Dear referee #1,

We appreciate your comments to improve the manuscript. Below you will find our response in blue and modifications to the manuscript in green.

General comment:

Selection of zenith reference spectra, calculation of *dBox-AMFs* and definition of sensitive range *S*. Since the method described here depends on the selection of the reference spectra it will be interesting if the authors could provide some extra information to what is shown in figure 1. Could Figure 1 be expanded, maybe in the supplementary material, showing the thickness of *S* in function of instrument altitude and zenith spectra height? How does the shape of *dBOX-AMFs* change with these 2 factors? Why is S lower boundary, *nL*, defined 1km below instrument altitude? How does it change when instead a zenith reference spectra a reference spectra with EA 10 is used?

This is a good point. We have added a supplementary figure showing the effect of instrument altitude and reference choice on *S* and explain the choice of upper and lower boundaries in the revised manuscript.

p4, line 13: Initial sensitivity studies have shown that the parameterization method typically works best when about 90 % of the sum over the dBox-AMF trace is included in *S*. Therefore the lower boundary n_L is set to 1 km below the instrument layer, n_{instr}, while n_U is set to the altitude layer before the difference between two consecutive dBox-AMFs is smaller than 10 %, and no more than 3.5 km above n_{instr}. Due to the distinct shape of the dBox-AMF (Fig. 1), the placement of the lower boundary is less critical and remains fixed, while the placement of the upper limit is more flexible. The dependency of *S* on altitude, reference and wavelength is shown in Fig. S1.

SI text:

