We thank referee #1 for taking the time to review our paper and for the many comments about details that we feel have helped to improve the paper. Detailed below are our answers.

GENERAL COMMENTS

Like in previous studies related to airborne scattered sunlight DOAS measurements, the work of Baidar et al. addresses the relevance of an accurate knowledge of the viewing geometry and further exploits the implementation of a motion compensation system on an AMAX-DOAS instrument to gain the maximum sensitivity at the layer where the aircraft flies. After describing the instrument, the authors present DOAS measurements performed over California in 2010. The authors do calculate VCD below the aircraft from the nadir viewing geometry (with geometrical approximation) and compare it to ground based measurements, and invert trace gas vertical profiles from the limb viewing geometry with RTM calculations based on O4 dSCDs measurements. Note that the motion compensation system embedded in the instrument represents a useful implementation to the experimental setup of the air-borne DOAS family (e.g. Petroli et al., Appl Optics, 2002; Bruns et al., Appl Optics, 2004; Weidner et al., ACP, 2005; Dix et al., AMT, 2009). Thus in fact the instrument presented in this work is not a novelty in itself. The adjective "unique" should be removed or modified in the manuscript.

We respectfully note that none of the references cited by the reviewer features a motion compensation system. Notably, the adjective 'unique' describes something that is existing as the only one of as the sole example. This applies to our instrument, as it is the only existing AMAX-DOAS instrument that features a motion compensation system.

We have modified the abstract to clearly associate the word 'unique' with the two features of our instrument. To do this, we have removed 'the first systematic implementation of MAX-DOAS on research aircraft' to make this link as direct as possible.

The paper is generally well written although, for clarity, the results should be divided into "nadir measurements" (with geometrical approximation and its validation) and "nonadir measurements" (vertical profile inversions based on O4 and RTM and WITHOUT validation). Hence the title of the manuscript should be clarified since the "ground based validation" it is only for the case of the nadir viewing measurements (which in fact is not a novelty). The work of Baidar et al. should be published in AMT only after addressing the following comments.

As suggested by the reviewer, we have divided the results into two sections: 1. Nadir Observations and, 2. Limb Observations. The new section number reads as follows:

4. Results
- *4.1 Nadir Observations*
  - *4.1.1 Horizontal distribution*
  - *4.1.2 Validation of NO2 vertical column*
- *4.2 Limb Observations*
  - *4.2.1 Determination of O4 SCD in the reference spectrum*

4.2.2 Aerosol extinction profiles
4.2.3 Trace gas vertical profiles

We have further removed 'ground based validation' from the title.

SPECIFIC COMMENTS

Page 7243 (Title): The DOAS community already suffers a very broad amount of abbreviations (LP-DOAS, MAX-DOAS, CE-DOAS, mini-DOAS, AMAX-DOAS, etc) mostly related to the measurement technique. Thus, since the relevant matter in this work is the measurement technique (MAX-DOAS), CU MAX-DOAS should be simplified into MAX-DOAS throughout the manuscript. Otherwise the community would end up with so many different MAX-DOAS as institutes or departments, when indeed we are talking of the very same measurement technique. The procedence of the instrument is to be considered in the main text, not on the title. The AMAX-DOAS instrument is validated through ground based NO2 comparison only for the nadir viewing geometry. No validation of any trace gas vertical profile is in fact presented ("ground base validation" should be "VCD validation").

We agree, and note that there is no new acronym for a technique introduced in this paper. The topic of this methods paper is a new instrument, and the name of this instrument is the CU AMAX-DOAS instrument. We have carefully checked the text, and have made sure that we refer to the technique as AMAX-DOAS, and to our instrument as 'the CU AMAX-DOAS instrument'.

Regarding the 'validation': Please see above.

Page 7244, line 6: Aren't the trace gases retrieved in a spectral window rather than at a single wavelength? What is the meaning of "360 nm, 477 nm, 577 nm and 632 nm"?

