Reply to Referee #2

First of all we want to thank this reviewer for the positive assessment and constructive comments.
We addressed these comments as explained in detail below.

General comments
In recent years, more and more authors have indicated the need for scaling factors in order to improve the agreement of measured and simulated O4 dSCD/dAMF. In the previous publication by Wagner et al. 2019, various factors were investigated to determine the possible cause of this disagreement. One of the key remarks made by reviewers and the community was that the uncertainty of aerosol information and ist impact on the oxygen dimer could not be ruled out as a possible cause of disagreement. In this novel study, Wagner et al. examine the difference of measured and simulated O4 dAMF for low aerosol loads measured during a ship cruise in the Atlantic in 2019. The authors claim that due to the low aerosol load possible aerosol uncertainties can be neglected and that the underlying differences must have another, as yet unknown reason.
The document is well written and structured and the analyses have been carried out thoroughly and consistently. However, I recommend publishing it after making some minor changes listed below.

We thank the reviewer for the positive assessment.

1. Please add a table including all uncertainties described in the document (e.g. pressure/temperature changes, aerosol parameterization, effective temperature, ...)

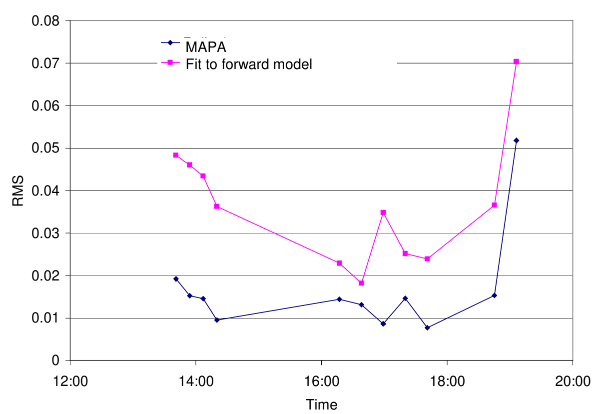
Many thanks for this good suggestion! We added such a table (new table 3) to the paper:

Table 3 Uncertainties related to the different analysis steps

Spectral analysis		
Effect	**Magnitude**	**Reference**
Spectral fit	1 - 4%	Result of spectral fit
Temperature dependence	1.5%	Wagner et al., 2019
Fit paramaters	3.5%	Appendix A1, and Wagner et al., 2019
Total	4 – 5.5%	
RTM without aerosols		
O_4 profile	1%	Wagner et al., 2019
albedo	1%	Section 6
RTM general	1%	Wagner et al., 2019
total	2%	
RTM with aerosols		
O_4 profile	1%	Wagner et al., 2019
AP & SSA	3%	Section 6
Strat aerosols	1%	Section 6.1
albedo	1%	Section 6
Profile shape	2% for elevation angles < 4°, negligible for higher elevation angles	Section 6.1
RTM general	1%	Wagner et al., 2019
total	4%	
O_4 VCD	2%	This study, section 5, see also Wagner et al., 2019

2. It would be interesting to have a time series of O4 dSCD/dAMF RMS values (similar to A13) for the data shown in Fig.6 and A11. I would expect a clear trend in the RMS differences over the day maybe similar to the one you showed for AOD and scaling factor? How is the correlation of these RMS values and the retrieved/measured AOD?

We prepared the requested figure:



The RMS values from the fit to the forward model show the same temporal trend as the RMS from MAPA, but the absolute values are smaller (as expected).

We also checked the correlation: No correlation was found (R^2=0.00)

3. You mentioned that sun photometer measurements allow to differ between the aerosol particle size. Please show the contribution of differently sized aerosol particles to the total AOD over the day as well as all AODs and corresponding Angström exponents.

