Interactive comments on
"Intercomparison of NO_3 radical detection instruments in the atmosphere simulation chamber SAPHIR"
by H.-P. Dorn et al.

MS No.: amt-2012-281

We thank the anonymous Referees for their valuable comments and their support to publish the paper in this journal. Here is our response to all specific comments and questions. Proposed changes in the revised manuscript are shown in blue.

Anonymous Referee #1

Received and published: 21 February 2013

Comment: Page 309, Line 21: Was the number of instruments seven or nine?

Response: Two additional instruments participated in NO3Comp which only measured N2O5 (or more precisely the sum of NO3 and N2O5). We will change the sentence to: Five instruments participating in NO3Comp …

Comment: Reactions R2-R6: remove commas and stops at the end

Response: Removed.

Comment: Were the other CRDS instruments corrected for NO2 offset during calculation of the ring down times similar to the UAF instrument?

Response: 1) As described in sections 2.11 to 2.13 NOAA, UAF, and MPI cavity ring-down instruments used the titration method with NO and all corrected for the additional NO2 absorption resulting from the NO + O3 reaction.
2) ULEIC-BBCRDS (section 2.1.4) provides a separate measure of NO2 concentrations (at least when NO2 is above its detection limit; cf. Fuchs et al 2010), so it does not require correction for NO2's absorption signal.
3) Different NO cylinders were used for titration during the campaign and only one – used by UAF – showed a marginal contamination with NO2. We made this clear by changing this sentence: Moreover it was noted, that the NO titrant cylinder used for the UAF instrument was contaminated with traces of NO2, …

Comment: Page 341? (314 !) line 27: Was the baseline drift quantified?

Response: The MPI instrument could be operated either in CEAS or in CRDS mode. The baseline drift limited the accuracy of the CEAS mode. However, all NO3 data of the intercomparison were exclusively measured in the cavity ring-down mode. (Details of the evaluation of the CEAS measurements can be found in the publication of Schuster et al., 2009.)

Comment: Page 315 line 7: what was the flow rate?

Response: Each of the cavity mirrors was purged with 0.15 slm of bottled, filtered, zero air. The text was extended accordingly: The mirror distance was 0.7 m, sheath flows of filtered zero air (0.15 slm) protected the mirrors.

Comment: Page 318 line 10 – please give more details on the box modeling

Response: As stated in the text (page 318, lines 5-15) no corrections of the measured NO3 mixing ratios were made for NO3/N2O5 re-equilibration, hence we do not describe the model in the text. The calculations showed that the size of the correction was (i) smaller than/comparable to other source of systematic measurement uncertainty (the typical under-measurement was 4%) and (ii) dependent on conditions that themselves could change rapidly during an experiment (NO2, temperature and NO3 concentrations) and hence it was difficult to identify what precisely the correction should be.

In detail: The zero dimensional box model considered how the gas phase composition evolved with time as it was sampled from SAPHIR into the colder ULEIC-BBCRDS instrument. The model included the reactions

between NO, NO2, NO3, N2O5, O3 (R1 to R5 in Section 1) and representative wall loss rates, and was run for various conditions representative of experiments performed in SAPHIR. For any given set of initial conditions, the model's chemistry was propagated forward in time at SAPHIR's temperature. The model was then run again, introducing a step reduction in temperature part way through the run to mimic sampling gas into the colder instrument. We assumed that light circulating inside the cavity samples the evolving gas composition down the cavity's length, between 0 s (ie just sampled into the cavity) and 2.7 s (the residence time inside the instrument). The fractional loss in NO3 due to NO3/N2O5 re-equilibration was computed from the ratio of modelled NO3 concentrations at the SAPHIR and instrument temperatures, averaged over the first 2.7s after temperature change.

Comment: Page 319 line 12: Is it really 61 s?

Response: Yes.

Comment: Page 320 line 9: Mention the cutoff wavelength

Response: We added 630 nm to Page 320 line 9.

Comment: Please give details as to how the L_eff was calculated for all the CRDS instruments.