Trace gases are indeed retrieved in spectral windows rather than at a single wavelength. Here, we are referring to the O_4 absorption bands with peak absorptions at the above mentioned wavelengths. We have added text in the manuscript to make it clearer.

Page 7244, line 7: what do you mean with "sensitively"?

The word has been removed.

Page 7244, line 8: the instrument is an AMAX-DOAS, therefore it is not "unique" but maybe "an improved" AMAX-DOAS.

Please see our response to the general comments.

Page 7244, line 14: How are the "2 km above and below the aircraft" estimated? Explain.

Box AMF calculations for EAs 5 and 10 ˚above and below the horizon show that the sensitivity of the two EAs falls below 50% of the peak value above and below 2 km of the aircraft. EAs 5 and

10 above and below the horizon are the largest limb scanning EAs used during flights. We have added a description of these criteria in Section 3.5 to make our statement in the abstract transparent.

Page 7244, line 27: How can the profiles be "independent of the signal-to-noise at which the trace gas is detected"?

The particular sentence has been removed from the manuscript. See also our response to comment #2 of the other reviewer.

Page 7245, line 16: Missing reference to the so called mini-DOAS stratospheric balloon instruments (e.g., Weidner et al., ACP, 2005 and references therein). Note that they performed thorough sensitivity studies on the forward parameters in profile inversion as well as their propagation into the radiative transfer models.

Indeed an oversight. We have added the reference to the manuscript.

Page 7246, line 9: the phrase "The DOAS technique. . ." should be moved to Sect. 3.1.

Done.

Page 7246, line 14: "larger" than? Explain.

We have modified the sentence to read: 'Pitch and roll angle of the aircraft is used during post-processing to calculate the range of EAs during the time of measurements. This can lead to reduced sensitivity to a given layer in the atmosphere, compared to when the EAs are actively controlled, for the vertical profiles retrievals.

Page 7246, line 18: All AMAX-DOAS are implemented in aircrafts by definition. The sentence "presents the first true AMAX-DOAS implementation..." should be removed or rephrased. All AMAX-DOAS used in previous works are also "true" AMAX-DOAS.

Please refer to response to the general comments.

Page 7248, lines 9-10: Is the Hg lamp also used for calibration? If so, how often were the spectra calibrated?

No, the calibration of the wavelength stamp of spectra is performed using Fraunhofer lines. Hg spectra were taken before, after and at regular intervals during flights to establish the stability of the spectrograph's optical resolution.

Page 7249, line 6: A sketch of the telescope system as well as the optical fiber switch box would ease the understanding of the instrumental setup.

We have modified Fig. 2 to show the telescope system, and the optical switch box.

Page 7249, line 24: I understand that the retrieval of O4 does not require such a high resolution system as does the retrieval of the other trace gases presented in the manuscript. However, the "TG" spectrometer-detector system presented in the manuscript would also be valid for retrieving O4 (at least the 360 nm absorption band). Considering the space/weight/power limitations usually linked to most aircraft deployments, what would be the reason for adding a spectrometer (i.e., needing thus 2 racks) dedicated only to measure O4 in an aircraft deployment? Possible deployment issues on different aircrafts and/or campaigns could be spared if only one spectrometer (1 rack) is used. Additionally the fiber setting would also be simplified. The only requirement in case the O4 absorption band at 477 nm would be needed is a wider spectral coverage of the TG spectrometer.

We agree, but have a hard time identifying a question here. Our instrument fits into 1 rack, and the trade-off in spectral range and optical resolution lead us to the design as described.

Page 7250, line 9-10: Why the filters are only considered for the TG spectrometer?

The filters in the TG spectrometer are used to suppress spectrometer stray light from longer wavelengths (above 470nm). The O_4 spectrometer measures a much broader spectral range (up to 720nm), and the decreasing Rayleigh scattering cross sections at longer wavelengths effectively act to suppress spectrometer stray light from longer wavelengths.

