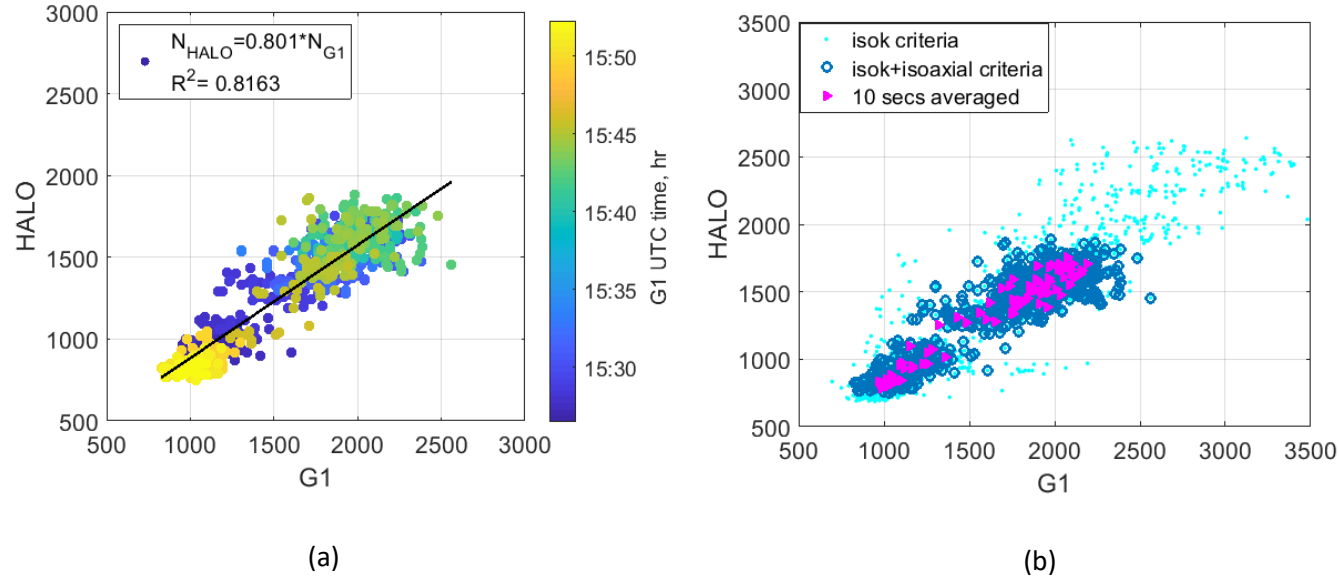


Figure 6. The G1 and HALO comparison for aerosol number concentration measured by CPC (>10 nm) on September 9: (a) with iso-kinetic inlet constrain; (b) with different criteria.

