

Response to editor's comments:

We thank the editor for his detailed review of our revised manuscript. We fully agree with the suggestions and correspondingly we fundamentally re-structured and rewrote the paper. The major modification includes the removal of the statistical modelling part and extension of the technical description. Now, the paper is focused on the analytical instrument, its characterization, validation, and application. In the following, we address the specific comments individually.

Detailed comments

Introduction

The overview of measurement techniques that are being used to measure atmospheric NO₂ is not complete. DOAS, LOPAP, LIF, luminol chemiluminescence should be mentioned. See for example, Villena et al., *Atmos. Meas. Tech.*, 4, 1663–1676, 2011; Sluis et al., *Atmos. Meas. Tech.*, 3, 1753–1762, 2010; Dari-Salisburgo et al., *Atmos. Env.* 43, 970–977, 2009; Dunlea et al., *ACP* 7, 2691, 2007; Thornton et al. *Anal. Chem.* 72, 528-539, 2000.

For the measurement of NO₂, the widely used CLD technique can have interferences in polluted air. This is not only the case for instruments with catalytic converters as mentioned in the introduction, but also for instruments with photolytic converters which can show interferences from PAN, HONO, VOCs (Reed et al., *Atmos. Chem. Phys.* 16, 4707, 2016 and Villena et al., *AMT* 5, 149, 2012).

We added a corresponding paragraph and discussed the cross sensitivities of CLD detection and alternative NO₂ detection methods.

Instrument

Some more technical details describing the instrument should be given, e.g. supplier and model of components (e.g. breadboard, etalon, mirrors, mirror holders, reflective coating of absorption cell), reflectivities of absorptioncell walls and of turning mirrors, laser beam diameter, spectral bandwidth of laser, total optical transmission of absorption cell.

The instrument description now features the requested details.

Evaluation of measured spectra

The NO₂ concentrations are calculated according to Beer's Lambert law and converted into mixing ratios. The values of all experimental parameters needed for the calculation should be specified together with their errors and uncertainties (absorption path lengths, absorption cross section, pressure and temperature in the absorption cell). How large is the accuracy of the resulting NO₂ mixing ratios?

An estimation of the total error of the spectroscopic measurement was added and discussed.

Figure 1c displays a measured absorption spectrum with NO₂ and H₂O lines. Was the spectrum measured in synthetic air or in ambient air? Besides the intense absorption lines, there seem to be modulations like ripples in the spectrum. Are these caused by the instrument or by other weak absorbers? A spectrum of pure synthetic air should be shown for comparison.

For clarity, we replaced the spectrum with two spectra recorded using NO₂-free air and ambient air, respectively. Both spectra show the water absorption line at 1600.47cm⁻¹.

The periodic ripples in the spectrum originate from optical interference fringes occurring in the multi-pass cell and the associated optical elements.

How is the spectral baseline intensity (I_0 , without NO₂ absorber) at the central position of the NO₂ line determined? What is the smallest detectable absorbance of the instrument and what is the corresponding limit-of-detection for NO₂?

The fitting routine includes the determination of a baseline by fitting a polynomial (up to order 4) to the spectrum left and right of the absorption feature. The normalization of the transmission spectrum is based on associated zero-level (laser off) measurements.

The residual of the Voigt profile fit is 3×10^{-4} at 1s averaging time. The 1σ detection limit can be retrieved from the Allan-Werle Deviation plot: At 1s it is 300 ppt and the minimum is at 30 ppt after 200 s averaging.

p. 6, line 4-5: "Moreover, data were omitted when the NO₂ spectrum was not clearly identified by the processing algorithm." Which criteria are used to decide, if a spectrum can be used or not?

Normally, if the absorption of the analyte is smaller than the noise of the baseline, fitted spectra should yield concentrations scattering around zero within their fitting error.

The water absorption line shown in the spectrum was used to determine whether the QCL is operating at the right frequency. When the measured water concentration was not within the expected range the data have been discarded.