The following figure was added (new Fig. A3):

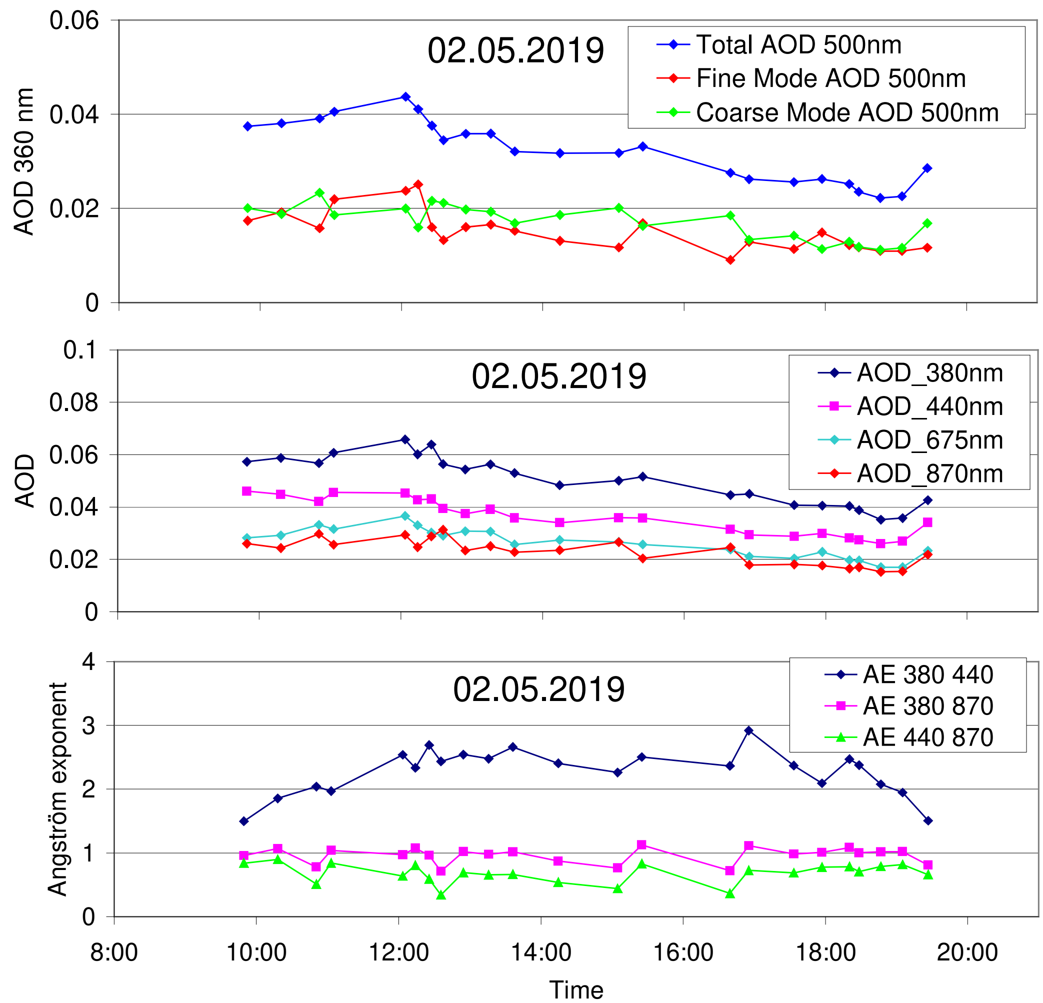


Fig. A3: Top: AOD at 500 nm attributed to the coarse and fine modes, as well as total AOD. Middle: Total AOD at different wavelengths. Bottom: Angström Exponents for selected wavelength pairs.

It should be noted that while creating these figures, it turned out that the AOD for the last measurement on 2 May 2019 (at 19:26) of this fully processed data set was about 30% higher than the value obtained from the initial AOD inversion (while all other measurements on that day were nearly identical). Even after consultation with the AERONET staff, no clear reason for this discrepancy could be identified. However, since the solar zenith angle is rather large (~84°), the extraction of the AOD, especially at short wavelengths is challenging, because of the strong Rayleigh extinction. These uncertainties affect the comparison of the last elevation sequence (19:06 to 19:25).

We added the following information to the paper:

Section 2.2:

‚It should be noted that the uncertainties of the last AOD measurement on 2 May 2019, 19:26, are rather large because of the high SZA of 85°. In particular it was found that for that measurement the AOD from the fully processed sun photometer data (Fig. A3) was about 30% larger than the AOD of the initial retrieval (Fig. 3), while the results for all other measurements are almost identical. The radiative transfer simulations presented below for the last elevation sequence (19:06 to 19:25) are based on the initial (low) AOD values, which are in agreement with AOD measurements 20 minutes earlier. Nevertheless, the comparison results for this last elevation sequence should be treated with caution because of the large uncertainties of the corresponding AOD measurement.'