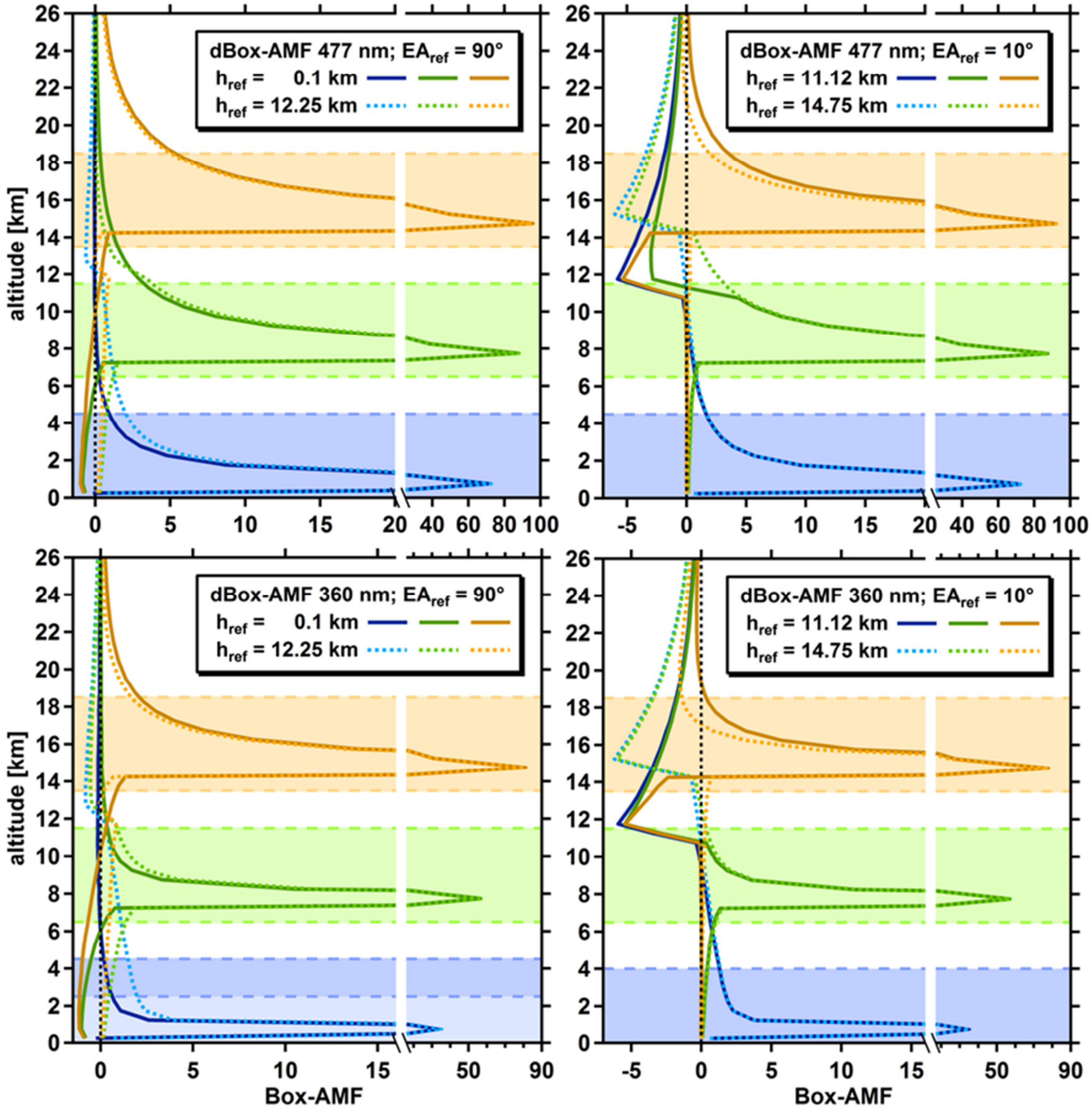

Figure S1 DBox-AMF traces at 477 nm (top) and 360 nm (bottom) nm for 0.75 km, 7.75 km and 14.75 km aircraft altitude and different references, calculated at solar zenith angle of 10°. Respective sensitive ranges are indicated by color shadings. Note that at 360 nm and 0.75 km, the upper boundary layer is at 1.5 km above aircraft altitude when using an EA 90° reference with a reference height, h_{ref}, of 12.25 km, and 3 km above for an EA 10° reference. Both n_U layers are below the 3.5 km limit. The more flexible upper limit accounts for different dBox-AMF peak shapes that depend on wavelength, altitude and choice of reference.

Specific comments:

Page 3, line 12. It will be interesting if the authors could provide a quantification of the speed up factor to be expected between parameterization retrievals vs. optimal estimation.

Typically OE is more time consuming, because the retrieval of the aerosol extinction profile, either iteratively or by non-linear inversion, requires multiple RTM runs of the same data set, whereas parameterization runs the RTM only once for the Rayleigh case. Actual processing times though very much depend on individual circumstances, e.g. the method used for aerosol retrieval or the type of RTM. Our RTM McArtim needs on the order of 5-15 minutes for one profile and it often takes more than 10 iterations to determine the aerosol extinction with the desired accuracy. Running the RTM for a complete flight in a Rayleigh atmosphere takes 5-6 hours, while running the actual parametrization on the dSCDs just takes a few seconds. More so than time savings, the fundamental benefit of the parameterization method is the ability to derive VMRs for each dSCD measurements, and along the complete flight track instead of only for profiling flight legs. In order to make this more transparent we have added text in several sections of the manuscript.

p.3, line 12: …, but it is typically time consuming, computationally intensive and requires vertically resolved SCD measurements as input.

p.28, line 13: Figure 9 shows BrO and IO VMR_{para} results along flight tracks for all 17 TORERO RFs (but RF08).

p.29, line 2: The parameterization retrieval is a robust tool to convert AMAX-DOAS EA0° dSCDs of BrO, IO and NO_2 directly into VMRs along a flight track.

Page 9, line 12. Is it there a publication presenting BrO and IO TORERO measurements? Please cite.

The parameterization results are created by this study and the reference has been added. TORERO BrO and IO optimal estimation results are published in Volkamer et al., 2015 and Wang et al., 2015 as had been mentioned in the manuscript.

p.9, line 12: BrO and IO c-profiles are smoothed TORERO campaign averages for the tropics based on initial parameterization runs (this work).

Page 9, line 17. Please provide information about the origin (model, measurements) of stratospheric columns and aerosol profiles for the sensitivity studies.