As the water line is always detectable in the spectrum, it was used as a frequency marker to allow fitting of the NO₂ spectral region even if no detectable NO₂ concentration was present. This results indeed in scattering around zero within the fitting error. The zero-point however may drift (caused by interference fringes) and therefore is regularly determined and later corrected for by measuring filtered NO₂ free air.

There are no spectral interferences. To ensure data quality, data where the zero-point correction exceeds ± 5 ppb were excluded from the analysis.

p. 9, line 12-15: "Based on these results, the measurement strategy was designed such that it applies a drift correction to the concentration measurements by measuring NO₂-free air for 2 min after 18 min of measuring. This offset correction is achieved by subtracting the linear interpolation of the mean of the zero-air measurements before and after each 18 min measurement period." What is the cause for the offset drift and how large (ppbv?) were the corrections? How can a concentration meas-

urement based on Beer's Lambert law have a drifting offset? Are there spectral interferences from other variable atmospheric compounds?

We rewrote the corresponding paragraph for more clarity. As mentioned above interference fringes originating from the optical setup cause a structure on the retrieved spectra that are interpreted as absorption and thus cause an offset from zero in NO₂ free air.

The processing of the measured atmospheric spectra, the determination of the NO₂ absorption cross section and the determination of the final NO₂ concentrations should be explained in detail.

We extended the paragraph describing the determination of NO₂ concentrations.

The QCLAS instrument should, in principle, allow absolute, calibration free measurements of NO₂. Have the authors tested the instrument against an independent NO₂ standard (e.g. certified gas mixture)?

How was the agreement?

Yes, the instrument should in principle be calibration free. However the uncertainties in cross section, path-length, pressure and temperature lead to a systematic error that makes a calibration necessary. Since calibration with NO₂ reference gases is a tricky and currently not established issue. We therefore decided to use a CLD (despite its shortcomings and possible cross sensitivities) as a reference method. This calibration strategy allows a comparison of the mobile QCLAS measurements on top of the tram with the CLD measurements at the two air quality monitoring sites situated next to the tram tracks. We rewrote the corresponding paragraph.

The performance data of the new instrument should be mentioned in the abstract: upper and lower limit-of-detection, time resolution, precision and accuracy. In the text, a table with specifications and instrumental parameters would be useful.

We mention the instrument's precision, time resolution and dynamic range in the abstract.

Could the instrument be used for absolute water vapor measurements? Figure 1 shows an isolated H₂O absorption line that could probably be used?

Yes, indeed. In fact, the water concentration retrieved by the fitting algorithm was used as a check to make sure that the laser operates in the right spectral region.

Possible limitations and interferences should be discussed in more detail.

o Are there any known spectral interferences?

No.

o The production of NO₂ from the reaction of NO and O₃ in the inlet line and absorption cell is mentioned. Can you quantify the expected percentage range of perturbation for low O₃ (10 ppbv) and high O₃ (100 ppbv) conditions?

The reaction of NO and O₃ is fast and leads to formation of NO₂ in the immediate vicinity of the source of NO emissions (i.e. along roads). In the inlet line of instruments, this reaction can continue and additional NO₂ can be formed. We calculated the formation of NO₂ in the inlet line of the reference CLD instrument at the urban traffic site SCH in Zurich, where concentrations of NO and O₃ are available. The residence time of the air sample in the inlet line of the CLD is about 5 s. For the time period discussed in the paper (September 21, 2015 until March 9, 2016) we find an average NO₂ formation in the inlet line of 0.5 ppb, the average NO₂ concentration during that time period at SCH was 24.3 ppb. For the conditions at SCH, the NO₂ formation in the inlet and absorption cell of the QCLAS would be somewhat larger, because of the longer residence time of the air sample (about 10 s), for the QCLAS we calculated an average NO₂ formation of 1.0 ppb. Note that the formation of NO₂ in the inlet line depends on the O₃ concentration, but also on the concentration of NO. The maximum formation of NO₂ in the inlet line does not occur at peak ozone conditions, because NO concentrations are then typically low. Largest NO₂ formation in the inlet occurs at conditions with moderate but sufficient levels of both O₃ and NO. The NO₂ formation in inlet lines due to reaction of NO and O₃ is largest at locations that are strongly impacted by traffic (or other sources of NO). The SCH site is an urban traffic site next to a major road and the calculated formation of NO₂ can be considered as an upper limit for NO₂ formation in the inlet system of the QCLAS during the described mobile deployment. We suggest not going into details here, because NO₂ formation in inlet lines is most of the time a small perturbation and occurs in the QCLAS but also in the reference CLD.

