

Interactive comment on “Intercomparison of aircraft instruments on board the C-130 and Falcon 20 over southern Germany during EXPORT 2000” by N. Brough et al.

N. Brough et al.

Received and published: 18 November 2003

Anonymous Referee #2 Received and published: 21 September 2003 Review of "Intercomparison of aircraft instruments on board the C-130 and Falcon 20 over southern Germany during EXPORT 2000", Brough et al., Atmos. Chem. Phys. Discuss., 3, 3589-3623, 2003.

Referee comment: Overall a good report of a useful comparison exercise between measurements of NO, NO_y, O₃, and CO between two major European research aircraft. The authors could improve the utility of this report substantially if they would include a brief section on the implications of their comparison results on any subsequent scientific use of the data. For example, the NO_y data are judged to be in agreement within the combined estimated uncertainties; all well and good. However, there

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

remains a substantial offset between the two NO_y data sets, on the order of 90 pptv at 500 pptv ambient NO_y, and a slope difference of ca. 15%. While this is within the combined uncertainties, two additional points could and should be made in this report. 1). The comparison does not fully test the stated accuracy of the DLR instrument, as the UEA uncertainties are a factor of ca. 3 larger. 2). This level of discrepancy in NO_y can have a substantial impact on interpretation of "missing NO_y", especially in light of the recent report of Day et al. ("On alkyl nitrates, O₃, and the missing NO_y", J. Geophys. Res., 108(D16) 4501, doi:10.1029/2003JD003685, 2003).

Response: The uncertainty of the NO and NO_y detectors is controlled by several parameters, such as the artifacts, the uncertainty of the mass flow controllers, the uncertainty of the gas standard, the uncertainty of the conversion efficiency and the variation in the count rate at constant NO/NO_y level. The relative contributions of each of these to the overall inaccuracy depends on the ambient levels of NO and NO_y. At low concentrations the artifact dominates the other uncertainties while at higher concentrations the uncertainty in the determination of the instrument sensitivity is the main contribution. The accuracy of the UEA instrument is acceptable when compared to the reported inaccuracies in similar instruments, For example: Williams et al. (JGR 1997) quote estimated inaccuracies of 19% in NO and 15% in NO_y; Parrish et al. (JGR 1993) quote an inaccuracy of 15% in NO_y measurements made under similar conditions. Stehr et al. (JGR 2000) using similar methods with a molybdenum catalyst report an inaccuracy in their NO_y measurements of 30%, Kondo et al (JGR 1997), also using a molybdenum catalyst report estimated uncertainties of 12% in NO and 27% in NO_y. This information has been added to the paper

Text added to results and discussion: Although the NO_y measurements of the two instruments are in agreement within the uncertainties of the measurements, the systematically higher values of the DLR instrument (90 pptv at 500 pptv) are such that care will have to be taken in using these measurements in the interpretation of missing NO_y especially in light of the recent report of Day et al. (2003)..

[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)[Discussion Paper](#)

Referee comment: Were an NO_y budget to be constructed from both aircraft data sets, despite the noted agreement in NO_y, two completely different conclusions on the magnitude of missing NO_y would be drawn. What would be the recommendations of Brough et al. in this regard?

Response: Without quantitatively knowing the contributions of each NO_y species to the total NO_y measured by the instruments, it is difficult to make quantitative estimates of the amount of missing NO_y. Since the difference in observed NO_y between the two instruments (~90 pptv at 500 pptv) amounts to ~18%, both data sets are within the combined inaccuracy of the two instruments (Tables 1 and 2) and therefore a reasonably accurate NO_y budget calculation can be permitted but the calculated amount of missing NO_y will also have a large uncertainty.

Referee comment: Are the DLR data sufficiently more accurate and precise (see Tables 1 and 2) to support that kind of analysis?

Response: Without a full set of chemical instrumentation on board each aircraft measuring the same air masses during an intercomparison flight, it is difficult to see which instrument is providing the most accurate NO_y budget.

Referee comment: Are the UEA data insufficient in this regard, at the ambient levels encountered in the EXPORT 2000 mission?

Response: No, there is good agreement on the NO_y partition budget with regards to the UEA instrument, considering only the concentrations of those species measured on board the C-130 (NO, NO₂, HNO₃ and PAN) and the differences in NO_y between the two aircraft are within the combined uncertainties of the instruments.