Page 7253, line 4: Introduce the reason of interest for measuring NO2, HCHO, CHOCHO and aerosols.

Text regarding the motivation to measure these species has been added to the manuscript.

Page 7254, line 8: What is the principle of the DOASIS software for calculating the Ring cross-section? How does that software calculate the Ring spectrum? It doesn't appear in the associated reference.

The Ring calculation in DOASIS is based on the work of M. Bussemer (1993). The appropriate reference has been added to the manuscript.

Page 7254, line 14: Not only "excess". It is a difference no matter the direction of the variance.

Agreed. "Excess" has been changed to "difference" in the manuscript.

Page 7254, line 27: Shouldn't be Eq (1) dVCD instead of VCD? Explain the difference.

We are referring here to the tropospheric VCD, as the primary purpose of the reference spectrum is to cancel out the stratosphere. We have added 'tropospheric VCD' to make this clear.

Page 7255, line 8: Assuring that an error is "very high" is quite ambiguous. Define "very high" in this case.

We have removed 'very high' as it is not needed to convey our point.

Page 7256, line 26: How is the value of the aerosol asymmetry parameter determined?

We used the typical urban values for asymmetry parameter (g=0.68) in urban atmospheres (Dubovik et al., 2001). This value is within the range of aerosol asymmetry parameter measured at AERONET station at the CalNex ground site, Pasadena, CA (0.63-0.80) for the entire campaign. Text has been added to section 3.2 for clarity. AMFs calculated for this range of g values were within 3% of AMF relative to g=0.68.

Page 7257, line 21: "higher" than?

Modified to 'high SA' as we do not mean to make a comparison.

Page 7257, line 27: It depends on the extinction of the atmosphere due mostly to aerosols and clouds. Are clouds considered at all? If so, please explain how.

Flights were mostly performed on cloud free days. We had two webcams to access the atmospheric state (occasional clouds) during the time of the measurements. There were no clouds during this particular flight. We have added text to the manuscript to make it clear that only cloud free conditions are considered.

Page 7258, lines 14-16: The phrase "For 0_ EA. . .. . ." is only true in the case of Rayleigh atmosphere and single scattering approximation. What happens when the aircraft crossed a cloud or an aerosol layer?

In light of the above, we do not see this question has a relation to our manuscript.

Page 7258, line 26: That approach also assumes the atmospheric layers to be horizontally homogeneous. How can you assure that in the case of a polluted environment with several emission sources?

Our measurement inherently averages over few 10 km and retrieved values are averages over those distances as stated on p.7265. This is a general issue with MAX-DOAS observations that is not particular to our instrument, but certainly a concern.

Page 7259, line 3: Actually above 10 km, what is the amount of O4 in the atmosphere? Would that 20-80 % error be due to the measurement error or to the pointing accuracy error?

20-80% error is only due to the error in pointing accuracy. The O_4 concentration is very small above 10 km (around 10%), but the SCD at EA0 can still be measured at good signal-to-noise, because of the much longer photon path lengths in the less dense air. The high sensitivity towards the pointing accuracy is therefore the limiting source of error. This has been added to the manuscript.

Page 7259, line 19-22: E.g., Kritten et al., 2010

Reference has been added to the manuscript.

Page 7260, line 1: Don't you measure dSCD and not SCD?

Yes, we do measure dSCD, but considering that we use a high altitude clean environment zenith reference spectrum, SCD of the reference is / can be considered to be negligible for tropospheric pollutants. See also our above comment about equation (1).

Page 7261, line 2: For clarity, the results should be divided into nadir measurements (with geometrical approx and its validation) and no-nadir measurements (profile retrieval based on O4 and RTM and without validation). Thus the title of the manuscript should be clarified since the "ground based validation" it is only for the case of the nadir viewing measurements (which in fact is not a novelty). Note that, for instance, the Sect 4.3 "determination of O4 SCD in the reference spectrum" is only relevant for the no-nadir measurements.