o The absorption cell is operated at reduced pressure favouring the thermal decomposition of labile components like PAN. Could this effect play a role in the instrument on a hot summer day in a polluted urban environment?

This effect could happen but would lead to a very minor interference. Measurements within the Swiss Air Quality Monitoring network show that PAN concentrations in the Zurich area reach during summer maximum hourly concentrations of 1.5 ppb, and concentrations well below 1 ppb during the cold season, when the mobile measurements with the QCLAS have been performed. Note that PAN is one of the interfering compounds when measuring NO₂ with CLD instruments.

o Page 10: "... a correction to the normally calibration free QCLAS method was necessary because of loss of NO₂ during the sampling and in the long path cell... The contact of NO₂ with metal or plastic surfaces leads to dissociation." Is there direct, independent experimental evidence for such a wall loss in the instrument, or is this just an assumption? Which kind of materials were used? How long was the inlet line? Has the loss been checked by using different lengths of inlet lines or by variation of the volume flow rate of the sampled air? If the NO₂ wall loss is high, it could change over time with surface conditions (adsorbed H₂O, trace gases, particles). Are there any indications for such a behaviour?

The NO₂ loss in the multipass cell has been observed experimentally, where the NO₂ concentration of a gas flow through an inlet line was monitored by a CLD while passing or bypassing the multipass cell. A

systematic loss of 10% in the latter case has been repeatedly measured. However, no systematic survey of different flows has been done, except the average condition that were used for the field campaign. Nevertheless, similar effects were found for dry NO₂ calibration gas.

Measurement comparisons

On page 10, the NO₂ measurements obtained by the QCLAS instrument and a CLD instrument are directly compared side-by-side at one AQMS station. The comparison and the conclusions drawn from it must be examined more critically.

The authors have apparently more confidence in the CLD measurements than in the QCLAS measurements and correct (calibrate) the QCLAS data to match the CLD measurements. This is somewhat surprising because the QCLAS is in principle an absolute technique, while the CLD needs calibration. Since the QCLAS measurements were scaled to the CLD data, Figure 4 cannot be said to demonstrate "excellent agreement". Rather, the scatter plot demonstrates a high linear correlation of the two techniques and a negligible offset. I suggest to plot the uncorrected QCLAS data against the CLD measurements and discuss linearity, slope and offset. The deviation of the slope (1.2) from unity should be discussed more critically with respect to known measurement errors of the QCLAS (see comments above) and CLD instruments (see comments below). Only if these cannot explain the 20% difference, it is meaningful to speculate about other possible reasons.

No, we do not have more confidence in the CLD measurements. The point is that calibration of direct and specific NO₂ instruments using NO₂ reference gases is still an unsolved issue. There are reference gases available (either static ones in cylinders or dynamically generated NO₂ using permeation sources), but it is difficult to avoid losses of NO₂ due to wall effects and therefore to produce a reliable reference. As emphasized in the revised manuscript, we therefore decided to correct the QCLAS to match the CLD. Another advantage of this strategy is that the QCLAS measurements can be directly compared with the CLD measurements at the two AQM sites that were regularly passed by the tram (see revised manuscript).

As suggested by the reviewer, Fig. 6 shows now the relation between uncorrected QCLAS and CLD, the reasons for the observed differences between QCLAS and CLD are discussed in more detail.