Respective information has been included.

p.9, line 17: Stratospheric profiles for BrO and NO_2 are based on Real-time Air Quality Modelling System (RAQMS) (Pierce et al., 2003, 2007) for the TORERO study area with $VCD_{strat} = 1.1 \times 10^{13}$ molec cm^{-2} and 1.3×10^{15} molec cm^{-2} for BrO and NO_2 respectively (see Fig. S2 for profile shape). Profiles were created by averaging and smoothing 30 stratospheric profiles, 5 each from RF01, RF04, RF05, RF12, RF14 and RF17, chosen from flight periods with a consistent tropopause between 17 and 18 km.

p9, line 22: Profile 1 is similar to extinction we found over pristine ocean during the TORERO project. Profiles 2 and 3 are constructed specifically for the sensitivity studies here to investigate the effects of higher AOD (2) and lofted pollution (3).

Page 10, line 12. Please provide information about source of atmospheric profiles.

Respective information has been included.

p.10, line 17: Trace gas, temperature and pressure profiles were created by averaging and smoothing 30 individual profiles, 5 each from RF01, RF04, RF05, RF12, RF14 and RF17. Stratospheric ozone and tropospheric NO_2 profiles are from RAQMS. For ozone flight periods with a consistent tropopause between 17 and 18 km are chosen, and for NO_2 areas with pristine background air. Tropospheric ozone, H_2O, temperature and pressure are from aircraft in situ measurements and averaged over the same periods as model profiles.

Page 11, line 9. Figure 1 shows dBOX-AMF for EA 90^o zenith reference however for O4 and IO EA 10^o zenith reference is used. It will be interesting to see a similar plot to 1 for EA 10^o. Similar to point to the general commnet.

Good point. This is addressed in our answer to the referee's general comment and in the new Figure S1.

Page 11, line 26. Please include reference for typical fit uncertainties for the University of Colorado (CU) AMAX-DOAS instrument.

Respective information has been included.

p.11, line 26: These limits are based on typical fit uncertainties for the University of Colorado (CU) AMAX-DOAS instrument as reported among other trace gases in Volkamer et al. (2015), i.e., 1.3×10^{13} molec cm^{-2} for BrO, 2.1×10^{12} molec cm^{-2} for IO, and 1.5×10^{14} molec cm^{-2} for NO_2.

Page 25, line 21. Why parameterization retrieval is using pressure, temperature and water vapor data averaged over each full flight instead of each profile as the OE?

Good point. The following information has been included.

p.25, line 21: The parameterized retrieval uses an almost identical RTM atmosphere, but pressure, temperature and water vapor data are averaged over each full flight, which saves RTM computation time. This averaging is possible for the TORERO study area, because air masses for each flight are rather homogenous.

Page 26, lines 7. How well compares the cloud information obtained by the ratio of modeled and measured O4 dSCDs and HSRL observations?

Good point. We have included this comparison in the text and have added an SI Figure.

p.26, line 15: Overall, the cloud information gained from the O_4 ratios is very consistent with HSRL observations. Figure S12 provides examples of O_4 ratios and HSRL data for different aerosol and cloud scenarios.