See above, and also our response to the other reviewer.

Page 7261, line 21: Again, why keep on adding abbreviations to the original name? The origin of MAX-DOAS is a ground based instrument. Why GMAX-DOAS? Please try not creating new acronyms when they are indeed not new instruments and/or techniques.

Good point. The manuscript refers to ground based MAX-DOAS as MAX-DOAS now.

Page 7262, line 4: Given the measurement error of MAX-DOAS instruments at 20_, how is that VCD 10% error estimated? Please explain the error propagation throughout all the calculations in the manuscript.

We are not sure where the reviewer finds 20% as error for MAX-DOAS? The measurement error of the MAX-DOAS dSCDs is ~3%. MAX-DOAS measurements were taken at 20 degree elevation angle, which the reviewer might have mistaken to be the measurement error. The error calculations are specified in a separate manuscript currently in preparation (Ortega et al., 2013).

Page 7262, line 18: Detail the cloud filter used.

See above.

Page 7262, line 29: How is then the validity of the geometrical approach affected under horizontally non-homogenous atmospheric layers?

Comparison with VCDs retrieved from ground-based MAX-DOAS measurements shows excellent agreement when both instruments are looking at the same air mass (azimuth constraint improves correlation between two instruments). Hence, the geometric approach is valid under

the given horizontally non-homogeneous conditions. We have made this clear in the updated manuscript.

Page 7263, line 17: the temperature dependence of O4 cross-section was actually reported much earlier. Correct reference.

We modified the manuscript to read '(Blickersdorfer and Ewing, 1969; Wagner et al., 2002 and references therein)'.

Page 7264, line 7: where does that regression appear?

The regression plot is shown in Fig 9.

Page 7264, line 23 (Fig. 10): How is the error propagation in the case of the aerosol retrieval?

An error propagation section has been added to the manuscript. See also our response to the other reviewer.

Page 7264, line 23: Is the retrieval performed for one single type of aerosol? If so, how realistic is that simplification?

What type of aerosol does the reviewer mean? Aerosols rarely exist as external mixtures, nor are they monodisperse in their size distributions. As such the atmosphere always contains multiple types of aerosols from a variety of sources. These are averaged into effective radiative properties, such as the extinction, or the g-parameter, which integrate over size distribution, refractive indices, hygroscopic volume mixing rules, chemical composition etc. As such, there is no simplification in our calculations.

Page 7265, line 10: How are those "10 km" estimated? That value depends on the wavelength, altitude and aerosol conditions. Clarify.

Based on the Koschmieder visibility formula for the visible wavelength region, visibility = 3.91/extinction coefficient, the visibility at 477 nm during the time of our measurement is ~39 km. Hence we believe it is appropriate to assume that our measurement in the forward direction averages over "few 10 km". See also Fig. 8 in Sinreich et al., (2012), where the horizontal and vertical extent of MAX-DOAS is quantified for Mexico City.

Page 7266, line 13: "aerosols as a source of glyoxal". Is there any experimental work in literature confirming that suggestion or the opposite? Please, provide some references supporting that statement.

We added two references (Kroll et al., 2005; Kampf et al., 2013) that discuss reversible aspects of glyoxal partitioning to aerosols.

Page 7266, lines 16-17: Is that elevated O3 layer considered in the RTM calculations for the aerosol and for the trace gas inversion?

Yes, O_3 measured by the NOAA TOPAZ lidar has been considered in the RTM calculations for the aerosol and trace gas inversion (see section 4.3).

Page 7266, line 22: How can the capability of inverting vertical profiles be independent of the signal-to-noise of the measurements when in fact the error of the measurement is part of the formula for the mathematical inversion? Please, clarify.

This sentence has been removed from the manuscript.

Page 7267, line 1: quantify "small".

We have quantified 'small' as relative percent change in the manuscript.