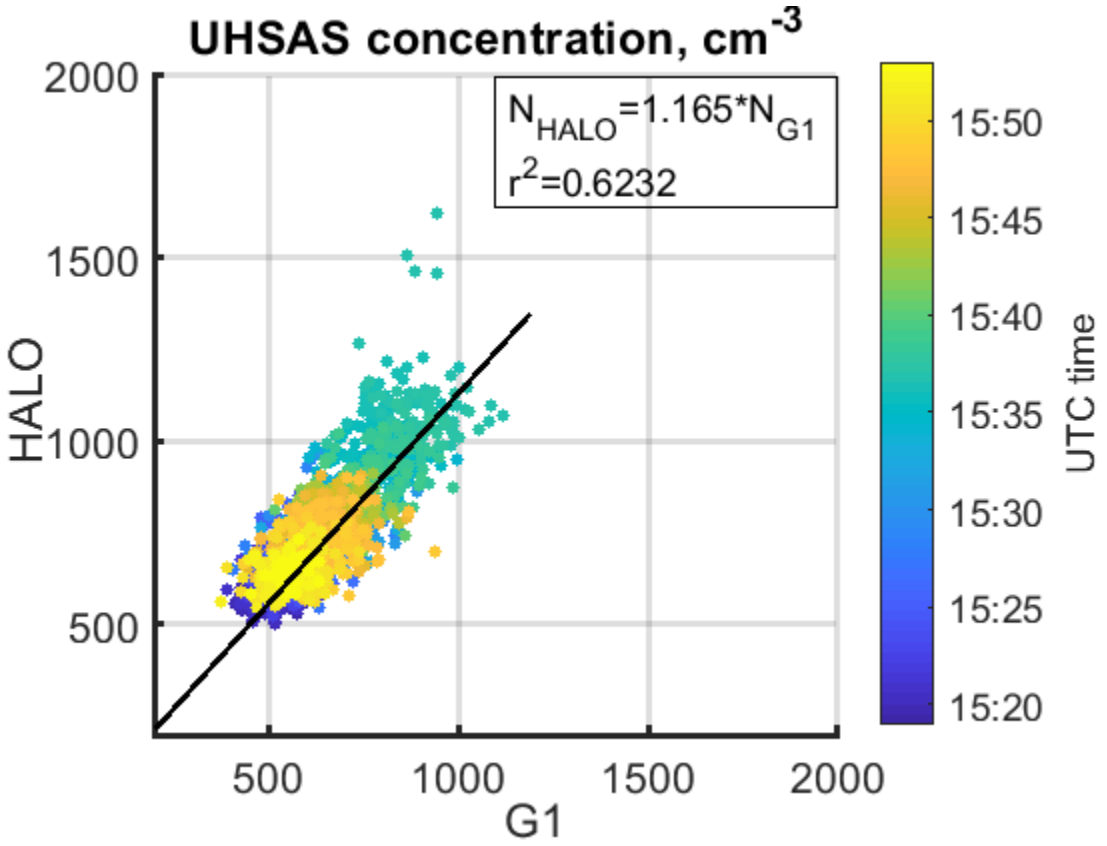


Figure 7. The G1 and HALO comparison for aerosol number concentration measured by UHSAS (90-500 nm) on September 9.

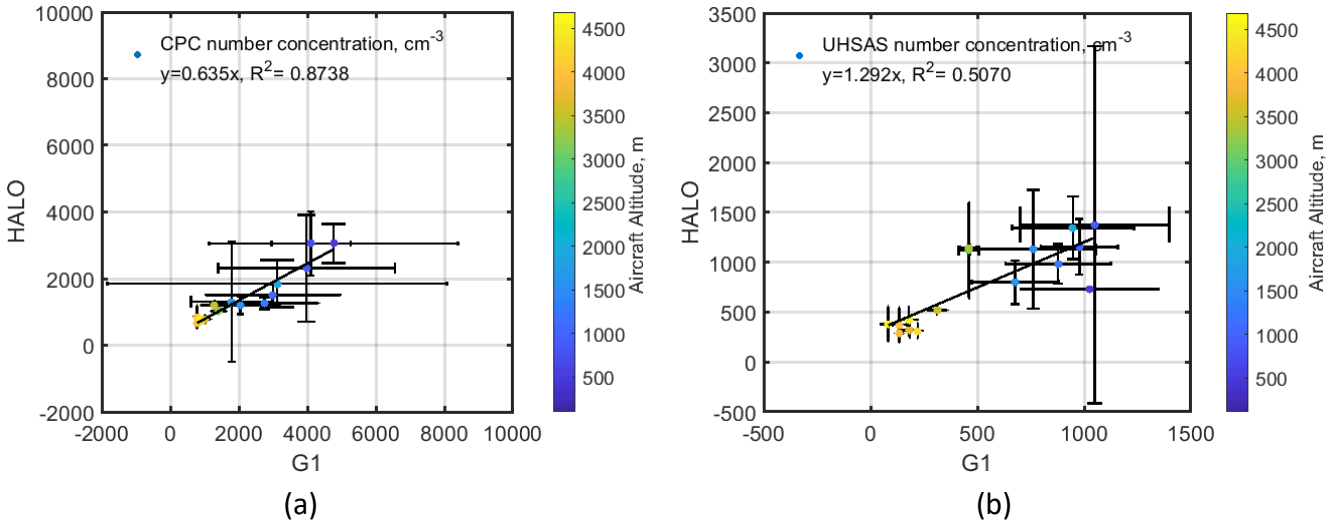


Figure 8. The G1 and HALO comparison for aerosol number concentration profiling measured by (a) CPC and (b) UHSAS (100-700 nm) on September 21.