Figure caption of Fig. 7:

‚Note that for the last elevation sequence, the AOD used in the forward model has large uncertainties, see section 2.2.'

Figure caption of Fig. A12:

‚Note that for the last elevation sequence, the AOD used in the forward model has large uncertainties, see section 2.2.'

4. Furthermore, I was wondering why you only showed results for one day? The AODs for the following days are also rather small. Do these days support your findings?

The extremely low AODs only occurred on the selected day. Only at the beginning of the following day, still low AODs were measured (< 0.05 at 360 nm). However, during this period, the measurements at low elevation angles were strongly affected by clouds. Nevertheless, we compared the MAX-DOAS O_4 measurements retrieved during that period with radiative transfer simulations. Here, we only made simulations for an aerosol-free atmosphere to limit the effort (and also because of the rapid temporal variation of the AOD). The comparison results are shown below:

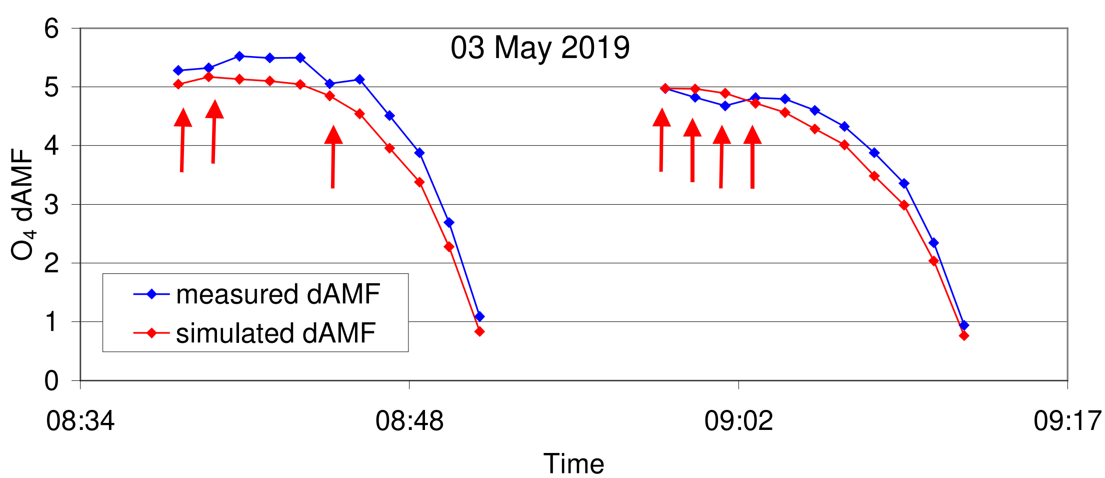


Fig. A13 Comparison of the measured and simulated O_4 dAMFs for two elevation sequences on 05 March 2019, when the AOD was rather small (<0.05 at 360 nm). The radiative transfer simulations were made for an aerosol-free atmosphere.

Like on 02 May 2019, the simulated O_4 dAMFs (for aerosol-free atmosphere) are smaller than the measurements (for cloud-free observations).

We added the following information to section 7.1.:
‚It should be noted that that during the entire ship cruise, only during the beginning of 3 May 2019, similarly low (but still larger) AOD were measured as on 2 May 2019. We also compared the measured O4 dAMFs for the first two elevation sequences on 3 May to radiative transfer simulations. For that comparison we only made simulations for an aerosol-free atmosphere in order to limit the effort (and also because of the rapid temporal variation of the AOD during that time period). The results (see Fig. A13) are similar to those on 2 May 2019: except for the cloud contaminated measurements, the simulations are smaller than the measurements.'

Specific comments

P1, L22, 25: Please add a selection of corresponding references to the sentences starting with (L22) "In recent years,..." and (L25) "Several studies found that a scaling factor...".