Page 7277, Table 1: The cross-sections used in this work are not updated. Please quantify the effect of using the latest HCHO (K. Chance and J. Orphal, 2011) and H2O (HITRAN 2009) on the retrieved dSCDs. Please, explain how you correct for the temperature dependence of H2O in the retrieval. Why don't you use only one O4 cross-section?

The mean difference in HCHO dSCD retrieved using Chance and Orphal, 2011 and Muller and Moograt, 2000 is 6e14, which is well below the DOAS fit error for our retrieval. For a mean dSCD of 3.0e16 for EA0 at all flight altitudes for one exemplary flight, the difference is only 2%.

The reason we use different O_4 cross sections is because the Hermans O_4 cross-section has some non-physical structures in the baseline, which affect the fitting of the weak O_4 absorption band around 440 nm, and which subsequently seems to affect NO_2 and glyoxal fits. The non-physicality of these structures has been verified by our laboratory. See our detailed response to the other reviewer. Hence, the O_4 cross-section by Greenblatt was used for those two trace gas retrievals. In the future, we will be able to use our own O_4 cross sections (Thalman and Volkamer, 2013, in preparation).

We do not see any systematic residual structures that would motivate correction for temperature dependence of H_2O (see Fig. 4).

Page 7278, Table 2: How are the "2 Rings" calculated? Please detail the why the Ring cross section is modeled as well as the influence of including 1 or 2 Rings in the HCHO retrieval. Why the 2 Rings were only included in the HCHO retrieval and not in the retrieval of the other trace gases? Since two NO2 cross-sections were included in the retrieval of NO2, how do you interpret the resulting dSCDs when in fact you are orthogonalizing them? Which dSCDs were used for the retrieval of NO2 profiles?

The 2 Ring spectra were calculated for two different temperature (298 and 230 K) using the MFC software (Bussemer, 1993). They were included to account for the temperature dependence

of Raman scattering, which leads to a so-called "filling-in" of the Fraunhofer lines (Ring effect) that the inclusion of Ring spectra accounts for. The inclusion of a second Ring cross section in the HCHO retrieval (1) improved the HCHO fit, (2) minimized fitting residuals and (3) reduced scatter in the retrieved HCHO slant columns. A second Ring cross-section is not needed to achieve comparable results for other trace gases, where the Ring structures are relatively smaller. The two NO_2 cross-sections were not orthogonalized and hence we used the sum of 2 NO_2 dSCDs for the retrieval of NO_2 profiles.

Page 7279, Table 3: How is the detection limit calculated from the DOAS measurements (formula)? The detection limit for H2O is missing. É˙Z360 is not shown anywhere in this work. Either include results or exclude from table.

Detection limit was calculated as 1 sigma RMS detection limit for the aerosol scenario presented in the manuscript. Detection limit for H_2O has been added. The profile for aerosol extinction at 360nm has been added in a new supplementary file.

Page 7280, Table 4: Please define "most probable conditions" and explain the error propagation.

The definition of "most probably conditions" has been described in section 3.3 in the original manuscript. We have also added it as a footnote to Table 4 now. The reported errors in Table 4 are errors in AMF for the individual atmospheric parameters, not propagated error. Note that the Table 4 has the character of a sensitivity study.

Page 7281, Table 5: What would then be the total error in the aerosol extinction coefficient? How would that error propagate into the trace gas retrieval?

We have added the section "Error Analysis" to the manuscript describing error propagation.

Page 7289, Figure 8: How do you calculate VCD from dSCDs from the ground-based MAX-DOAS. How does the VCD calculated with the nadir measurements compare with the VCD after the inverted mixing ratio NO2 vertical profile?

Please refer to section 4.2 for the details on ground based MAX-DOAS AMF calculation.

Simple linear regression analysis between geometric VCDs and VCDs obtained using AMFs calculated with the retrieved NO_2 and aerosol extinction profiles for this particular flight show excellent agreement (R^2=0.99, slope = 0.96, statistically insignificant intercept).

