Response to Reviewer #1 comments to "The CU Two Dimensional MAX-DOAS instrument - part 1: retrieval of NO_2 in 3 dimensions and azimuth dependent OVOC ratios" by Ortega et al., 2014.

We thank Referee #1 for the evaluation of our paper, and are responding below to all the comments in detail.

Black: Referee's comments
Blue: Author's reply
Green: sentence added/modified in the manuscript

This paper suffers from two major deficiencies, which need to be addressed before publication

1) There is no independent validation data presented to show that the profile and geographic distribution results represent a real atmosphere.

During the access review a similar point had been raised. In response we added a separate section 4.6 'Challenges with validating 3-dimensional measurements', and referred the reader to our independent validation data shown in Fig. 7a of Sinreich et al., 2013. We have now re-organized Section 4.6, and added results from a recent intercomparison study that has compared the RMT parameterization with in-situ CE-DOAS in homogeneous air (Volkamer et al., 2015). These results are now discussed at the beginning of the revised Section 4.6. The added text reads:

"The RTM parameterization approach to convert dSCDs into near surface VMR used in this study has previously been compared with independent validation data for NO_2 (Sinreich et al., 2013) and glyoxal (Volkamer et al., 2015). We refer to Fig. 7a of Sinreich et al. (2013) that compared the near surface NO_2 VMRs from the RTM parameterization in three different azimuth views by means of two co-located Long Path (LP-) DOAS instruments (facing in opposite directions) in Mexico City. In that particular case, there was a generally good agreement of NO_2 between the two methods, although differences were found when the air mass probed was different for the two LP-DOAS instruments, indicating strong NO_2 concentration gradients in air surrounding the site. For glyoxal, the near surface VMR from Ship MAX-DOAS was compared with in-situ Cavity Enhanced DOAS (CE-DOAS), and airborne MAX-DOAS (AMAX-DOAS) over the remote tropical Eastern Pacific Ocean (Volkamer et al., 2015). In this case, the RTM parameterization was applied to glyoxal dSCDs measured by the Ship MAX-DOAS using EA of +1.5, which is slightly lower, but similar to the EA used in this study. The in-situ glyoxal, ozone and other measurements on the ship are constant over the course of several hours (Coburn et al., 2014), and indicate homogeneous air during this case study. Under such conditions good agreement (within 10%) was observed between in-situ CE-DOAS and Ship MAX-DOAS measurements of glyoxal near surface VMR (see Sect. 3.4 and Fig. 8 in Volkamer et al., 2015)."

Further text was added in response to comments from all reviewers in Section 4.6 to avoid confusion about the intention of comparing with satellite data in Fig. 10, and to make our rationale for further active steps that we have taken to add validation data in Section 4.6.1 transparent:

"The comparison with OMI data in Section 4.5 is not meant as an attempt for validation. Rather the better agreement found for the comparison of data sampled on similar spatial scales poses questions as how to best validate 2D-MAX-DOAS observations (and satellites) under conditions when the air is inhomogeneous. In Section 4.6.1 we compare the NO_2 VMR from OE and parameterization in an attempt to assess different retrieval approaches of dSCDs that were generally measured in the same general direction (fixed AA direction of the standard view), but are nonetheless not identical. Section 4.6.2 then discusses the challenges with validating 3D distributions of NO_2 and other gases in inhomogeneous air that need to be actively addressed to achieve the best results from ongoing and future comparison efforts."

We have added a new sub-section 4.6.1 that compares the near surface NO_2 VMR from the parameterization approach with that derived from optimal estimation. The added text reads:

