**Anonymous Referee #1**

In the experimental description of the NOx QCLAS great detail is given about the precision of the instrument and potential uncertainties due to background structures. To summarize this, it would be nice if the authors would give a quantitative estimate of the total uncertainty for NO and NO2 measurements, respectively.

For the measurements described in this paper, the main uncertainty contributions are: i) zero-air determination, ii) drifts due to temperature fluctuations, and iii) the uncertainty of the span values, i.e. the employed calibration gases.

Potential systematic errors when using the NOx-scrubber for background definition were investigated by comparing the measured values for high purity $N_2$ and NOx scrubber. Small, but still detectable systematic differences between the two samples could be quantified. Their value during the whole measurement campaign was -32 ± 15 ppt for NO and 15 ± 9 ppt for $NO_2$. Measured every two hours, they were included as offset correction for the QCLAS data. Uncertainties due to background drifts, mainly caused by temperature fluctuations, can be significant (> 40 ppt), if no action is taken. However, the configuration of the sampling strategy as we have used, these uncertainties were minimized (see measurement strategy) and their contributions were estimated to 5 ± 4 ppt and 11 ± 7 ppt for $NO_2$ and NO, respectively. The estimated uncertainties resulted from the dilution process were 0.1% and 0.4% for NO and $NO_2$, respectively. To summarize, the total uncertainty of the QCLAS is dominated at large NOx values by the accuracy of the calibration gases, which lies in the percent range (±2−4%), while at sub-ppb levels it is mainly given by the uncertainty of the ”zero-level” measurements, which is below 10 ppt.

Added text to include this detail of total uncertainty.

From the Allen variance plot it seems that the stability of the instrument with respect to background drifts is of the order of 2 min. In the results and discussion section the authors state that background measurements are performed every 10 min? Is this sufficient to account for background drifts?

Analyzing the Allan variance plot makes evident that at least two slow fringe components (see text) prevent long averaging times by introducing an instrumental drift. This shows up in the time-series of the background signal as a linear trend and would indeed deteriorate the measurement accuracy. Obviously, averaging over 10 min would involve additional error due to drifts. Our strategy to minimize this effect was to monitor this drift by re-measuring the “zero-level” after the ambient air measurements. Therefore, the measurement cycle was as follows: 1) background definition within 10 s (total flush time 110 s), 2) ambient air measurement (10 min), 3) NOy measurement (10 min) and 4) zero-air measurement. The difference between the latter and initial “zero-level” measurements was considered as real drift in the system and a linear interpolation between the two values was used to correct the 1 s measurement values within this interval prior averaging.

We modified the text to include a more detailed description of our measurement strategy for more clarity.

The comparison between the QCLAS and the CLD in-situ measurements presented in Figures 5 and 6 is quite impressive, but rather qualitatively. I would appreciate a more quantitative analysis, including a regression analysis.

We agree that a scatter plot is more adequately presented when regression analysis results are included. However, in our case (especially for NO) an ordinary least-square fit is "meaningless", because of two issues: 1) the span is rather low and 2) the distribution of the data is not homogeneous, but strongly concentrated around zero, close to the detection limit of the CLD, which accordingly influences the degree of freedom of a linear fit. Nevertheless, we addressed the request of the reviewer and performed an orthogonal distance regression, which includes the errors in both dependent and independent variables (X and Y-values) as well the weighting with the expected magnitude of errors for each individual data point. The fit results are now included in the Figure 6.

Finally, the authors discuss the differences in the NOy measurements, and claim that an (unlikely) 14 % difference in the conversion efficiency for PAN would be necessary to explain the difference. The conversion of PAN in a gold converter is quite straight forward (thermal decomposition followed by reduction of NO2). Therefore the conversion efficiency for PAN should be very similar to the conversion efficiency for NO2, which I guess has been measured for the two converters. The conversion of HNO3
(which might represent a large fraction of NOy at this altitude) is more complex, and thus prone to errors in the conversion. Has the conversion efficiency for HNO3 been determined for the two converters?

The conversion efficiency (> 98%) for $NO_2$ was tested in detail, as also mentioned in the manuscript, and found to be similar within 1.5%. Unfortunately, we had no means to quantify the $HNO_3$ conversion and had to rely on previous data already published using the gold converter installed at the monitoring station (Zellweger et al., 2000, Zellweger et al., 2003, Pandey Deolal et al., 2012). These indicate uncertainties in the reported NOy values of about 9 to 15%. Thus, the difference that we observed between the two gold converters falls to the higher limit of uncertainty of the method, and as correctly pointed out by the Reviewer, it may well be caused by issues related to $HNO_3$ such as conversion efficiency of the gold converter, memory effects or inlet transmission efficiencies. Further investigation of the NOy data revealed that the temporal variability of its relative difference closely follows the humidity fluctuations in the ambient air. The relative differences in NOy were systematically getting larger when the relative humidity dropped below 70%. We have currently no plausible explanation for this behaviour.

We modified the text correspondingly.