Referee comment: Is neither data set judged to be accurate at the relevant levels to permit an NO_y balance calculation, if the individual NO_y species measurements were available?

Response: To assess the implications of this intercomparison on the missing NO_y

[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)[Discussion Paper](#)

budget is difficult this is only possible by taking into account the inaccuracy of each instrument measuring one species of the NO_y family. Since the difference in observed NO_y between the two instruments (~90 pptv at 500 pptv) amounts to ~18%, both data sets are within the combined inaccuracy of the two instruments calculated from the values quoted in Tables 1 and 2 and therefore can be judged to permit a NO_y balance calculation, however, the uncertainties in the measurements will lead to corresponding uncertainties in the NO_y budget.

Referee comment: The authors should be encouraged to include their assessment of the remaining discrepancies, which could add a very useful dimension of information to the present report.

Text added to summary: The small discrepancies for the missing NO_y could be due to the slight differences in the inlet design and orientation (perpendicular to direction of flight for the NO_x, rear facing for the DLR), which may result in different efficiencies of aerosol sampling by the two instruments leading to the sampling of different amounts of particulate nitrate, although slightly different air mass sampling cannot be totally ruled out.

Specific comments.

Referee comment: Introduction, p.3591, line 22: "To obtain a good instrument comparison it is advisable that the mixing ratios of the compounds to be measured are not consistently near to the detection limits of the instruments, as this can result in erroneous statistical analysis [sic]." True, insofar as the comparison of data is intended to provide information only on instrumental calibrations. But wouldn't an ideal comparison include ambient data throughout the full range of atmospheric relevance, including (at least for NO and NO_y instruments) levels at and below detection limits?

Response: I agree, an ideal comparison would be below and above detection limits but within this data set most of the air encountered was higher than the detection limits of the instruments and comparing them below their detection limits if there was data

[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)[Discussion Paper](#)

would provide larger statistical errors. Text added: §However, it is important to test instrument behaviour through intercomparison exercises at both high and low ambient mixing ratios¶ after sentence quoted above.

Referee comment: Certainly for CO and CO₂ the ambient background is typically well above detection limits throughout the atmosphere, but accurately quantifying a difference between ambient levels of, e.g., 1 and 10 pptv of NO can be critically important for determining the photochemical ozone tendency of an air mass. The above sentence, limiting a comparison exercise to relatively high mixing ratios of a compound, seems to dismiss the importance of accurate knowledge of instrumental zero, or background, levels. For the present report, most of the background tropospheric concentrations of NO were close to both instrumental quantitation limits, so this would seem especially relevant for the data sets in question. Could another sentence be added here discussing the importance of testing instrument behavior through intercomparison exercises at both high and low ambient mixing ratios?

Response: In this data set most of the background measurements obtained for NO (~10 pptv) were above the detection limits of each instrument but to compare instruments below the detection limits will give much larger estimates of error although a clearer picture of the individual instrument performance would be obtained. Text added: §However, it is important to test instrument behaviour through intercomparison exercises at both high and low ambient mixing ratios¶ after sentence quoted above.

Referee comment: Experimental, p. 3593, line 20: "... and there was no verbal in-flight discussion of concentration profiles." Was the rest of this comparison performed blind, as well? For the data presented here, was there any communication during the data reduction process, and were any of the data resubmitted after the initial data evaluation took place?

Response: Intercomparison data was not discussed during the flight and the only changes after the initial data evaluation were to the ozone which was reported in the

[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)[Discussion Paper](#)

paper due to a temperature sensor malfunction.

Referee comment: Experimental, p. 3594, line 7: "The inlet was specifically designed to minimize sample-surface interactions and consequently had a low-volume permitting fast sampling of aerosols up to 4 [micrometers] thus allowing discrimination of nitrate aerosols (Ryerson et al., 1999)." On the contrary, the inlet described by Ryerson et al. (1999) was reported to experience extensive sample-surface interactions: "... all the sampled air contacts an inlet surface, and NO_y instrument response time to HNO₃ is governed by sample-surface interactions." No quantitative description of aerosol transmission was provided, other than the assumption that the majority of aerosol by mass was excluded due to inlet design and orientation. The extent to which the NO_y inlet in the present study was similar to that described by Ryerson et al. is not clear; however, the cited reference does not support the conclusion in the sentence quoted above.