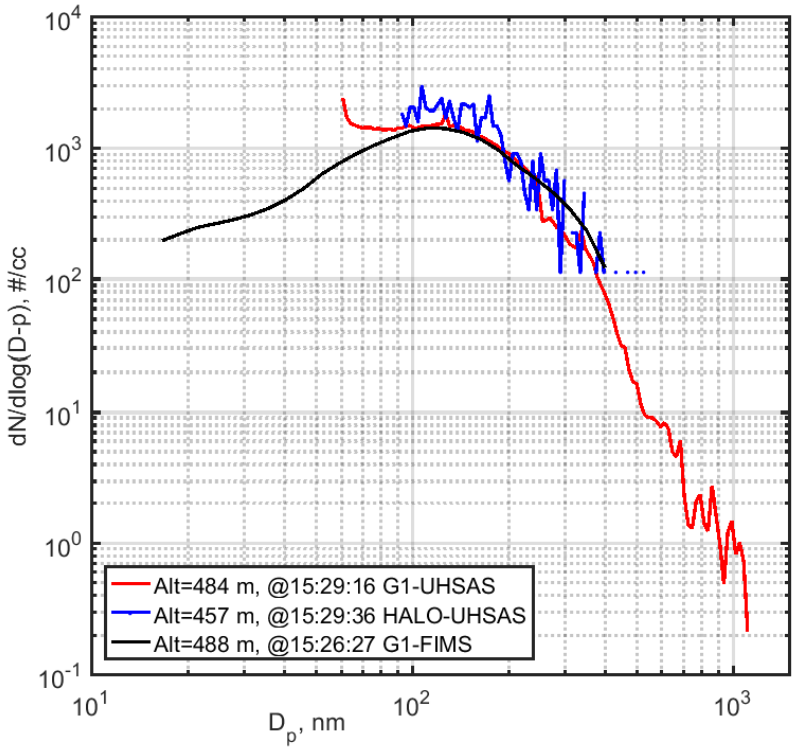


Figure 9. The G1 and HALO comparison for aerosol size distribution measured by UHSAS (from both aircraft) and FIMS (on the G1) on September 9.

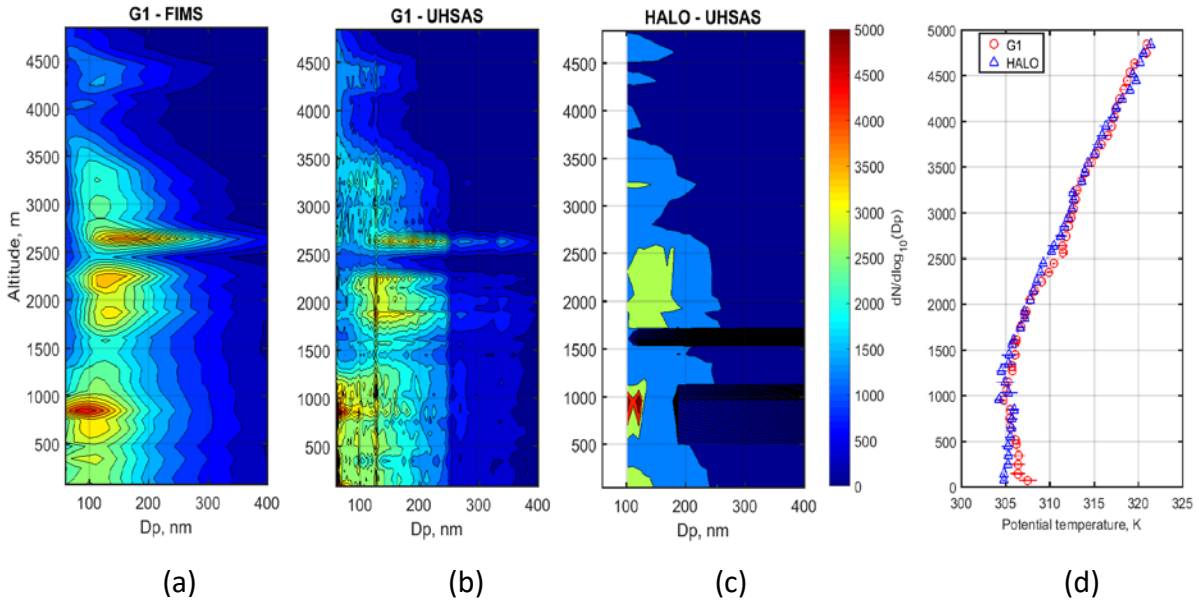


Figure 10. Aerosol size distribution vertical profiles measured by (a) the G1 FIMS, (b) The G1 UHSAS, (c) the HALO UHSAS, (d) Potential temperature aboard the G1 and HALO on September 21.