"The multi-wavelength NO_2 near-surface VMR retrieved with the parameterization approach are compared with those derived from OE for the standard azimuth view in Figure S5. Note that the two approaches are independent, and do in particular not use identical NO_2 dSCD. As pointed out in sections 3.2.3 and 3.3 the NO_2 derived with the parameterization approach uses only data from the lowest EA, and the mixing ratio is characteristic only over the differential effective path length and height in the lower portion of the MLH. By contrast, the OE uses NO_2 dSCDs from a full set of EA. In reality, the comparison of MLH_{NO2} in Section 4.1.2 shows that there is some variability in NO_2 with altitude. The near surface VMR is subject to smoothness and assumptions about a priori profile shape/magnitude and error covariance matrices, and thus differences in Fig. S5 may represent actual gradients in the chemical state of the atmosphere, and should not be interpreted exclusively as an 'error' of the retrieval strategy'. In order to account for the different horizontal range sensitivity we have averaged the NO_2 profiles from figure 6 to a similar representative height of the parameterization approach for each wavelength. In general, both methods capture the NO_2 variability within the spatial scale for each wavelength. The slopes of a linear regression (parameterization versus OE) varied from 0.83-0.88 and intercepts from 0.11-1.76 ppbv NO_2 (see insert on Figure S5). There is further a general good correlation, with the R^2 ranges between 0.55-0.78, for each wavelength considering the difference in sensitivity of air masses probed by the two methods."

Note: In the revised manuscript we adopted terminology consistent with that of Sinreich et al. (2013), however we refer to the NO_2 mixing layer height (MLH_{NO2}) instead of planetary boundary layer (PBL) to avoid confusion with the meteorological term.

We have further added an explanation to Section 4.6.2 to make our reasoning transparent why we did not use surface-based in-situ measurements of NO_2 for validation. The added text reads:

"The regional air quality network in Mainz provides in-situ NO_2 measurements. A comparison with our 3D distributions is not straightforward, and has not been attempted, for at least two reasons: (1) previous comparison of NO_2 and HCHO in-situ observations and open-path spectroscopic measurements show a rather fair level of correlation ($0.32 < R^2 < 0.77$) in ambient air (Harder et al., 1997; Cardenas et al., 2000; Jimenez et al., 2000; Kim et al., 2003; Thornton et al., 2003; Dunlea et al., 2007), while comparisons under controlled conditions show excellent correlation (also for ambient air) if the air is well mixed ($R^2 > 0.98$) (Thalman et al., 2014). The comparison with network sensors thus is unlikely to carry much meaning in light of the different air masses probed. The inhomogeneity along a constant AA is 50-100%, and a lower limit for the

actual inhomogeneity given the averaging nature of 2D-MAX-DOAS measurements. Any better agreement would be coincidental. Moreover, this fundamental sampling problem is not unique to a comparison with in-situ sensors. For example, car MAX-DOAS measurements have the benefit of integrating vertically, but are conducted in air that is offset horizontally from the measurement site of the 2D MAX-DOAS, and is also subject to the inhomogeneity along a constant AA. (2) Additional artifacts can arise with in-situ sensors that measure NO_2 indirectly, e.g., molybdenum converters (as used in the Mainz City network) suffer from interferences of NO_y species that high-bias the sensor signal attributed to NO_2 (Dunlea et al., 2007; Villena et al., 2012). For example, Villena et al. (2012) found that values of NO_2 can be overestimated by up to a factor of four using molybdenum converters. On the other hand, photolytic converters can show even stronger negative interferences than the molybdenum converters, especially under high pollution levels (high NO_x concentrations) (Villena et al., 2012). The sensors used to validate 3D NO_2 distributions should be chosen to actively avoid such potential for bias. In order to validate azimuth distributions of NO_2 a first step consists in comparing the NO_2 SCDs from multiple 2D MAX-DOAS instruments during the MAD-CAT campaign. An according manuscript is currently in preparation (Remmers et al., 2015). The subsequent validation of 3D distributions of NO_2 is non-trivial. In particular, the heterogeneity of NO_2 poses the following challenges to designing a sampling strategy for retrieval validation."

The following Table is added for review purposes, and citations have been added to the text.

Table. Summary of previous studies where trace gases are compared using in-situ vs spectroscopic long path instruments.