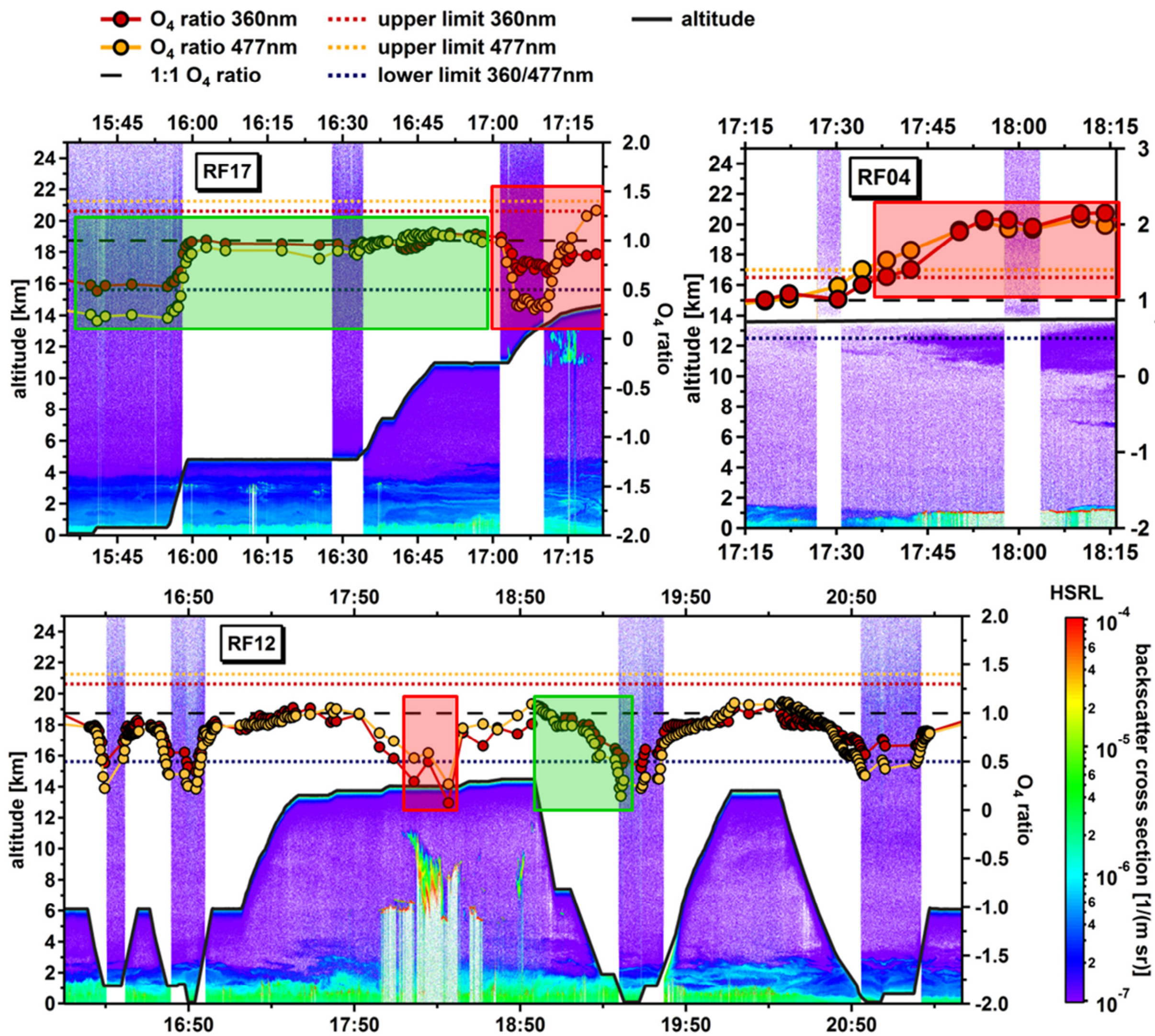

Figure S12. Comparison of O_4 ratios at 360 nm and 477 nm with HSRL particulate backscatter cross section data for RF04, RF12 and RF17. Altitude resolved HSRL backscatter data is plotted and color coded along the flight track. Larger signals denote the presence of aerosol/clouds. HSRL is either measuring above or below the aircraft. The shading directly around the flight track seen in part of RF12 and RF17 is a near field effect that leads to erroneous large back scatter signals by HSRL. DOAS O_4 ratios along the flight track are plotted on the right axis. Upper (red and orange dashed lines) and lower (blue dashed line) O_4 ratio limits denote where aerosol/cloud conditions are considered too complex and respective trace gas dSCD data is not used for parameterization (Section 5.2). Note that the lower limit is only relevant when the aircraft is flying above cloud layers and does not apply to cloud free boundary layer legs. The one to one line is added as reference and signifies Rayleigh conditions. Red boxes show cases where dSCD data was filtered based on cloud conditions. Green boxes in RF12 and RF17 mark data periods that were used for BrO, IO and NO_2 OE profile retrievals as published in Volkamer et al. (2015).

Regular HSRL upward scans show that for these time periods no aerosol or cloud layers were present above the aircraft.

The red box in RF17 displays an example where data is filtered because the aircraft is within 2 km of an elevated cloud layer. The cloud shields O_4 concentrations below the cloud, which leads to very low measured O_4 dSCDs and thus a very low O_4 ratio. For the time period between 17:01 and 17:07 UTC, where HSRL scans upward, filtering is aided by aircraft video data. A similar effect is observed during RF12, marked by a red box. Here data is filtered based on exceeding the lower O_4 ratio limit. The red box in RF04 shows an example where the aircraft flies across a rather solid low cloud layer situated at ~1.2 km (red HSRL backscatter data points), while almost simultaneously an optically thin aerosol layer right below the aircraft is encountered. The elevated aerosol layer is not sufficiently optically thick to shield O_4 below. Instead, the increased albedo caused by both the boundary layer clouds and the lofted aerosol layer leads to measured O_4 dSCDs that are up to a factor of two higher than those simulated for a Rayleigh case. Here, data points are filtered by the upper O_4 ratio limits.