Response: The ratio R_L of the geometrical cavity length and the length over which the trace gas is present when the cavity mirrors are flushed was determined by adding an absorbing reference gas (usually NO2, O3, or H2O) in different, precisely measured concentrations to the cavity. The ratio is determined from the fitted slope of the linear regression line of a plot of externally measured reference gas concentration vs the concentration measured in the CRDS cavity. The accuracy of R_L depends on the accuracy of the concentration measurement of the reference gas and on the uncertainty of its absorption cross section. The procedure is described by Fuchs et al. 2008. We extended the text on page 311, line 22: The value of *R*L had been determined previously in laboratory experiments to be 1.15±0.03 (Fuchs et al., 2008).

Comment: Page 323 line 6: Should the mirror reflectivity not be measured more regularly to account for any deposition effects? Could this be why the CE-DOAS is less accurate on some days with more aerosols?

Response: The UHD-CEDOAS instrument used an open cavity. It was therefore only possible to measure the mirror reflectivity in the morning of each experiment day after the chamber was flushed over night with dry synthetic air. Contamination of the mirrors was minimized by a continuous flow of dry synthetic air (5 slm). On days with high aerosol load the optical path length is reduced and hence the sensitivity of the CE-DOAS instrument is lower.

Comment: Page 329 – the pressure was slightly over the ambient pressure.

Response: The chamber is held at an overpressure of 40 Pa which corresponds to 0.4 mbar. We changed the text to: All simulation studies were performed under virtually ambient pressure and temperature so that the performance of the instruments was investigated under realistic, near-atmosphere conditions.

Comment: Page 330 line 11 – should be 4 and 2 pptv, respectively.

Response: Changed.

Comment: Page 334 line 6 – why did the sensitivity change?

Response: Sudden changes of the instrumental response under otherwise constant measurement conditions are often a hint for fluctuations of the light source. Probably the sensitivity change (on 14 June) was caused by problems with the temperature regulation of the LED as already observed during the experiments on 10 and 11 June. We adapted the text on page 334, line 7: However, the sensitivity of the UHD-CEDOAS instrument seemed to have changed after 10:00UTC probably due to a temperature change of the LED.
A corresponding remark was also added in Section 4.3, page 348, line 6: The deviation of the UHD-CEDOAS instrument can be explained partly by non-linear saturation effects of some water vapour absorption lines which were not treated properly in the retrieval process and partly due to insufficient temperature stabilisation of the LED on this day.

Comment: Page 335 line 4 – why was the arc so unstable – this is an easily solvable problem with DOAS instruments?

Response: Unfortunately, we do not have an explanation. We assume a problem during the lamp production because all lamps used during NO3Comp came from the same new batch of Xe lamps.

Comment: Page 335 line 24 – why did the differences increase? I couldn't find it in the paper.

Response: The explanation – assumed loss of NO3 on aerosol deposited on the walls of inlet lines and/or on the filters – is given in Section 3.3.3 (page 343, line 18ff, discussion paper).

Comment: Please give similar information for all instruments – e.g. which detectors, spectrometers and light sources were used. It would be useful to have a table with the similarities and differences between the same techniques, especially for the CRDS instruments. Also, maybe you can describe the common features of the CRDS instruments in one place and then detail the difference, more importantly why the groups chose different configurations. Essentially the pros and cons.

Response: See our comments for referee #3.

Comment: The use of the goodness-of-fit parameter (q) indicates that the measurement errors are underestimated for most of the instruments on most of the days. The authors should specify whether its because of the assumption of a linear model or whether its a non-normal distribution. If indeed the linear relationship is in question for most of the experiments, a different statistical method should be used to calculate the correlations.