Specie	In-situ vs spectroscopic	Findings	Comments	Reference
NO_2	CL vs (TILDAS and LP-DOAS)	Up to 50% positive interference in the CL	Loss of spatial coherence when comparing a long path measurement with a point sampling data for a reactive species	Dunlea et al., 2007
NO_2	MS2 vs LP-DOAS	y = 0.655x +7.437; R^2 = 0.768	Meteorological conditions may be required to properly assess the DOAS performance due to its capacity to cover spatial scale over the open path length	Kim et al., 2003
NO_2	(LIF and PCL) vs LP-DOAS	y = 1.07; R^2 = 0.77 LP-DOAS =f(in-situ)	Comparison is just possible within well mixed environment	Thornton et al., 2003.
NO_2	CL vs LP-DOAS	Fractional differences of +30±110% were observed	The comparison is complicated by the influence of strong sources and large ambient variability.	Harder et al., 1997
HCHO	LP DOAS vs DNPH-sampler	y = 0.78 +1.96; R^2 = 0.32 DOAS vs. the DNPH	DNPH measurements were accepted as true and DOAS normally underestimated the concentration of HCHO	Jimenez et al., 2000.
HCHO	LP DOAS vs (Hantzsch and TDLAS)	y = 0.62; R^2 = 0.45 DOAS=f(Hantzsch)	Large uncertainty in measurements at background locations	Cardenas et al., 2000

TILDAS - Tunable Infrared Laser Differential Absorption Spectroscopy.

CL - Chemiluminescence.
MS2 - conventional point measurement system, MACSAM-2 (Kim et al., 2003)
LP-DOAS – Long Path DOAS
LIF - laser-induced fluorescence
PCL - photolysis to NO followed by chemiluminescence (PCL)
DNPH – chromatographic technique employing C18-DNPH-cartridges (2,4-dinitrophenylhydrazine)

2) The spectral fitting does not demonstrate uniqueness with respect to the wavelength window chosen by the authors. For example, the window for HCHO is a frequently used spectral window, but experience has shown that quite different answers can be obtained with a slightly wider window. The paper needs to independently address the uniqueness issue using the results from the new fairly coarse resolution (0.75 nm) MAX-DOAS instrument.

In the revised supplementary information we have documented the sensitivity in our HCHO fits by systematically varying the spectral window used to fit HCHO, and comparing it to the results discussed in Pinardi et al (2013). In the latest version of the manuscript we added the following statement in section 3.1:

"Sensitivity studies of the spectral window chosen for the fit of HCHO were performed (see Fig. S1), and confirm that this spectral window is stable for different polynomial degrees, and minimizing the residual and cross correlation with BrO."

On the other hand glyoxal is one of the gases for which detailed comparison of glyoxal dSCDs are planned for MAD-CAT, where numerous other instruments had been collocated with our CU 2D-MAX-DOAS instrument.

General Comments

The paper is very well written with generally clear figures. The technique presented is new, and appears to have significant promise for ground-based measurements.

• I think the term solar stray light is meant to be solar scattered light.

The term "solar stray light" has been changed to "solar scattered photons".

• A minor note: Aside from the apparent compactness, this instrument is not "novel". For example, the Washington State University MF-DOAS instrument has all of the same capabilities. At a minimum, this earlier instrument should be referenced.

We have added reference to the MFDOAS instrument (Herman et al., 2009). We note that this reference does not discuss azimuth distributions of trace gases. All the groups/institutions with 2D instruments, including Washington State University, are now named in the introduction section.

To our knowledge there is no previous instrument that systematically looked at 3D distributions. We have changed 'novel' to 'innovative' in the abstract, and refer to a 'novel retrieval' to measure 3D distributions in the text.

• The OMI measurements of NO_2 do not qualify as validation, since OMI always underestimates NO2 amounts in polluted regions. This is true in the case presented in your Figure 11.

This is a misunderstanding, as we do not intend to use OMI to 'validate' our instrument. The point we make in comparing the OMI data (Fig. 11, Section 4.5) is that the NO_2 VCD obtained from azimuth distributions of NO_2 measured at 560nm wavelength represents an area that is more directly comparable to the area probed by the OMI pixel (top panel in Fig. 11). We have modified language in Section 4.5 to avoid any possibility of misunderstanding.