## Anonymous Referee #2

***General comments:***

Perhaps the only general comment is that the instrument description lacks a specification of the QCL instrument's size, weight and power requirements. Since one aim of the paper is to establish the advantages (or disadvantages) of this approach relative to others, some description of these parameters is warranted.

Instrument size: two individual 19-inch rack modules (6U) incorporating optics and electronics, respectively. The QCLAS weight is 75 kg and its power requirement is 300 W.
Note: Accessories such as thermochiller, vacuum pump and gas-handling unit are not included.

Added text:
*"Overall, the spectrometer consisted of two individual full-size 19-inch modules (6U) incorporating optics and electronics, respectively. These units were built into a rack together with the custom-made gas handling and calibration module (3U), an UPS (2U, APC Smart series) to bridge short power outages, and a rack-mount thermoelectric liquid chiller (4U, Thermorack 300, Solid State Cooling Systems) used for cooling the lasers, IR-detectors and temperature stabilizing the optical module." The total weight, including the vacuum pump, of this autonomous setup was 175 kg and its power consumption was approximately 0.8 kW."*

***Specific comments:***
Page 8971, line 7-8: Although I largely agree with the author's statement, it is also fair to say that chemiluminescence instruments can be considered routine tools, at least for NO. The paragraph that follows says as much.

Modified sentence to adopt this suggestion: "*At present, chemiluminescence detection (CLD) can be considered as the standard technique, which is routinely used in field ….*"

Page 8974, line 10: "cells" rather than "cell's"

Done.

Page 8975, lines 14-16: Minor comment: Do the authors have a quantitative measurement of the relative humidity change through the NOx scrubber? Presumably it could introduce some time lag in situations where relative humidity changes rapidly.

Yes, we have the water measurements along the NOx data (as also mentioned later in the text) measured by the QCLAS (see Figure 1 in the manuscript). The relative humidity of the outside air is also monitored independently at the station. Thus, we can compare the two humidity data (ambient air and air through the NOx scrubber). Water vapor corrections of the NOx data are, however, based on the spectroscopically measured $H_2O$ values and thus, potential lag times have no influence on the data quality.

Added text to mention this detail.

Page 8979, line 4: Absolute humidity is given as a volume mixing ration – it would be useful to the reader to know the corresponding relative humidity at this site. Presumably these effects would be larger at lower elevation, higher absolute humidity sites.

Added text:
*"The relative humidity (RH) at JFJ varied between 20 and 100%, depending on meteorological conditions and wind directions, but predominantly (≃ 87% of all values) it was above 75% RH. During the campaign, the temperature gradually increased from −11 °C (in March) to around 0 °C (in June), while the air pressure remained rather constant at 654±8 hPa."*

As a closing comment only, the authors may wish to point out that the NOy comparison highlights the need for development of more robust conversion schemes or measurement techniques for NOy.

Added as suggested.

**Anonymous Referee #3**

Considering referee#1's comments, this referee (#3) agrees that (a) the total uncertainty for NO and NO2 measurements should be quantified, and (b) it would be nice, if the comparison between the QCLAS and CLD measurements could be expressed in a quantitative matter.

In agreement with this request and Referee #1 comments we added a paragraph to address the issue of total uncertainty and performed a more quantitative analysis (see also our answer to Referee #1) when comparing the two data-sets.

Following referee#2: (a) the specification of the QCLAS physical size (dimensions, weight, power requirements) is missing, and (b) it would be truly fruitful, if the mentioned measurement details about relative humidity could be addressed.

More details regarding instrumental specifications are added. More details regarding the relative humidity is also given (see also the answer to Referee #2).

Page 8980, lines 1–4: „*In order to minimize the scatter due to synchronization differences between instruments, the error of the mean of each individual value was taken as a measure of the variance during that time window and was used to filter out data points which exhibited too large short-term variations*." The procedure is unclear and magnitudes of scatter/error/variance/short-term variations are far from being quantitative.

We re-phrased the statement for more clarity:
*"The calibrated QCLAS data were averaged to 10 min (600 points) to match the time resolution of the CLD and the error of the mean of each individual averaged value was used to filter out those data points which exhibited too large (3$\sigma$ of the CLD detection limit) variations. This minimized the scatter due to synchronization and time response differences between instruments in situations where the NOx concentrations changed abruptly during the averaging time interval."*

Finally, the NOy problem. The referee admits frankly, he doubts basically the value of NOy measurements, due to the obvious conversion problems of the Au-converter which are well known for decades now! ("NOy" should be termed as "NOwhy?"). However, the authors made most likely the best out of their results, however some more quantitative argumentation would be desirable (s.). The authors are encouraged to follow referee#2's recommendation as to make strong statement(s) for the need of the development of more reliable, robust conversion techniques for NOy.

Systematic inconsistencies related to NOy analysis were observed before at JFJ during intercomparison campaigns (see references in text). Our study clearly shows that they are independent of the CLD technique (i.e. of quenching, $O_3$ and $H_2O$ effects, etc.). The observed discrepancies seem to be related to water vapour, however we cannot explain these effects, which may well be related to $HNO_3$ conversion.
See also our answer to Referee#1's comments.