We changed the text to:
‚Several studies found that a scaling factor (SF<1) had to be applied to the observed atmospheric O_4 absorptions in order to bring them into agreement with radiative transfer simulations (e.g. Wagner et al., 2009; Clémer et al. 2010). Other studies, however, did not find the need to apply such a scaling factor (e.g. Spinei et al., 2015; Ortega et al., 2016). A more detailed discussion and overview on existing studies of both groups is provided in Wagner et al., 2019.'
We added the following references:

Spinei, E., Cede, A., Herman, J., Mount, G. H., Eloranta, E., Morley, B., Baidar, S., Dix, B., Ortega, I., Koenig, T., and Volkamer, R.: Ground-based direct-sun DOAS and airborne MAX-DOAS measurements of the collision-induced oxygen complex, O2O2, absorption with significant pressure and temperature differences, Atmos. Meas. Tech., 8, 793-809, https://doi.org/10.5194/amt-8-793-2015, 2015.

Wagner, T., Apituley, A., Beirle, S., Dörner, S., Friess, U., Remmers, J., and Shaiganfar, R.: Cloud detection and classification based on MAX-DOAS observations, Atmos. Meas. Tech., 7, 1289-1320, doi:10.5194/amt-7-1289-2014, 2014.

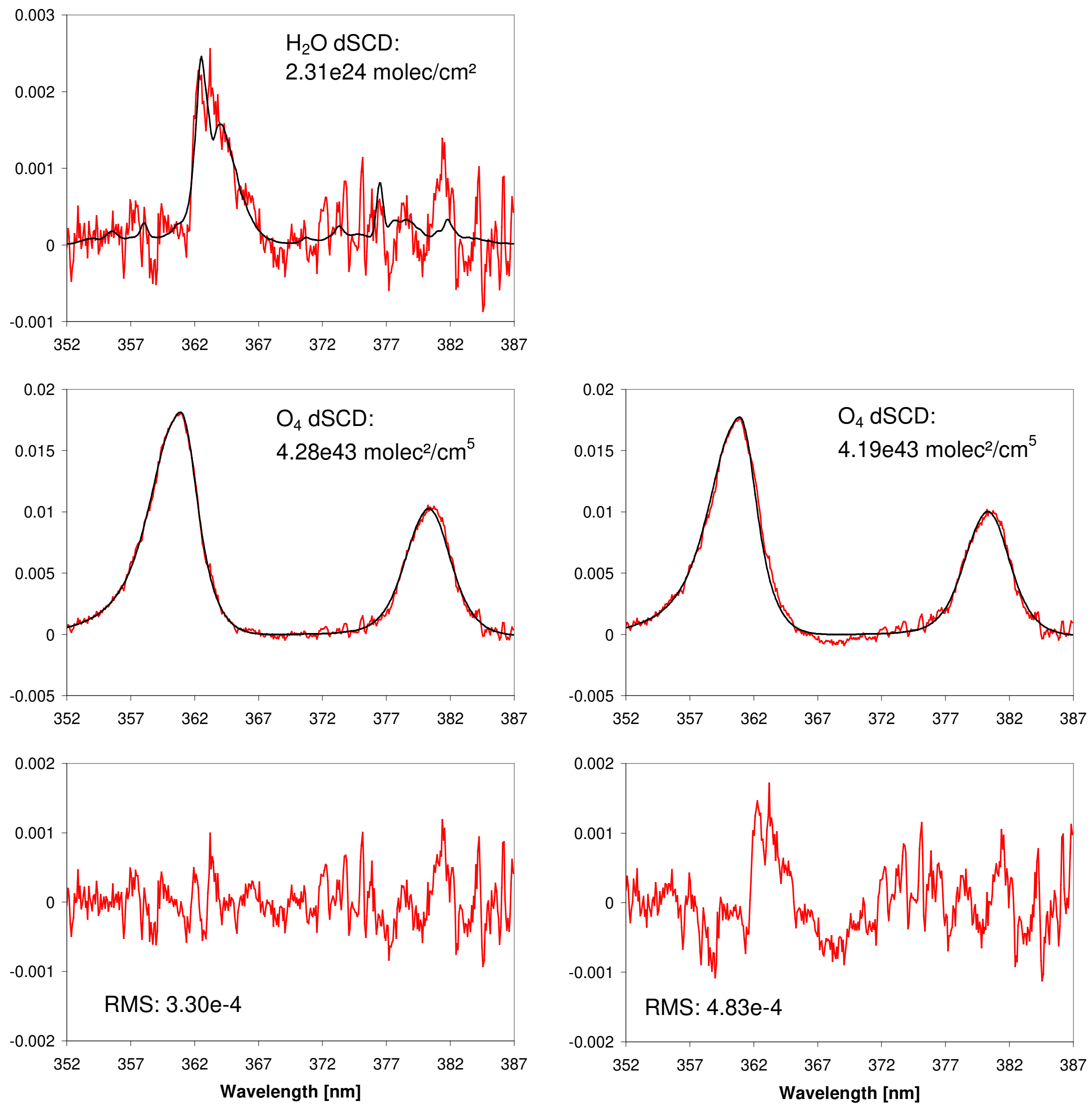
P2, Sec 2.2 and Fig.3: Since AODs at other wavelengths are available, please add them to Fig. 3.

