The replies to the reviewer comments are marked in blue

General comments

This manuscript discusses the statistical significance of the gap between observed and simulated AMFs of O4 on selected two clear-sky days during MADCAT campaign. Thorough and detailed analysis of various factors producing uncertainties in the observed and simulated AMFs was made. The authors pointed out the importance of proper usage of temperature and pressure for the condition, proper account of aerosol optical parameters (phase function, aerosol profile extraction) in the simulation, and standardization of DOAS settings (spectral range, degree of polynomial etc) for observations. Considering these factors altogether, the authors conclude that the gap was insignificant on one day (June 18) but was significant on other day (July 8), supporting conclusion from some previous works. Recognizing that there is a hot debate in the community if the scaling factor is necessary, the manuscript is valuable since it provides as thorough analyses as ever provided.

We thank the reviewer for the positive assessment of our paper and for the good suggestions. We addressed them as described in detail below.

Nonetheless, I would like to request revision on the following points. First, I find the studied uncertainties could be classified into two types: those from apparently ill treatment (i.e., 203K O4 cross section, US standard atmosphere without temperature correction, no offset in the DOAS analysis etc) and those unavoidable even with the stateofthe-art analysis. For the purpose of evaluating spread of results from multiple groups and of determining best practice to avoid potential hazard during the analysis, determination of the former type uncertainty helps. But when discussing the significance of the gap between observed and simulated AMFs of O4 critically, only latter type uncertainties should be used. In such a way better control of the determined uncertainties is recommended.

We agree that such a separation of different types of uncertainties would be helpful. Therefore we added two columns to tables 9 and 10 in which we quantify the uncertainties if optimum settings were used and sufficient independent information was available. For the radiative transfer simulations of the O4 dAMFs the uncertaintes for these optimum settings are about ±4% compared to ±(6 – 9)% for two days of the MAD-CAT campaign. For the spectral analysis the uncertaintes for the optimom settings are about ±6% compared to ±(11-13)% for the two selected days of the MAD-CAT campaign.
These findings indicate that for future campaigns the comparison of measured and simulated O_4 absorptions can ptobably be carried out with much better accuracy (if these optimum settings were used). Here it should, however, be noted that the optimum settings for the radiative ransfer simulations will require LIDAR measurements at the same wavelengths as the MAX-DOAS measurements and without a sensitivity gap close to the surface. Such measurements are currently hardly available. This information was added to the new section 4.4.

Secondly, it should be more clarified in Abstract that the precise determination of the uncertainties (+/- 0.16 and +/-0.12 here) is the main point. Careless readers may not realize the importance.

We agree and modified the abstract to make this point more clear. We also changed the title to: 'Is a scaling factor required to obtain closure between measured and modelled atmospheric O4 absorptions? An assessment of uncertainties of measurements and radiative transfer simulations for two days during the MAD-CAT campaign'.

Thirdly, possible influence of horizontal heterogeneity of aerosol optical parameters should be mentioned. When the aerosol abundance over the line of sight is becoming less with distance (which may be likely when instrument is located in a city looking out of it), the observed higher O4 dAMFs might be better explained by considering such inhomogeneity even on July 8. I understand that with 1-D radiative transfer models homogeneity needs to be assumed and detailed discussion would be beyond the scope. However, some simple analysis such as that on spatial distribution of AOD from satellite with a fine resolution maybe possible.

We agree that this is a potentially important aspect. However, for the two selected periods the wind direction and wind speed were rather constant. On 18 June the wind direction was between 80° and 150° wrt North, and the wind speed was about 2 m/s. On 8 July the wind direction was between 70° and 90° wrt North, and the wind speed was about 3 m/s. Thus on 8 July the wind came from almost the same direction at which the instruments were looking. Taking the wind data into account, during the 4 hours of the selected period on 8 July, the air masses moved along a distance of about 40 km. During the 3 hours of the selected period on 18 June, the air masses moved along a distance of about 20 km. These distances are larger than the distances for which the MAX-DOAS observations are sensitive. Since also the AOD and the aerosol extinction profiles were rather constant during both selected periods, we conclude that for the measurements considered here horizontal gradients can not explain the discrepancies between measurements and observations. It should also be noted that the discrepancies were simultaneously observed at all 4 azimuth directions. We added this information to section 4.2.1.

Lastly, conciseness should be attained during revision. I would suggest shortening section 4.1 and section 5 (paragraphs before section 5.1).

We moved several parts of section 4.1 to the appendix. We also shortened the paragraphs before section 5.1.

Overall, I would suggest minor revisions on the general comments above and some specific comments listed below.

Specific comments
1. Line 359. Probably appendix A2?

Corrected

2. Line 526. US standard atmosphere

Corrected

3. Figure 10. What are the differences of the first three series, with same legend "HG AP 0.6?"

The correct labels (0.60, 0.68, and 0.75) were added.

4. Figure 11. Although the panel is for showing noise influence, the gap related to the main conclusion of this study is well represented as the difference in the O4 optical depths in the first two panels. Such discussion should be added in section 4.3.1.

We added the following sentence to section 4.3.1:
'Here it is interesting to note that the ratios of the results for the measured spectrum and the simulated spectra are between 0.68 and 0.74, similar to ratio for the dAMFs on 8 July shown in Table 8.'

5. Table A12 in line 1922 is mislabeled. (Table A10)

Corrected

6. Table A11. MCARTIM

Corrected

7. Lines 846-848. Second and third points should be exchanged, considering the order of Fig. 14b and c and the following discussion.

The order was changed

8. Line 906. Overall uncertainty calculation deriving 0.12 is not clear. When considering 3% uncertainties for VCD, 6.1% from radiative transfer simulation, and 10.8% from spectral analysis, the overall uncertainty may be 13%. When it is around 0.71, it can be 0.09?

Many thanks for this hint! We agree and updated the calculations accordingly (with slightly modified uncertainties, see tables 9 and 10.

9. Line 944. 8 July

Corrected