Response: The inlet design used on the NO_x is practically identical to that described by Ryerson et al (1999) except that on the UEA NO_x the inlet is heated to 75°C and temperature controlled. All the sampled air contacts an inlet, and the NO_y response time to HNO₃ is governed by the sample surface interactions. True, but the inlet was heated to 75°C, minimizing the amount of HNO₃ adsorbed on the walls. At a temperature of ca 25 C the transmission of gas phase HNO₃ through the inlet is close to 100% efficient and remains above 90% efficient up to relative humidities of ca 60%. (Ryerson et al., 1999). The higher temperature used would only serve to increase this efficiency (Neuman et al 1999). As the relative humidity of the sampled air was never likely to exceed 70% RH it is reasonable to assume that the HNO₃ was transmitted through the inlet at close to 100% efficiency without significant adsorption to the inlet wall. Text modified to: The inlet and its orientation was specifically designed to minimize NO_y loss and consequently had a low-volume permitting fast sampling (Ryerson et al., 1999).

Referee comment: Experimental, p. 3594, line 23: "The sensitivity of the NO channel was 6.1 +/- 0.9 cps pptv-1 ...". The imprecision given here of $(0.9/6.1) = +/-15%$ in the

derived sensitivity seems rather high for a well-operated instrument. When combined with an uncertainty in the calibration tank gas mixing ratio of $(0.1/1.01) = \pm 10\%$, plus an estimated uncertainty in sample and calibration mass flow controller calibrations of ca. 4% absolute, and even assuming perfectly stable background count rates in the NO instrument, these values seem to be inconsistent with the accuracies for the UEA NO data stated in Table 1 of $\pm 12\%$ at 50 pptv. The scatter in the derived NO sensitivities alone can account for more than the total imprecision given in Table 1.

Response: This is a typographic error, the actual error in the NO sensitivity was 0.09 (0.1) not 0.9. Given an imprecision of 1.5% not 15 % as previously stated. The uncertainty in the calibration gas mixing ratio was also incorrect in the original MS, the actual value is 1% not 10% as quoted originally. The calibration cylinder was obtained from BOC and not Scott Speciality Gases as quoted in the original MS.

Referee comment: Experimental, p. 3595, line 22: "The major contributor to the inaccuracy for each detector was the instrument artifact signal." At low levels of ambient NO, yes; however, an imprecision of $\pm 15\%$ in the NO channel sensitivity derived from the in-flight calibrations suggests that at levels well above the detection limit, the calibration uncertainty dominates the measurement inaccuracy.

Response: This uncertainty is 1.5% not 15% as stated above.

Referee comment: Experimental, p. 3598, line 21: "The O₃ standard source was calibrated against O₃ measurements at the global watch station at HohenPeienberg [sic], Germany." Although unstated, I assume this O₃ standard source was based on UV absorption at 254 nm. If so, this provides a direct measurement of O₃ number density and is subject only to uncertainty in the measured cross section, gas temperature, pressure, and cell path length - it is most emphatically not capable of being calibrated, unless Beer's Law is itself being "calibrated". Please be careful in the wording here; comparison or referencing of two nominally absolute measurements cannot be a calibration.

Response: In deed all mentioned ozone instruments are based on UV absorption and this kind of instrument depends upon the Beer-Lambert Law. But unfortunately most of these ozone instruments used for measurements in the atmosphere are not so accurate that the ozone measurement only depends on the cross section, temperature, pressure and cell length. This is caused by the rather short length of the cells (30-40 cm). At this length the attenuation of UV light is very small and the change in intensity is of the same order as the change in intensity of the UV lamp. Additionally other instrument parameters can influence the accuracy of the measurement. Usually a second cell is used to correct for these changes. Nevertheless these measurements can never be as accurate as measurements using a UV photometer with absorption lengths of more than 2m (as used for the primary standard mentioned above). So, my feeling, although the ozone measurement on the aircraft and the primary standard are based on the same principle, is that the procedure described above is more a calibration than a comparison.