Response: Provided that the instruments' operators calculated the correct precision of each data point, very small q-values simply provide a hint (but no proof) that the assumption of a linear dependency of both data sets within the limits given by the instrumental precision is doubtful. The statistic cannot answer the question what the cause could be. The latter is discussed in Section 4.4.
Whether a non-normal distribution of the measurement errors (cf. Figure 2, UAF and UCC) could (at least partly) explain the low q-values is difficult to quantify.
We selected the 'fitexy' routine to calculate slope and intercept of the linear correlations in Section 3.3.3 because we showed in Section 3.3.2 that this routine can be successfully applied for correlations for which the reduced chi-squares value $X^2/(n-2)$ does not exceed ≈15 (which is the case for 90 % of the data pairs in Table 4).
Only few exceptions with higher values appear in Table 4. However these cases can be traced back to the technical issues discussed in previous sections like problems resulting from missing background spectra (UCC, c.f. Sections 2.1.5 and 3.2) and the ingress of ambient air into the detection cell through a leak (c.f. Sections 2.1.4 and 3.2) which lead to more systematic data fluctuations not covered by the instrumental precision.
We adapted the text on page 341, line 27 to: The regression parameters are summarized in Table 4. Slopes and intercepts of the regression lines were calculated by the routine "fitexy" because the majority of the data sets showed reduced chi-squares values <15 (cf. Section 3.3.2).

Anonymous Referee #2

Received and published: 28 February 2013

Comment: Because of the large number of participating instruments, readers need to refer to section 2.1 frequently to pick up information of instrumental properties (e.g., with or without filters, transmission efficiency, inherent consideration of aerosols using broadband spectral features, and method to determine zero level and its frequency, etc).
I would recommend adding several rows in Table 2 to summarize these features.

Response: see our response to referee #3

Comment: Abstract, page 305, line 22. Does +/-5.3 pptv represent 1-sigma range?

Response: We changed the text in the abstract and the summary and specify all statistical parameters consistently as median, 1st and 3rd quartile and min/max values.

The median of the coefficient of determination (r^2) over all experiments of the campaign (60 correlations) is r^2=0.981 (1. quartile (Q1): 0.949, 3. quartile (Q3): 0.994; min/max: 0.540/0.999). The linear regression analysis of the campaign data set yielded very small intercepts (median: 1.1 pptv; Q1/Q3: -1.1/2.6 pptv; min/max: -14.1/22.2 pptv) and the slopes of the regression lines were close to unity (median: 1.01; Q1/Q3: 0.92/1.10; min/max: 0.72/1.36).

Comment: Abstract, page 306, line 3. Important NO3 loss likely occurred on filters, in addition to cell walls.

Response: We adapted the text to: In individual experiments the discrepancies increased with time, pointing to additional NO3 radical losses by aerosol deposited onto the filters or on the inlet walls of the instruments.

Comment: page 309, line 16. The meaning of "calibration" is not very clear here. How many instruments are calibrated by generating known concentrations of NO3 during the campaign?

Does the calibration mean determination of transmission efficiency through filters and inlets, on a relative basis, and/or path lengths in some cases?

Response: All instruments used absorption spectroscopy, hence their absolute calibrations all rely on the accuracy of the NO3 absorption cross sections. The Yokelson et al. (1994) NO3 cross sections were used in the data work-up for all instruments (cf. Section 2.1), so any uncertainty in the Yokelson cross sections does not affect the correlations between pairs of instruments. No instrument was calibrated by a known amount of NO3. Only the NOAA-CRDS instruments used N2O5 synthesised during NO3Comp to determine the transmission efficiency of the NO3 and N2O5 detection channels. However, a contamination in the N2O5 source made the measurements unreliable (see Section 2.1.1).
We want to mention that the instruments performed a dedicated inlet transmission loss experiment on 21 June, 07:30 - 08:15 UTC). NO3 was generated inside SAPHIR (160 ppbv O3 and 1 ppbv NO2), and sampled as a common NO3 source by all the "outside" instruments for different flow rates/residence times through their lines (the small, slow changes in SAPHIR NO3 during this experiment were corrected using the NOAA measurements).

Other factors controlling the accuracy are the transmission efficiency through filters and inlets and the effective path length over which the gas is present in the cavity (see our answers to ref. #1).
We changed the text to: Each of the participating instruments adhered to different calibration schemes of the radical transmission efficiency through filters and inlets and of the effective absorption path length in the cavity and was likely to exhibit different sensitivity to potential artefacts such as reactive trace gases or aerosol.

Comment: page 309, line 21. Probably five of seven, not five of nine

Response: See corresponding reply to referee #1.