• What is the signal to noise ratio for the various measurements? Was the 30 second exposure an average of multiple measurements or just a single exposure?

The 30s integration time described in the manuscript is the result of multiple measurements recorded by actively controlling the CCD target saturation level. For instance, a typical 3-4s exposure time was needed to reach the 40% saturation level used during MAD-CAT, hence 7-8 scans were acquired within the 30s total integration time. As described in section 4.2.2 the typical root mean square (RMS) residual achieved in the fit of HCHO (and NO_2), CHOCHO (and NO_2), and NO_2 is on the order of $4.5x10^{-4}$, $3.5x10^{-4}$, and $3x10^{-4}$ for the 360, 450, and 560nm fits, respectively limited by the Fraunhofer lines and high NO_2 optical depth for some days. The detection limit (DL) of the NO_2, HCHO, and CHOCHO, calculated with the 1-sigma noise mentioned above, are roughly $1.2x10^{15}$, $8.50x10^{15}$, and $7.0x10^{14}$ in units of molec·cm^2, respectively. The measurements of these gases are normally above the DL.

These aspects are discussed in detail in Coburn et al. (2011). Text was revised as follows:

Section 2.1: "The CU 2D-MAX-DOAS instrument is a further development of the one-dimensional CU MAX-DOAS instrument (Coburn et al., 2011). For a detailed characterization of variations of the instrument line shape with wavelength, temperature, integration times and noise limitations see (Coburn et al., 2011)."

Section 2.3: "These spectrometer/detector units are identical to the ones described in Baidar et al. (2013); see also Dix et al. (2013); Oetjen et al. (2013) for airborne MAX-DOAS applications."

• What is the FOV for direct-sun measurements?

The direct-sun observations are not part of this manuscript (see line 34). A detailed description of the direct sun observations is planned for a follow up paper (part 2) that is focused on aerosol optical properties. We have added reference to this manuscript draft in the revised manuscript.

• What is the slit function shape as a function of wavelength?

The reviewer is referred to Coburn et al. (2011) Fig. 2.

• Does the fiber optic cable move? If so, what is the sensitivity to motion? Motion effects would seem particularly important for the large range of motion for this instrument.

The optical monofiber attached to the 2D telescope does not move; only the azimuth component moves. We have included a sentence in Fig 1. to make it clear:

"In order to avoid mechanical stress in the optical fiber and allow free mobility of the telescope the optical fiber is attached in the lowest level of the 2D telescope and does not move."

• The detector is 1340x400 pixels, suggesting that there are more than enough pixels (~7 and 9) to determine the slit function as a function of wavelength. What was the slit function and how did it compare to the assumed slit function in the retrieval program (usually Gaussian).

We do not make an assumption about the slit function, but rather use the measured instrument line shape of a nearby atomic emission line as is described in the manuscript, as is stated on section 2.3 (line 23).

• NO_2 and HCHO are highly variable in time. How does the 13-14 minute duration affect the azimuthal distribution results? What kind of error is introduced? What is the variation in time for continuous measurements in one direction?

We have added a new Figure S5, which shows the time series of NO_2. The current time resolution (14 min) of the measurements seems to capture the temporal variability, which are also visible in the dSCD of both gases. The 14 min time resolution for a complete 360˚ azimuth view is deemed sufficient to document changes in the atmospheric distributions of NO_2, HCHO, CHOCHO that occur on this time or longer timescales.

• How sensitive are you to the assumed P and T profiles (US Standard Atmosphere) compared to the real P and T profiles at the observing site? Equation 7 should be quite sensitive to the P and T profile assumptions.

We have conducted a set of sensitivity studies, and found that small changes of pressure and temperature relative to the US standard atmosphere have small effects on the RTM, in agreement with other studies (Clémer et al., 2012; Friess et al., 2006). The error in the forward model parameters, including P and T, are weighted in the error budget of section 4.2.2 and the new table below:

"Table 6 shows a summary of the errors for both optimal estimation and parameterization."