Referee comment: Results and discussion, p. 3601, line 4: "At an indicated aircraft speed of 180 knots this would give plumes [sic] widths of between 0.7 km and 4.3 km." This calculation assumes the aircraft crossed the plumes at 90 degrees to the long axis of the plume. Depending on the intercept angle, an aircraft could take e.g., 48 s to cross a 1-km-wide plume if the flight track vector had a substantial component along the plume, rather than across it. Variations of plume mixing ratio with altitude would also add uncertainty to this calculation.

Response: Text added \S assuming flying at 90 degrees to the plume

Referee comment: Results and discussion, pp. 3601 and following: In the discussions of uncertainty, the observed differences in fitted slopes between the UEA and DLR instruments are described, e.g., as "... well within the overall instrument uncertainty of the UEA NO_y (21% at 450 pptv) and just outside those for the DLR instrument (8% at 450 pptv)." The distinction is meaningless; the applicable metric is whether or not the observed discrepancy in the fitted slope can be encompassed within the combined

[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)[Discussion Paper](#)

uncertainty of the two instruments, presumably obtained by addition in quadrature. The only statistically meaningful conclusion is that the observations are not different from a slope of 1.0 within the combined uncertainty of the two instruments. It is further true that the relatively larger UEA uncertainty does not provide as robust or rigorous a test of the DLR instrument as could be hoped, but that is the nature of the comparison exercise. It follows that conclusions drawn from data obtained using the instrument of lower uncertainty will be defensible to a greater degree, but the comparison suggests that no substantial, unknown sources of error exist in either measurement.

Response: Text changed to: §This 17% difference is well within the 22% combined uncertainty of the two instruments obtained by addition in quadrature. This suggests that no substantial unknown sources of error exist in either measurement.

Referee comment: Results and discussion, p. 3603, line 10: "[A significant increase in CO] was clearly not observed by the DLR instrument and would appear to indicate real variations in ambient CO since good agreement was obtained for the rest of the comparison period." Another possibility that could account equally well for the observed but transient instrumental difference would be short-term instability in one or the other of the CO instruments. Without additional evidence for "real variations in ambient CO", there is no a priori reason to exclude these data just because they differ. If that were the case, every comparison data point that disagreed for whatever reason could be written off, making the comparison exercise meaningless. Are there ancillary data available that further point to this period being different?

Response: The CO data for this strange period was not excluded from the statistical analysis, all data points were used for the intercomparison period. The ancillary data did not help any further to provide evidence for this period of the flight being different. Text added: §...although short term instability in one or other of the instruments cannot be fully ruled out during this period.Ť

Referee comment: The O3 data might help in this regard, but more support for this

[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)[Discussion Paper](#)

assertion must be presented here or it weakens the entire comparison exercise. "Air mass sampling differences due to the spatial arrangement of the aircraft" are mentioned - well, were these spatial arrangements any different for this 3-min period than for the rest of the exercise? Please present some aircraft positional data to support this assertion, and at least mention the possibility of short-term instrument instability as another, if less palatable, explanation for the observed discrepancy.

Response: From the individual aircraft positions there was no spatial difference observed for this period, as far as short term instability, see response to point above.

Referee comment: Summary, p. 3608, line 23: "The degree of agreement lends confidence to the accuracy of all observed measurements and indicates the accuracy to be within 12% and 15% for the NO and NO_y respectively." These uncertainties are the percentage differences of the fitted slopes from unity from the comparison exercises, but they are not the accuracy of the instruments - those remain as given in Tables 1 and 2, of 12% and 21% for the UEA measurements and 8% and 7% for DLR for NO and NO_y, respectively. The comparison exercise can not result in improvements (for UEA, at least) in accuracy of the instrumental data, as the above statement implies; the exercise can only provide an assurance that the uncertainties are appropriately stated. Surely the DLR data quality has not been degraded by a factor of 2 as a result of this exercise? The uncertainties remain as given in Table 2, and this statement is somewhat misleading; better to simply restate the two instruments uncertainties here in the summary.

Response: I agree with this statement. Text changed to: "This 17% difference is well within the 22% combined uncertainty of the two instruments obtained by addition in quadrature. This suggests that no substantial unknown sources of error exist in either measurement."

Technical corrections made as suggested.