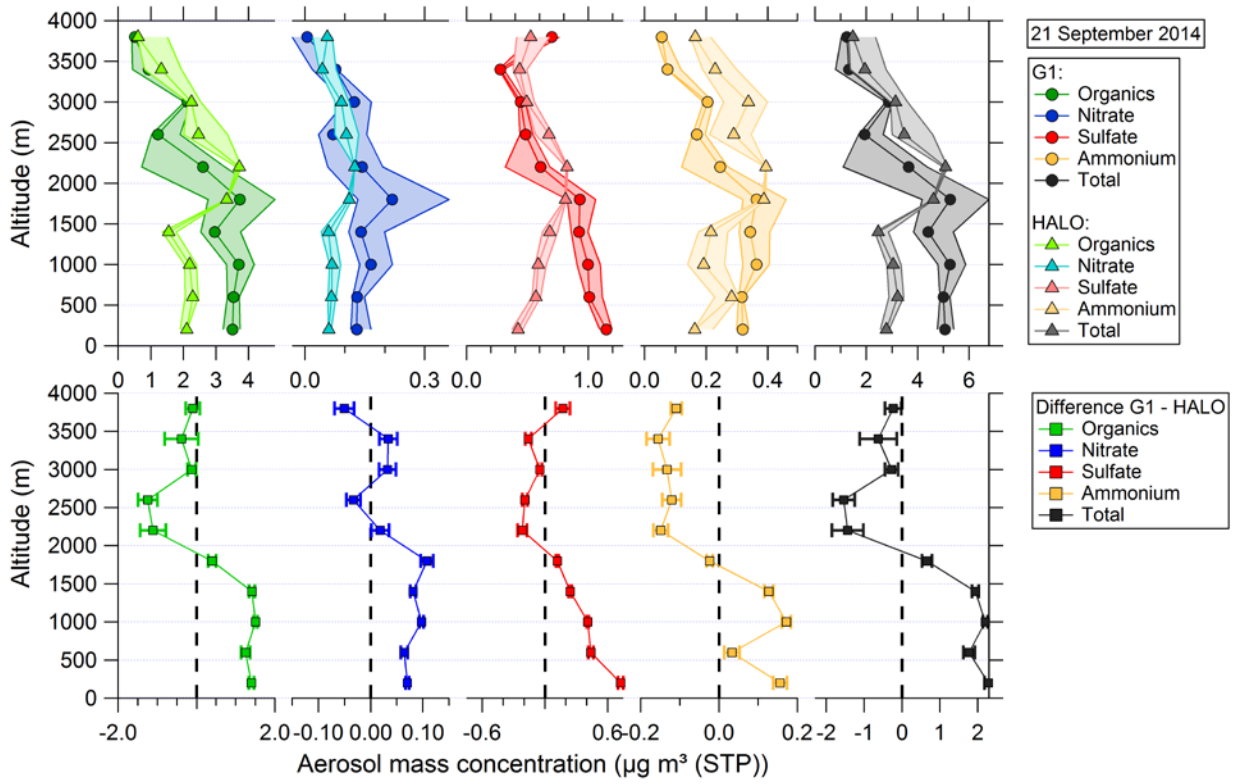


Figure 11. The vertical profile of aerosol mass concentration measured by the G1 and HALO AMS on September 21.