Method	Error overview (%)	Total error (%)

Table 6. Overview of the most important errors in the determination of NO_2 vertical profiles.

- Figure 6…"The PBL height is estimated as the 1/(2e) decrease of the near surface NO2 VMR at each wavelength." Does this mean that you have assumed an exponential profile shape? Frequently this is not the case. How does your assumption affect the PBL height calculated?

We do not assume an exponential profile shape, but rather use the measured profile shape a-posteriori to determine the height where the NO_2 concentration has decreased to 1/2e. This has been clarified in the revised manuscript. Also, we use terminology consistent with that of Sinreich et al. (2013), however we refer to the NO_2 mixing layer height (MLH_{NO2}) instead of planetary boundary layer (PBL) to avoid confusion with the meteorological term.

We have further tested sensitivity to the MLH_{NO2} using an independent approach that assumes a NOx box-profile, assigns the near surface VMR from an optimal estimation retrieval, and assigns an effective MLH height based on the measured NO_2 VCD using the following equation:

$$VCD_{NO_2} \approx VMR_{NO_2} \cdot MLH_{NO2,eff} \cdot CF$$

where VCD_{NO_2} has units of molec·cm^{-2}, VMR_{NO_2} is in ppb, $MLH_{NO2,eff}$ has units of cm, and CF is the conversion factor between mixing ratios and molec·cm^{-3}. The retrieved $MLH_{NO2,eff}$ was compared with the 1/2e decrease of the actual profile shape described in the manuscript. We found a good correlation (R^2=0.81) with a linear fit of $MLH_{NO2,eff} = 0.93 \pm 0.06 \cdot MLH_{NO2} + 0.10 \pm 0.04 \, [R^2 = 0.81]$. The resulting MLH$_{NO2,eff}$ is within the error near the lower limit of the standard deviation calculated with the multi-wavelength retrieval of the MLH$_{NO2}$, which is used to calculate the variability associated with the MLH$_{NO2}$ as mentioned in the revised manuscript. We have added a short description of this comparison exercise in section 4.1.2:

"For the calculation of f$_c$ the time resolved average MLH$_{NO2}$ obtained with the three wavelengths is used and the standard deviation is used to estimate its variability. The sensitivity of the height distribution of NO_2 on f$_c$ was tested with an independent method. This approach assumes a NO_x box profile shape, assigns the near surface NO_2 VMR from OE, and assigns an effective mixing height (MLH$_{NO2,eff}$) based on the measured NO_2 VCD. The retrieved MLH$_{NO2,eff}$ was compared with the MLH$_{NO2}$ calculated before. We found a good correlation (R^2=0.81) with a linear fit of $MLH_{NO2,eff} = 0.93 \pm 0.06 \cdot MLH_{NO2} + 0.10 \pm 0.04 \, [R^2 = 0.81]$. The resulting MLH$_{NO2,eff}$ is within the error near the lower limit of the standard deviation, which is used to calculate the variability associated with the MLH$_{NO2}$ as mentioned before."

•　　　　Figure 3C… It is not clear that the fit to CHOCHO is good except for the peak near 455 nm

In the revised manuscript we have changed the fit examples from 17 June, 2013 (the actual date of measurements).

•　　　　Figure 1B… What keeps the direct-sun light from scattering off the edges of the prism after coming out of the integrating sphere? Are there lenses and a collimator that are not shown?

We have included the following sentence in Figure 1B:

"A black anodized collimator tube (not shown) is inserted in the prism hole to avoid scattering off the edges of the prism after coming out of the integrating sphere."

Equations:

Equation 7 (line 404) does not seem to make sense….First, what does the symbol O4 mean?

When combined with eqn. 6 the Leff seems to be just the ratio of the differential AMFs times the vertical PBL height. I do not see how this is related to the horizontal effective path length. The authors need to explain this in more detail.

Thanks for catching this typo. It has been corrected.

References added:

Cárdenas, L. M., Brassington, D. J., Allan, B. J., Coe, H., Alicke, B., Platt, U., Wilson, K. M., Plane, J. M., and Penkett, S. A.: Intercomparison of formaldehyde measurements in clean and polluted atmospheres, J. Atmos. Chem., 37, 53–80, 2000.