Comment: page 314, lines 1-3. How constant was the decay rate of -0.4% h-1 for different experiments with different aerosol loadings? At which condition was it determined?
Trends in Figures 11 and 12 for UAF-CRDS indicate that the decay rate changed with aerosol loadings for these experiments?

Response: The experiment to determine the initial transmission efficiency of the instrument was performed in aerosol free air at a nearly constant NO3 mixing ratio in the chamber. The decay rate given is supposed to account for a general “aging” of the inlet lines and filter used and was assumed constant for the campaign. The correction factor was very small and essentially not significant. We compared the regression results of the UAF-CRDS instrument based on the submitted data set (with the correction factor applied) with the regression results of a data set for which the loss factor was back calculated and could not find a significant difference of the regression parameters.

Comment: page 315, line 1. Delete "emitting"

Response: Deleted.

Comment: page 315, lines 24 and 25. What is the meaning of the factors exceeding unity? Were they unchanged with time? In figure 12, there is almost no trend with aerosol loadings for MPI-CRDS. Does this indicate that the transmittance was almost constant at least for this experiment?

Response: This factor is applied to the measured NO3 mixing ratio in order to account for the losses of NO3 in the instrument.
Loss rates of NO3 during transmission through the inlet tubing and the cavity was made on four separate occasions during NO3Comp by variation of the flow rate of air drawn through the instrument. The tests were carried out in the absence of aerosol and when stable concentrations of NO3 were observed in the chamber and with different NO3 concentrations and NO3/N2O5 ratios. The NO3 inlet loss was fairly constant over a period of 7 days (see Schuster et al., AMT, 2009). We adapted the text to: An averaged correction factor of 1.13±0.1 was determined. Filter losses were measured later in the lab resulting in a correction factor of 1.18±0.1. Both factors were taken constant for the campaign and were applied to correct the measured mixing ratios.

Comment: page 316, line 21. What are the four tubes for?

Response: It was assumed that fast flows through 4 narrow tubes might mean smaller wall losses of NO3 than a slower flow through one wider tube.

Comment: page 319, line 20. is measured and spectrally resolved

Response: Corrected.

Comment: page 327, line 28. What is the typical size distribution of the generated ammonium sulphate particles?

Response: We extended to the text to: Ammonium sulphate, (NH4)2SO4, aerosol was added directly by spraying an aqueous solution into the chamber using a nebuliser. The resulting size distribution had a count mean diameter of 70 nm, and the geometrical standard deviation was 1.56.

Comment: page 331-332. NO3 mixing ratios observed by UCC-IBBCEAS are always larger than others on 10, 11, 13, and 20 June. What is the possible explanation for this? Does the zero level always drift in the same direction?

Response: On the days concerned (10, 11, 13, and 20 June) measurements of reliable spectra of I_0 were not possible for reasons outlined in item 1 on page 331. Overestimating I_0 can hence lead to a systematic increase of the true mixing ratio. There is no possibility for these days to independently verify whether the I_0 spectra used for the evaluation of mixing ratios were too large or too small. There is no reason for I_0 (presumably referred to as "zero level" in the comment) drifting systematically in one direction. There were variations in the lamp intensity during the campaign, however, the evaluation of the reflectivity spectra revealed that the intensity changes of the lamp occurred non-systematically during the duration of the campaign. In order to be less ambiguous about a potential drift of the lamp intensity in one direction we changed the text to: This indicates that the long-term intensity variation of the lamp was a limiting factor for the UCC-IBBCEAS instrument during some experiments of NO3Comp.

Comment: page 332, line 1. Which kind of "drift" did the lamp have?

Response: We changed the text to: This indicates that the long-term drift of the lamp emission spectrum was a limiting factor for the UCC-IBBCEAS instrument during some experiments of NO3Comp.

Comment: pages 333-336. For each experiment, which reaction controls loss of NO3 in the chamber? What is the typical NO concentration in the dark chamber? Does it influence NO3 loss? Can the NO3 lifetime be varied from experiment to experiment by factors of >10, for example, and thus affect the transmission efficiency through inlet lines?