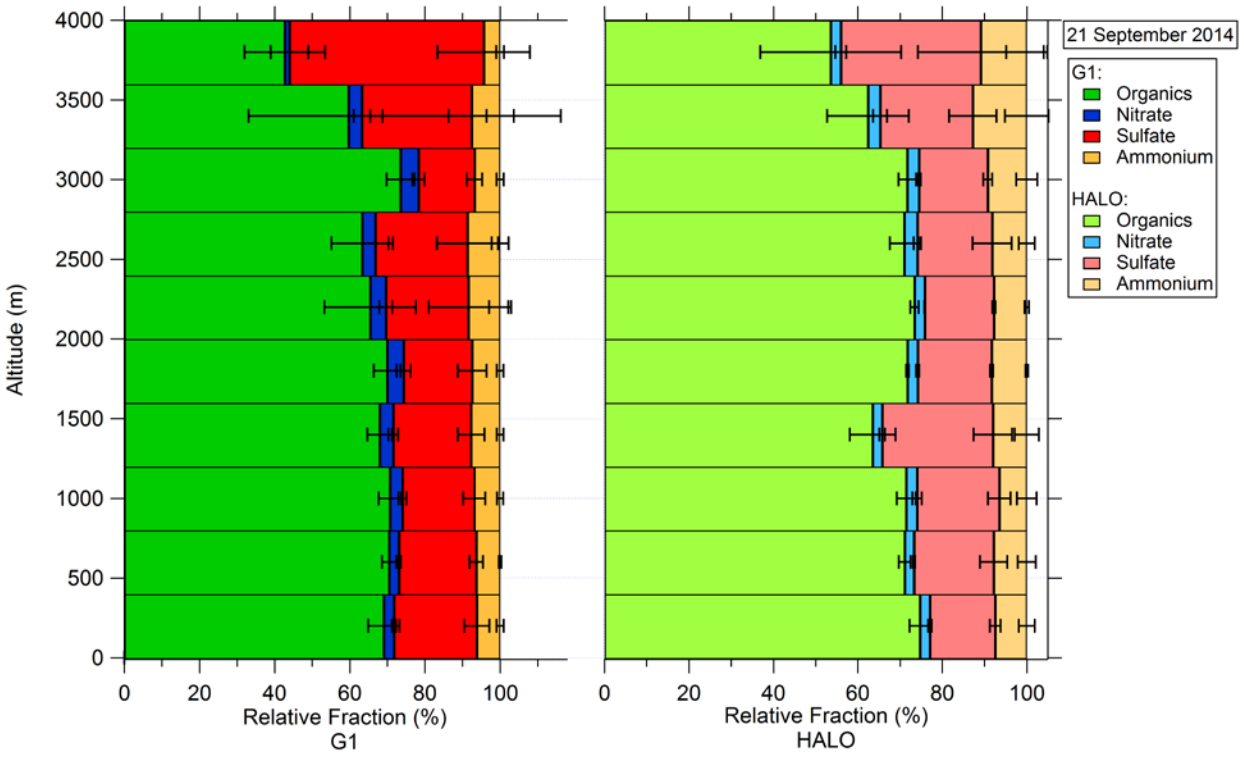


Figure 12. The vertical profile of relative mass fraction of major aerosol chemical species measured by the G1 and HALO AMS, respectively, on September 21