Page 7291, Figure 10: The authors do infer an aerosol vertical distribution based on the O4 dSCD retrieved at 477 nm (Fig. 10) and use it for the inversion of the vertical profiles of several trace gases. Is the inversion of HCHO vertical profile based on the aerosol extinction coefficient at 477 nm or at 360 nm? If the earlier, an estimation of the wavelength dependence (and error propagation) of the extinction coefficient would be needed. If the later, an additional figure of the aerosol extinction profile at 360 nm (similar to Fig. 10) would assist to understand the

retrieved HCHO. Please explain how the error bars are calculated in the extinction profile. Please specify the asymmetry parameter and single scattering albedo used in the RTM as well as the effect of changing those parameters on the retrieved aerosol profiles (and hence on the traces gas profiles). Do you consider only one type of aerosol? Do you distinguish between aerosols and clouds?

The HCHO vertical profile is based on the aerosol profile at 360 nm (now included in the supplemental section).

For the error bars in the extinction profiles, the reviewer is referred to the new Section 'Error analysis'. See also our response to the other reviewer.

The other questions had already been addressed above.

Page 7292, Figure 11: Showing the dSCDs as well as the AMF plots would assist the reader. Please include at least the AMF calculations. The FWHM of the averaging kernels seem to be of 0.5 km. That would be therefore the vertical resolution of the measurements, isn't it?. Please, define the a priori covariance considered on the inversions. Include detection limits of the shown profiles. As commented before, what aerosol profile is used in the inversion of the HCHO vertical profile? É˙Z360 or É˙Z477? Why is the water mixing ratio given in %? How do the authors go from mixing ratio to relative humidity? Where IO and BrO detected during the flights?

Does the reviewer mean weighting functions (instead of AMF)? We have added the weighting functions for glyoxal as an additional figure in the supplementary file. The weighing functions look very similar for the other gases. We do not show the dSCDs, because the vertical profile inversion uses data from various EAs, and thus we feel that the figure does not add anything useful.

We have verified that the FWHM is in fact 0.25 km and not 0.5 km.

Please refer to the answer to the previous comment regarding the aerosol profile at 360 nm.

The unit for water is volume mixing ratio and not relative humidity to be consistent with the volume mixing ratio unit used for all other profiles.

IO and BrO were not the focus of this study.

TECHNICAL CORRECTIONS

Page 7244, line 3: please delete "remote sensing" from the phrase "uses solar stray light remote sensing to detect"

Changes have been made as suggested.

Page 7244, line 6: change “dimmers” to “dimer”.

Changes have been made as suggested.

Page 7244, line 26: correct “are” to “is”.

Changes have been made as suggested.

Page 7244, line 27: the extinction coeff. at 360 nm is not presented in this work. Introduce or remove.

See above.

Page 7246, line 3: change “only one” by “the”

Changes have been made as suggested.

Page 7246, line 10: Reword “These instruments also lack. . ..” By e.g. “However these instruments lack. . .”

The sentence has been reworded as suggested.

Page 7246, lines 21-24: The phrase “The motion component. . ..” is grammatically wrong. Rephrase.

The sentence has been split into two sentences.

Page 7248, line 4 and line 16: Delete “on the”.

Text has been deleted.

Page 7254, line 2: Start a new paragraph for introducing VCD.

As suggested, introduction of VCD now starts with a new paragraph.

Page 7255, line 1: define dAMF.

Definitions of AMF and dAMF have been added to the manuscript.

Page 7283, Figure 2: (E), (F) and (G) cannot be found in the picture.

We have added picture of the optical fiber switch (E). (F) and (G) are commercial Hg calibration lamp and dell laptop and their pictures can be found easily available over the internet.

Page 7285, Figure 4: Include the RMS plot of each fit, the units of the retrieved dSCDs as well as the SZA of the measurements.

RMS plots have been included in the figure.

Page 7286, Figure 5: Please include “nadir measurements” in the capitation.

The phrase has been included in the figure caption.