Response: During experiments with added hydrocarbons, the NO3 lifetime was controlled by the HC. In clean-air experiments NO3 is likely limited by deposition/reactions on the chamber walls. Fry et al. 2009 estimated an NO3 deposition lifetime in the order of 30 min in the dry chamber The typical mean NO mixing ratio in the presence of NO3 is (0.0 ± 1.2) pptv. At 1 pptv the NO3 lifetime due to reaction with NO is 14.5 min at 298 K. With residence times of NO3 in the range of 100 ms to 2.7 s inside the instruments, it is unlikely that the transmission efficiency of NO3 is notably affected by the residual NO concentration in the chamber.

Comment:. page 334, lines 1-3. I would like to see another time series plot with increased time resolution to highlight fast variation in supplement.

Response: We think it will not be necessary to add an additional figure with high time resolution, because the figures in the final PDF version of the paper will be scalable to any size. A new figure would just contain redundant information compared to Figures 3 and 4.

Comment: page 334, line 5. What is the role of butanal in this experiment? Does it limit NO3 lifetime?

Response: We repeated a butanal oxidation experiment which we performed earlier in the chamber in 2005 (Bossmeyer et al., GRL, 33,L18810, 2006). The butanal oxidation is slow and the resulting NO3 lifetime (14 min) was in the order of the lifetime due to wall losses.

Comment: page 348, line 5. Does the water vapour absorption influence path length determination over the whole wavelength range studied with the UHD-CEDOAS instrument, or only at discrete H2O lines?

Response: Generally the path length reduction is the stronger the larger the absorption cross section is. Due to the low resolution (1.06 nm) of the spectrometer all narrow-band water vapour absorption lines within the observed wavelength interval are broadened and their peak absorption is reduced accordingly over full spectral range studied making the path length differences between water lines smoother compared to higher spectral resolution. For details please see Figure 7 in Langridge et al., 2008.

Comment: page 350 line 6. I would like to know pressure and residence time between sampling and detection in the UAF-CRDS instrument.

Response: The instrument sampled at ambient pressure. The residence time of air in the inlet line (L=0.4m, i.d.=6.3 mm, flow rate: 8 slm) was about 100 ms.

Comment: page 351. Does the co-presence of aerosol particles and high humidity affect the NO3 detection?

Response: When comparing aerosol experiments with and without added water vapour we don't see any systematic effect in the measurements which would support this question.

Comment: page 353, line 11. The maximum of the intercept values from Table 4 is 28 ppt (15 June, FZJ).

Response: Corrected !

Comment: Figure 12. Can the x-axis be changed to filter-laden aerosol amount or time from the beginning of the experiment? The data points with the lowest dry SOA surface area frequently show different tendency. Is there any specific reason for this? What does the tendency for the ULEIC-BBCRDS mean, where a filter is not used? Wall loss?

Response: As stated in the manuscript, the underlying chemical mechanism of the NO3 loss in the presence of aerosol was not identified. The observations during NO3Comp and the findings of Fry et al., 2009, 2011 (who investigated the formation of reactive organic aerosol during the oxidation of β-pinene and limonene by NO3) support our hypothesis that NO3 reacts with (organic compounds on the) SOA aerosol which is deposited on the filter surface or on the walls of the inlet systems/detection cells of the instruments.
Concerning the x-axis of Figure 21: Because some instruments used filters some didn't we did not try to produce a graph as proposed by the referee. Prior to the submission of the manuscript we compared the histograms of the NO3 loss plotted versus experiment time with the ones plotted vs dry SOA surface measured in the chamber. Both plots gave nearly identical results with some more pronounced trends of the NO3 loss when plotted vs SOA surface. Hence we preferred the presentation of the direct dependency of the calculated relative difference of NO3 on the measured SOA concentration, since time is just a correlated quantity.

Comment: One general question: In ambient air near surface, NO3 concentrations should be typically lower than those studied in the experiments. Is there any systematic difference in the regression line parameters for the whole range and for a low range (e.g., 0-20 ppt), for pairs of instruments having low-enough detection limits?

Response: On page 344, lines 6-20 we compared the regression parameters for the two episodes on 18 June. The slopes, intercepts and correlation coefficients during the high NO3 (< 150 pptv) and low NO3 (< 12 pptv) phases show no significant differences for the instruments having low-enough detection limits (UCC-IBBCEAS, UAF-CRDS, and NOAA-CRDS).