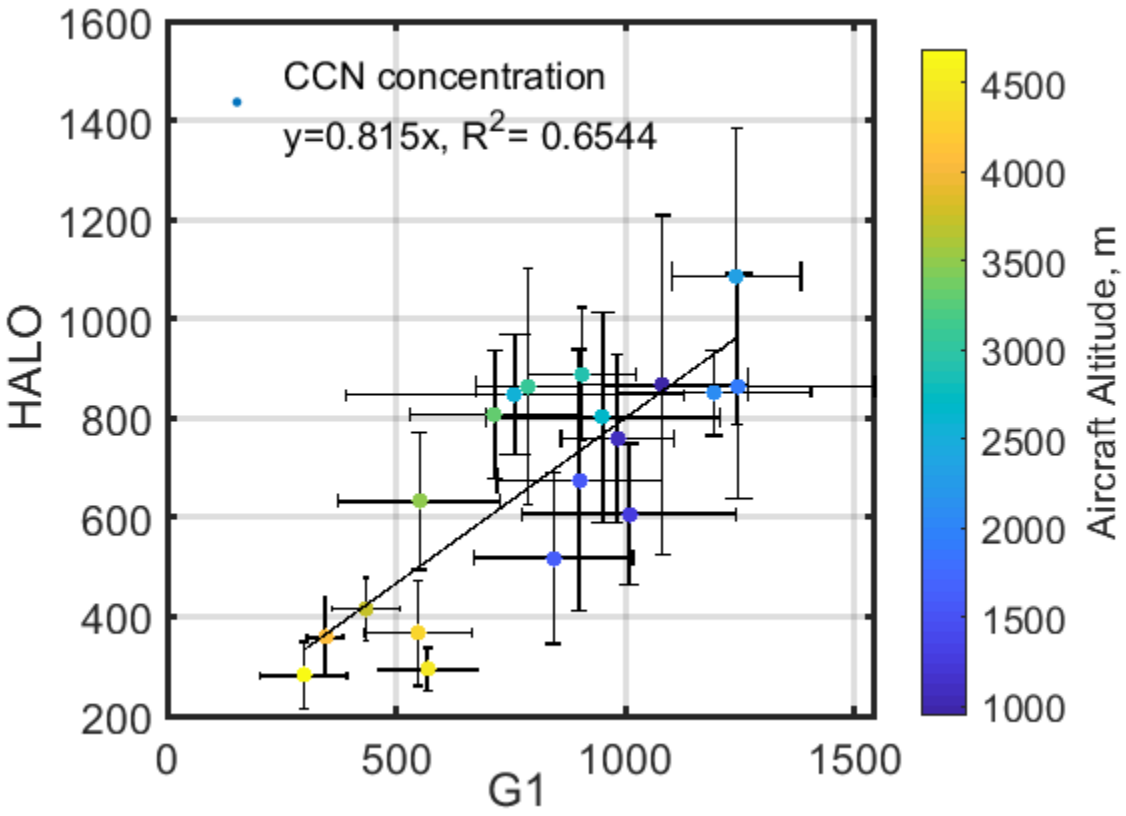
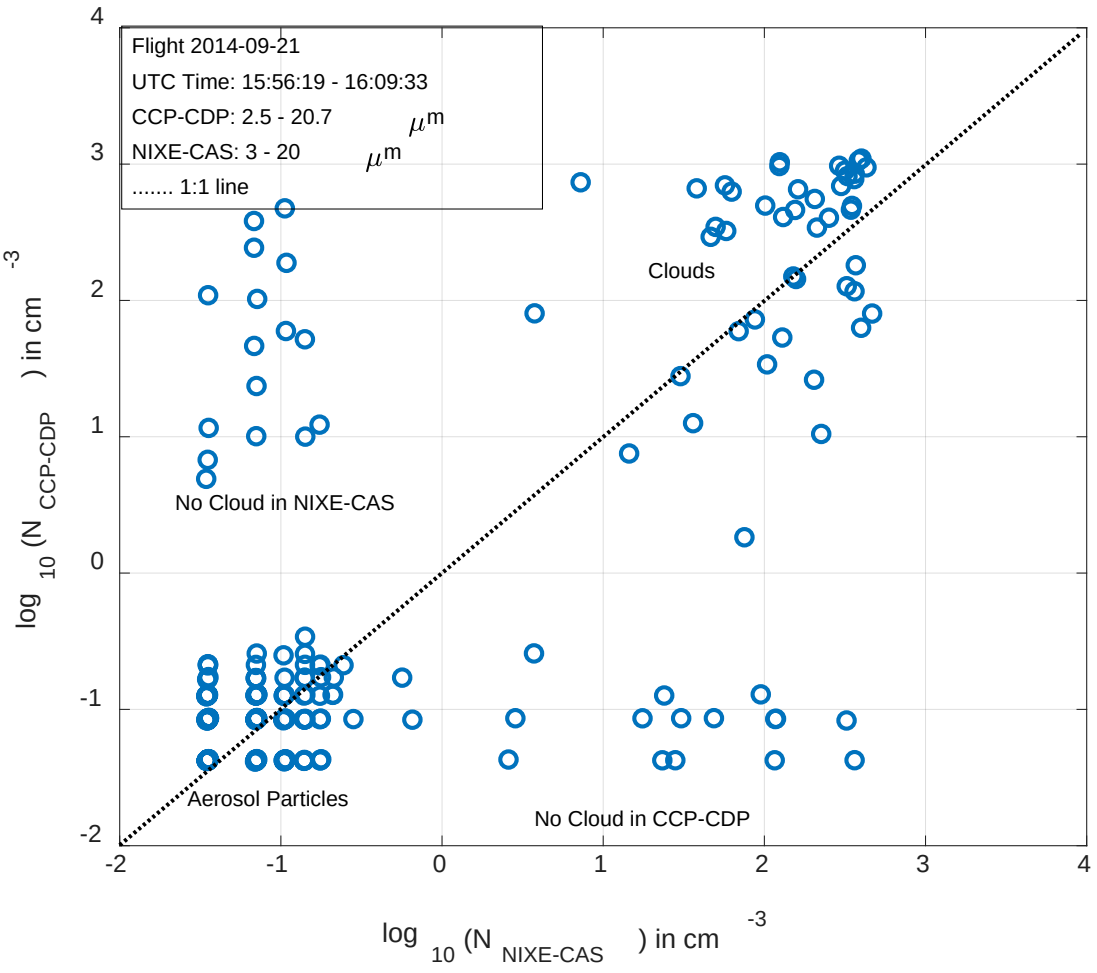