We added a new figure (new Fig. A3) showing the AODs at all wavelengths, see reply to comment 3 above.

P3, L87: "(with..." => "(which...?
Corrected

Fig. A1: Your fit uses the wavelength range 352 - 387nm but Fig. A1 shows only wavelengths up to ~384nm. Please change the x-Axis according to the applied fitting window.

Many thanks for this hint!
In this study, we restricted the spectral range to 352 - 385 nm, because for some measurements (not on 2 May 2019) large spectral structures were found > 385 nm). However, for 2 May 2019, almost identical results (differences < 1%) were found for both spectral ranges.
We added this information to section 3.
The figure below shows the fit results for the spectral range 352 – 387 nm for the same spectrum as shown in Fig. A1. The results are almost identical to those shown in Fig. A1.



Fit results for a spectrum taken on 2 May 2019, 13:14:50, at an elevation angle of 1° (SZA: 33.6°). Left: results if a H_2O cross section is included in the spectral analysis; right: results if no H_2O cross section is included in the spectral analysis. The black lines represent the fitted cross section, the red lines indicate the residual (bottom) or the residual plus the fitted cross section.

Furthermore, I was wondering about the shown residual. It appears to me that there are still some residual structures left. Especially three peaks around 372-376 nm look familiar and could be attributed to Fraunhofer-Lines. Could this be somehow related to your Ring-treatment or do you have another explanation?

We have no plausible explanation for these structures. For the O4 analysis, these small remaining structures are not critical.

P5, L177: Why was the albedo set to 0.05? Please add a references here. As far as I know, we can expect a small dependence on SZA. How large is the impact on O4 when changing the albedo according to possible values?

We chose the value of 5% to be consistent with the MAPA inversions, and because it is appropriate for many parts of the global ocean. However, by having a closer look at maps of albedo (Kleipool et al., 2008) and chlorophyll content (e.g. from the NASA Earth Observatory: https://earthobservatory.nasa.gov/global-maps/MY1DMM_CHLORA), we found that at the specific location of the measurements, very clear waters exist, for which the surface albedo is typically higher (about 7 to 8%). The presence of very clear waters is also supported by the in situ chlorophyll measurements made aboard the ship.
We therefore made additional radiative transfer simulations using a surface albedo of 8%. We found that the obtained O_4 dAMFs were almost identical with those obtained for 5% surface albedo (differences <1%). The reason for the good agreement is that the effect of the surface albedo is similar for the O_4 AMFs for different elevation angles. Thus the effect of varying surface albedo almost cancels out.
We added this information to section 6.

P6, L215: "the the", please remove either first or second "the".

We removed both the first and second ‚the' and added a new one.

P6, L216: You wrote that Fig. A8 includes constant and linearly extrapolated values for lower altitudes but the greenish line does not look like a linear extrapolation to me. Why is that?

Many thanks for this hint. We corrected the figure.

P6, L221: Why is the Angström exponent "assumed" to be 2 when you have AODs at several wavelengths available to calculate more accurate values?

The AOD measurements represent the total AOD, but what is needed is the Angström exponent for the stratospheric aerosols. Therefore, the sun photometer measurements cannot be used for that purpose.

P6, L224: decsribed => described
Corrected

P7, L263: "smaller than" => "larger than"
Corrected

P7, L276: "bebetween" => "be between"
Corrected

Fig. A7 Could you please add a similar figure for the geometry with the smallest RAA to better assess the impact of AP and SSA variations throughout the day?

The figure for 13:30 (RAA ~0) is added.