The CLD technique is known to have several sources of measurement errors which are not discussed in the paper, but could contribute to the observed discrepancy. How was the CLD calibrated (against which standard?) and what was the accuracy of the calibration? The detection sensitivity of CLD instruments have a water vapor dependence. Was this corrected? CLD measurements of NO₂ rely on chemical conversion of NO₂ to NO. For catalytical and photolytic converters, interferences have been reported which can be quite significant in polluted urban air. To my knowledge, the Horiba APNA-370 instruments are equipped with catalytic converters which have a significant sensitivity to other nitrogen species (e.g. PAN, HONO etc.) yielding too high NO_x readings. Have these interferences been characterized for the CLD used in this study and could these interferences explain the difference of 20 % between the QCLAS and CLD measurements?

In the revised manuscript we discuss the cross sensitivities of CLD instruments equipped with catalytic (molybdenum) converters in more detail. However, one should keep in mind that this measurement technique is the reference method in Europe and is mandatory for regulatory air quality measurements. The interferences of the CLD instrument have not been characterized within the scope of this study, but from earlier work we know that the conversion efficiency for other nitrogen compounds such as PAN and HONO is very high. From parallel measurements with two CLD instruments, one equipped with a catalytic and one equipped with a photolytic converter, we know that the interference of the standard CLD instruments cannot explain the observed difference between QCLAS and CLD. In Zurich (urban background) we find in an unpublished study, that on annual average 1.7 ppb (or 10%) of measured total NO₂ can be attributed to interfering compounds. This is explained in the revised manuscript. CLD instruments are calibrated against NO reference gases. This is also mentioned in the revised manuscript. The contribution of the uncertainties of the calibration gases (primary and transfer standard) to the total uncertainty budget is small. On the other hand, the interference of CLDs to water as mentioned by the reviewer is the largest component of the total uncertainty of NO₂ measurements using CLD (see NABEL 2016 for the detailed uncertainty budget estimation). The CLD measurements are not corrected for the interference of water, because the relationship is unknown.

Possible NO₂ losses in the inlet line of the QCLAS instrument are mentioned (s. above). What about possible losses in the inlet line of the CLD? Were other materials used than for the QCLAS instrument? Losses of NO₂ in the inlet line of CLD instruments are not known. As outlined above, it is currently difficult to provide instruments with precisely known NO₂ reference gases. For CLD measurements within regulatory AQ measurements it is currently only recommended to use materials for the sampling line that are fit for this purpose (e.g. PTFE), see e.g. EN 14211 (CEN, 2012).

How far apart were the inlets of the two instruments? Were both inlets exposed to the same amount of direct and diffuse solar UV? If the NO₂ values are different, the NO₂/NO ratio may be different even if NO_x is the same.

The QCLAS was placed directly next to the inlet of the NABEL station (d<0.5m). Both were exposed to the same solar radiation.

Statistical model

In addition to the report of Referee #1, I have the following question.

Has the "statistical model approach" ever been tested with simulated NO₂ concentrations from a regional atmospheric-chemistry model? One could simulate stationary or mobile measurements taking modelled NO₂ values or at fixed locations or along selected trajectories (roads), respectively, and superimpose artificially measurement noise. The statistical model could then be fed with simulated measurements and its output could be compared with the complete spatial NO₂ distributions from the atmospheric-chemistry model. The whole test could be done for variable ozone background and varying solar radiation.

Following the Reviewers' recommendation the modelling part has completely been removed from the publication.

Technical comments

- p. 2, line 14: give reference for electrochemical sensors.
- p.7, definitions of DTV: do you mean "at a distance of..." or "within a distance of..."?
- p. 9, Figure 3: the explanations of the colors in the legend and in the figure caption disagree.
- Figure 4: are the errors of the fitted parameters 1sigma values?
- Figure 7: errors of the fit parameters should be given.
- p. 13, line 11: one of the reasons for the large scatter in Figure 7b is explained to be due to "a larger distance" than in Figure 7a. However, for both panels, a radius of 30 m is specified.
- NABEL 2016: where is the report available? The source, publisher, authors, ISSN etc. should be given in the reference.

In the revised manuscript, all technical comments have been addressed.