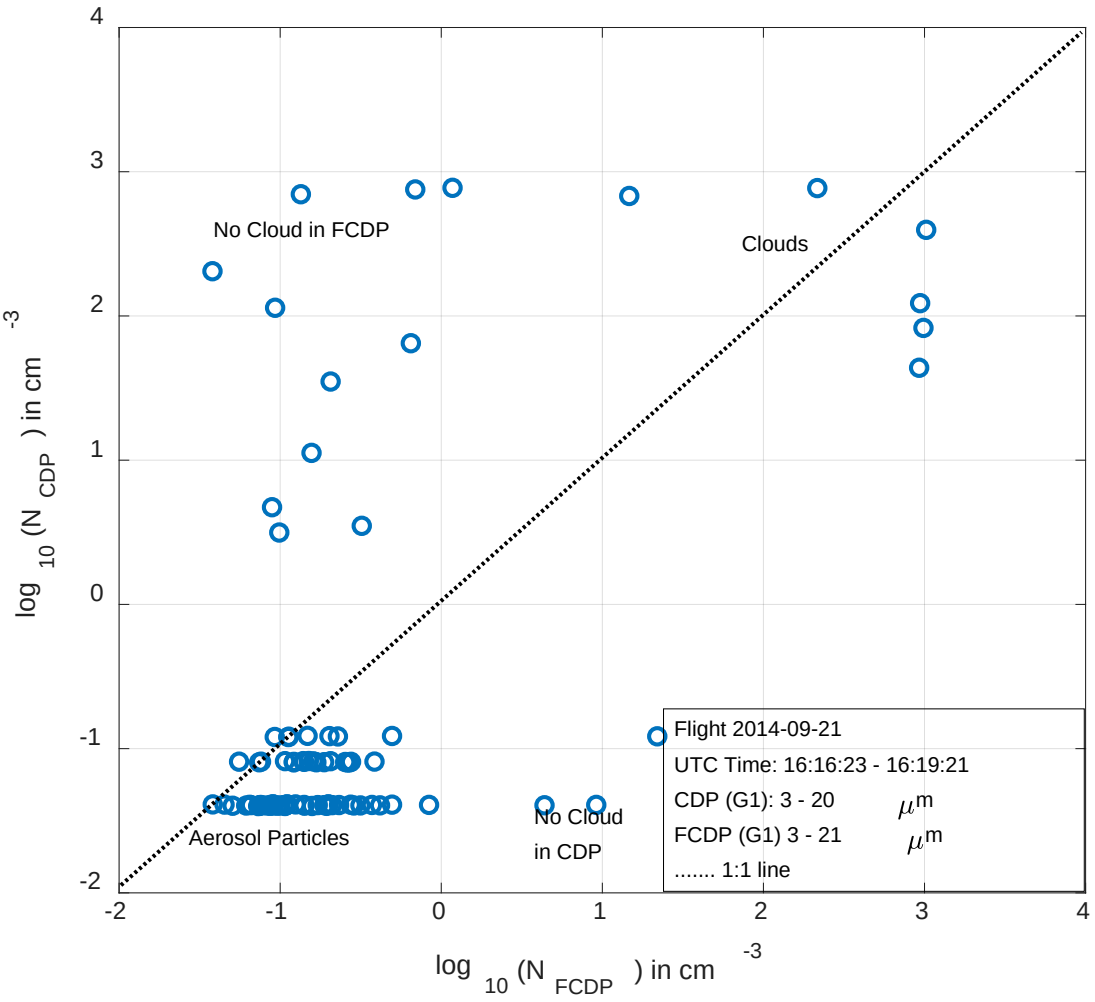