Anonymous Referee #3

Received and published: 6 March 2013

Comment: My main comment on the information presentation is that the authors should provide a clearer description of the lessons learned from the experiments. For example, what would be an optimal exchange frequency for the aerosol filters? What further improvements do these instruments need? What error sources should be further reduced? How do the lessons learned apply to atmospheric deployments and what do the intercomparison results mean for past and future atmospheric observations of NO3. Without clearer and more general conclusions this manuscript will only serve the participants as a proof of the quality of their individual instruments, but have no lasting value for the scientific community.

Response: We do not agree with the referee's general statement that the manuscript in its current status will not have a lasting value for the scientific community. The value is that the instruments agree well over a wide range of experiments, and that gives confidence that the instruments also operate well for their own ambient measurements. See Ref #2 "..a firm basis of a synthesis across findings for individual field campaigns performed at different time[s] and place[s] using different instruments". See also Ref #3's own comment about p309, lines 9-13 below.

We agree it would be nice to make an evaluation of NO3 measurement capability, both backward- and forward-looking; but that is almost a whole review paper in itself. Also can the present experiments really say with certainty that the data from campaign X were wrong by Y%? There are too many other possible variables.

However, we took up his suggestion and extended the conclusions (last paragraph on page 354, line 15ff):The filters should be exchanged frequently, however, there is no general recommendation for the maximum filter "age" because the filter "lifetime" may vary with the aerosol composition and the aerosol mass load in ambient air. The inlet and filter losses are specific for each instrument and hence need to be characterised rigorously and individually for the individual instruments under the different operating conditions in the field.
The NO3 transmission efficiency due to losses on the filter, in the inlet system, and inside the cavity have to be precisely determined for each instrument. Generally losses to the walls can be minimized by reduction of the residence time of the sample air in the instrument. Fuchs et al. used sampling lines with 6.4mm i.d. and reduced the pressure to 350 hPa to establish a residence time of 100 ms between the filter and the centre of the

detection cavity. The resulting wall losses of NO3 were 3±1%. Together with NO3 loss on the filter, an overall NO3 transmission of 92±2% was achieved.
Due to the strong temperature dependence of the N2O5 equilibrium constant (R5) the difference between the temperature of the ambient air and the (usually higher) temperature inside the NO3 detection cell should be minimized in order to avoid decomposition to NO3 especially when the N2O5/NO3 ratio is large.

Comment: At times the manuscript is unnecessarily lengthy and the many details obscure the main points of the text. For example, the CRDS instruments and the CEAS instruments are so similar in their principle that Section 2.1 could be considerably shortened by providing a general description of each instrument principle and tables with the details of the different instruments. The current detailed description could be moved into the supplemental material. This would also make it easier to understand where the differences between the various instruments really lie. Also the description of the individual experiments could be shortened since much of the information is shown in Table 1 and Figures 3 and 4.

Response: We agree that the different groups of instruments participating in the intercomparison seem to show some principle similarities as far as their detection scheme is concerned. However, some instruments were in an early stage of development some actually were prototypes whose setup was not published. Furthermore the instruments used different measurement strategies (e.g. single wavelengths vs broadband detection), different hardware, acquire their reference measurements differently, etc. Hence we disagree with Ref #3 that the instrument descriptions and operational details can be moved to Supplementary Material. In our opinion it is entirely justified to provide such detail in the main body of the paper and we think that a journal like *Atmospheric Measurement Techniques* provides a suitable forum for this approach.

However, we agree that an additional table summarising the main technical features can provide readers with a quick and easy reference to the instruments. We will add such table to the Supplement of the revised manuscript.

Comment: Page 309, lines 9-13: I agree with the authors that intercomparisons in environmental chambers are a great way to learn about instruments and their behaviour. However, this is only one side of the story. One could also argue that the highly simplified conditions in a chamber do not allow for a full investigation of possible, and often unknown, interferences with other atmospheric components. I suggest adding a sentence that more clearly explains this issue.