Dunlea, E. J., Herndon, S. C., Nelson, D. D., Volkamer, R. M., San Martini, F., Sheehy, P. M., Zahniser, M. S., Shorter, J. H., Wormhoudt, J. C., Lamb, B. K., Allwine, E. J., Gaffney, J. S., Marley, N. A., Grutter, M., Marquez, C., Blanco, S., Cardenas, B., Retama, A., Ramos Villegas, C. R., Kolb, C. E., Molina, L. T., and Molina, M. J.: Evaluation of nitrogen dioxide chemiluminescence monitors in a polluted urban environment, Atmos. Chem. Phys., 7, 2691-2704, doi:10.5194/acp-7-2691-2007, 2007.

Harder, J. W., Williams, E. J., Baumann, K., and Fehsenfeld, F. C.: Ground-based comparison of NO2, H2O, and O3 measured by long-path and in situ techniques during the 1993 Tropospheric OH Photochemistry Experiment, Journal of Geophysical Research: Atmospheres, 102, 6227–6243, doi:10.1029/96JD01729, http://dx.doi.org/10.1029/96JD01729, 1997.

Jiménez, R., Martilli, A., Balin, I., v. d. Bergh, H., Calpini, B., Larsen, B. R., Favaro, G., and Kita, D.: Measurement of Formaldehyde (HCHO) by DOAS: Intercomparison to DNPH Measurements and Interpretation from Eulerian Model Calculations, Proceedings of the 93rd Annual Conference & Exhibition, Air & Waste Management Association, Paper #829, Salt Lake City, Utah, June 18–22, 2000.

Kim, K.-H. and Kim, M.-Y.: Comparison of an open path differential optical absorption spectroscopy system and a conventional in situ monitoring system on the basis of long-term measurements of SO2, NO2, and O3, Atmospheric Environment, 35, 4059 – 4072, doi:http://dx.doi.org/10.1016/S1352-2310(01)00216-3, 2001.

Remmers, J., Wagner, T., and the MADCAT team.: Azimuthal variability of trace gases and aerosols measured during the MADCAT campaign in summer 2013 in Mainz, Germany: EGU General Assembly Conference Abstracts. Vol. 17, EGU2015-5018, 2015.

Thalman, R., M.T. Baeza-Romero, S.M. Ball, E. Borrás, M.J.S. Daniels, I.C.A. Goodall, S.B. Henry, T. Karl, F.N. Keutsch, S. Kim, J. Mak, P.S. Monks, A. Muñoz, J. Orlando, S. Peppe, A.R. Rickard, M. Ródenas, P. Sánchez, R. Seco, L. Su, G. Tyndall, M. Vázquez, T. Vera, E. Waxman, and R. Volkamer, Instrument inter-comparison of glyoxal, methyl glyoxal and NO2 under simulated atmospheric conditions. *Atmos. Meas. Tech. Discuss.,* 7, 8581-8642, 2014, accepted for *Atmos. Meas. Tech.* (27 March 2015), in press.

Thornton, J. A., Wooldridge, P. J., Cohen, R. C., Williams, E. J., Hereid, D., Fehsenfeld, F. C., Stutz, J., and Alicke, B.: Comparisons of in situ and long path measurements of NO2 in urban plumes, J. Geophys. Res., 108, 4496, doi:10.1029/2003JD003559, 2003.

Villena, G., Bejan, I., Kurtenbach, R., Wiesen, P., and Kleffmann, J.: Interferences of commercial NO2 instruments in the urban atmosphere and in a smog chamber, Atmos. Meas. Tech., 5, 149-159, doi:10.5194/amt-5-149-2012, 2012.

Wagner, T., Beirle, S., and Deutschmann, T.: Three-dimensional simulation of the Ring effect in observations of scattered sun light using Monte Carlo radiative transfer models, Atmos. Meas. Tech., 2, 113–124, doi:10.5194/amt-2-113-2009, 2009.