Figure 13. The G1 and HALO comparison of aerosol CCN concentration (S=0.5%) measured on September 21.


Flight 2014-09-21
UTC Time: 15:56:19 - 16:09:33
CCP-CDP: 2.5 - 20.7 μm
NIXE-CAS: 3 - 20 μm
..... 1:1 line
Clouds
No Cloud in NIXE-CAS
Aerosol Particles
No Cloud in CCP-CDP
log 10 (N CCP-CDP) in cm -3
log 10 (N NIXE-CAS) in cm -3
4
3
2
1
0
-1
-2
-2
-1
0
1
2
3
4

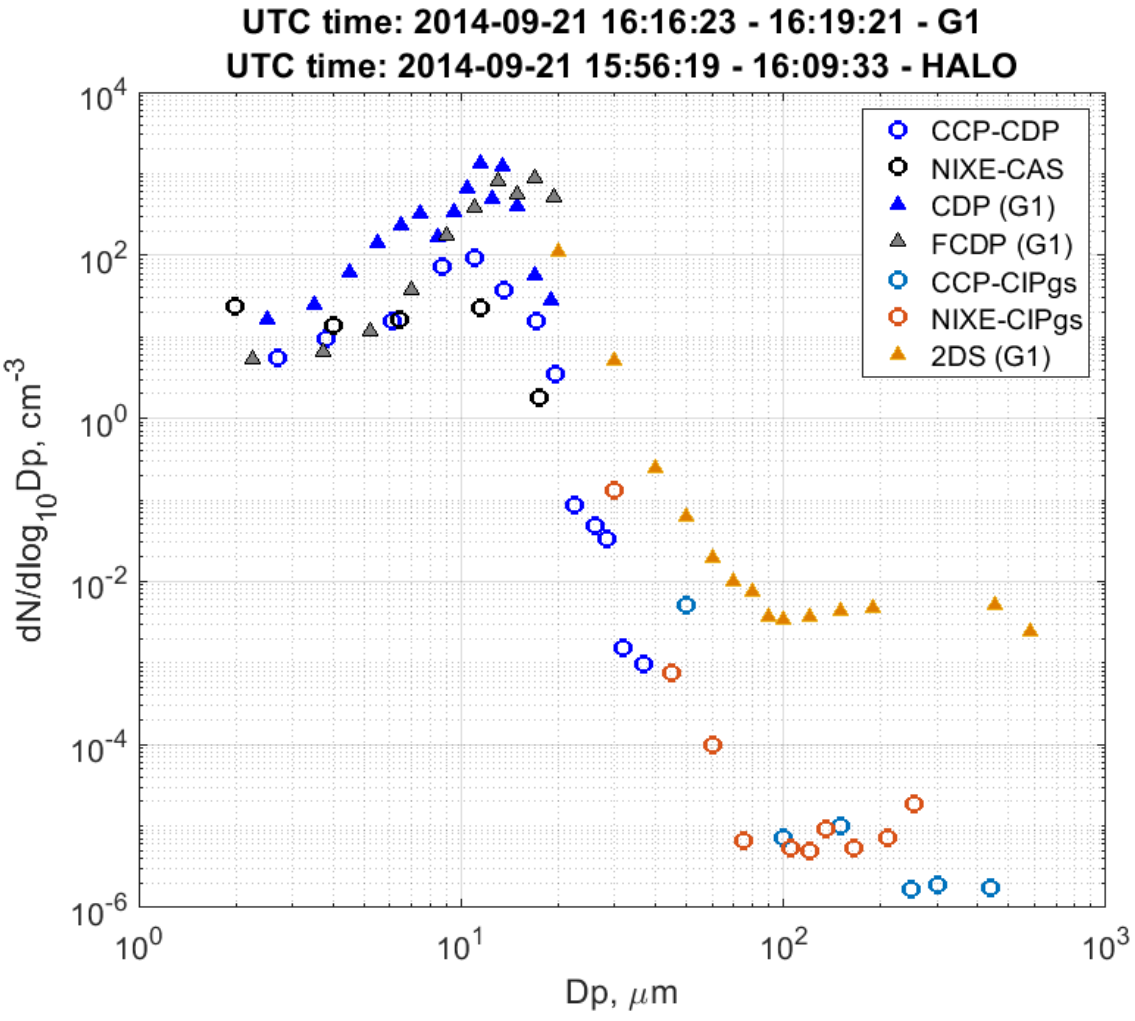
71

72 (a)



(b)

Figure 14 The comparison of cloud droplet concentrations in the same aircraft (a) between NIXE-CAS and CCP-CDP on board HALO; (b) between CDP and FCDP on board the G1.



77

78 Figure 15. The cloud droplet size distribution from the cloud probes on the G1 and HALO.

79

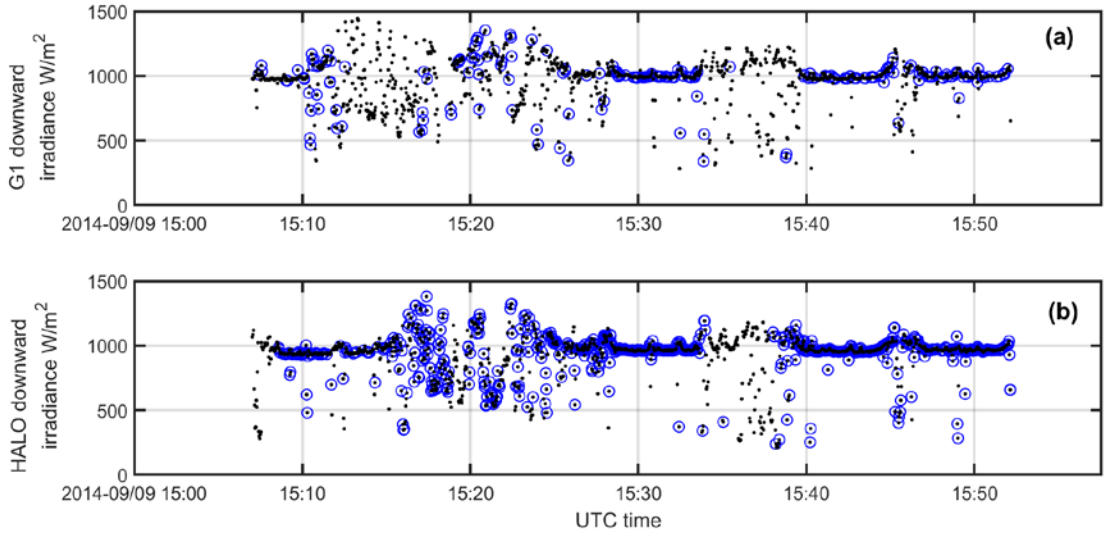


Figure 16. Time series of the G1 and HALO downward irradiance on September 9. The (a) by SPN-1 and (b) by SMART-Albedometer. Black dots represent all data under the general inter-comparison criteria. The blue circles represent the restricted navigation criteria.