Response: We adapted the text on page 329, lines 1-2 accordingly: In order to allow for a full investigation of possible interferences with other unknown atmospheric components, ambient air was pumped into the chamber on 11 June.

Comment: Page 310 line 21. This is a contradiction with the introduction that states that DOAS has been used for the past 3-4 decades to measure NO3.

Response: Text was adapted to: The most mature cavity enhanced instrument employed in this campaign…

Comment: Page 315, line 18ff: To be consistent the various error sources in the MPI-CRDS should also be quantified. Again, it may help listing all the error sources of the instruments in a table to give the reader a better overview.

Response: We added the required information to the revised version: The overall 1σ accuracy of the NO3 measurement by MPI-CRDS of 14% is given by the uncertainties of the determination of filter loss (8.5%), inlet/cavity loss (8.9%), cavity length (2.9%), and absorption cross section (5%).

Comment: Page 324, line 3: ”deconvolved” should probably be ”convoluted”?

Response: Changed.

Comment: Page 325, line 2: Should the error of the cross section be 5% (1σ)?

Response: Correct! The resulting accuracy was adapted in the text accordingly to be 8%.

Comment: Page 325 bottom and page 326 top: Please comment on how this systematic variability was identified. A better explanation on how arc-lamp instabilities can cause such effects would be helpful here.

Response: Compared to the data of the concurrently measuring NO3 instruments, the NO3 data measured by DOAS often showed sudden systematic changes of the mixing ratio by typically ±(20-30) pptv which introduced a bias in the DOAS data (cf. Fig. 3, 14 June and Fig. 3, 14/15 June). This behaviour can not be explained by the ‘known’ random noise of the instrument, which is <10% (see Fig. 2). The changes were neither caused by addition of reactants nor by aerosol. We explain this observation by intermittent shifts of the arc emission point on the surface of the electrodes of the Xe lamps. As a result of the geometrical displacement of the arc, the light transmission of the White cell will be different and also the spectral emission of the lamp changes slightly. Both effects

lead to a change of the spectral baseline compared to the zero-intensity spectrum which was measured at the beginning of the experiment. A systematic bias of the retrieved mixing ratios will be the consequence. All Xe lamps used for DOAS measurements during the intercomparison were taken from the same new batch of lamps ordered prior to the campaign. Hence, we assume a problem during the manufacturing process of the Xe lamps. We adapted the text on page 326, lines 2-3 to: We explain this effect by intermittent shifts of the arc emission point on the surface of the electrodes of the Xe lamps used during NO3Comp. As a result of the geometrical displacement of the arc, the light transmission of the White cell was different and also the spectral emission of the lamp changed slightly which led to a different spectral baseline compared with the zero-background spectrum recorded at the beginning of the experiment. A systematic bias of the retrieved NO3 mixing ratios was the consequence.

Comment: Section 3.1: The precision is not defined as the error at zero concentrations, but rather, speaking in simplified terms, as the ”reproducibility” of a measurement at any concentration. Please explain why the precision at concentrations near zero are representative for the precision at higher concentrations.

Response: We did not determine the precision of the instruments from measurements in zero air. In Section 3.1 we tested if the precision of the various instruments as defined by the respective operators is consistent with the observed scatter of measurements in zero air.
The precision of an optical absorption instrument, i.e. the size of the error bar of each single data point, is limited exclusively by the various random statistical noise sources which contribute to the detection signal. If the nature of these noise contributions and their statistical dependencies (for instance, Poisson statistics describing the noise of counting events) are known the detection precision for a defined data acquisition time interval can be fundamentally calculated and attributed to each measured data point. This process is different for the different types of instruments and has to be quantified by the respective operators.
A test of the correctness of the determination of the instrumental precision can be performed for the limiting case of signal-to-noise ratio of unity, i.e., in zero air. The data points ‘measured’ in zero air are supposed to results in a mean of zero and the statistical distribution of the data scatter should follow a Gaussian distribution with a 1σ-width equal to the pre-described precision. This test is presented in Figure 2 and in Table 2.

Comment: Page 331 top. This paragraph is very similar to the caption of Figure 3 and can be shortened accordingly.

Response: Done.