Section 5.3. It will be nice to explain the reasons why some flights are excluded from OE vs. parameterization comparison. Are these flights linked to specific trace gases, aerosols or clouds conditions?

The choice of flights that are included in the OE vs parameterization comparison is based on existing high quality OE profiles for identical time periods and all three trace gases. Since OE retrievals are less time consuming in cloud free conditions, all of the profile comparisons are in predominantly cloud free atmospheres. We have added an explanation in the manuscript. Note that we added NO_2 profile comparisons (see also answer to referee #2's comment "Figure 8 and Section 5.3").

p.28, line 2: To increase statistics for the correlation, trace gas profiles from more flights are included, i.e. from RF01, RF04, RF05, RF12, RF14 and RF17 for all three trace gases (see also Wang et al. (2015) and Volkamer et al. (2015)). Profile selection is based on availability of high quality OE profiles.

Conclusions, page 29 line 24. Retrieval duty cycle is mentioned in the conclusions but never explained anywhere else.

This is a good point. We have modified the conclusions and added a paragraph above that introduces the duty cycle.

p.28, line 26: Based on the number of available EA 0° spectra, a parameterization duty cycle is defined, which expresses the fraction of EA 0° dSCD measurements that are converted into VMRs. A total of 7124 BrO dSCD measurements and 7168 IO dSCD measurements were processed. 24.3 %, 22.4 %, and 10.3 % of the BrO data points, and 10.6 %, 26.2 %, and 3.2 % of the IO data points were filtered due to data quality, cloud, and parameterization method filters, respectively. The resulting duty cycle for the quality assured VRM retrievals were 55.4 % for BrO and 81.4 % for IO before cloud filters, and 44.3 % and 60.1 % for BrO and IO after cloud filters. Notably, the method filter removes the least number of data points for both trace gases, which underlines the statistical advantage of the parametrization method over OE. Based on these numbers, the development of an advanced cloud treatment has good potential to further improve the duty cycle.

p.29, line 24: The TORERO VMR retrieval has a duty cycle of 55.4 % and 81.4 % for BrO and IO before cloud filtering, and 44.3 % and 60.1 % after. Less than 11 % of data points are removed by the method filters, which underlines the statistical advantage of the parametrization method over OE.

Conclusions, page 30 line 1. What are the numbers for the removal of NO2 data? How can the data removed due to cloud filtering be different for different species (BrO and IO)?

We did not run the parameterization retrieval for NO_2 on the complete TORERO data set and have added the following explanation in the revised manuscript.

p.26, line 25: For NO_2, a typical TORERO $dSCD_{strat}$ value at 60° SZA is 1.5-3 x 10^{15} molec/cm^2, which is about a factor of 2-4 higher than a tropospheric EA 0° dSCDs measured in pristine background air (e.g. NO_2 c-profile). Here, our EA 10° dSCD fit error of 3-5 x 10^{14} molec/cm^2 is almost on the same order as the tropospheric EA 0° dSCDs, underlining the need for a highly accurate characterization of the stratospheric VCD by EA 10° measurements. Since TORERO specifically targeted very pristine air masses, we refrained from running the NO_2 parametrization retrieval on the complete data set, and focused instead on select case studies with $dSCD_{strat}$ << EA 0° dSCDs. For these cases evaluation of NO_2 is still possible down to as low as 10 pptv, as has been demonstrated in Fig.10 in Volkamer et al. (2015).

There are two reasons for a difference in the cloud filter percentage for BrO and IO. 1) The different overlap of DOAS analysis quality filtered data points with cloud flagged points entering the VMR retrieval. 2) In addition to the initial color ratio and video cloud screening, points get cloud flagged based on the ratio of measured and modeled O_4 (see p.26, line 11ff), which is wavelength dependent and differs for BrO and IO.