Dear Reviewer and Editor,

Thank you very much for your comments. Please find below our replies.

Major comments

My comment "Supplement 3 is meant to demonstrate a good agreement of the small-scale ozone variability in OMI and SILAM data. Looking at Figs. 8-10 I cannot follow how the authors come to the conclusion that the agreement between the modeled and experimental data is very good, e.g. I see nothing in common between black or between red curves for 60°S – 90°S in Figs. 9 and 10. As this part is not highly relevant for the rest of the study this supplement can be removed. Otherwise comparisons and justification of the conclusions must be improved." was addressed only in part. The strong disagreement at high southern latitudes is properly explained and the remaining agreement is claimed to be good. However, one still sees a lot of discrepancies, e.g. in Fig. 9 and 10 for 30°S – 30°N region the behavior of most of the curves is quite different for distances larger than 500 km, for 30°N – 60°N the blue, green, black and red solid curves are grouped clearly different, for 30°S – 60°S red and black dashed curves do not look similar. For this reason I do not agree with rating the agreement as good without any additional comments.

We wrote a more detailed discussion on comparison of structure function from model and OMI data. In the revised version, instead of using a subjective word "good", we explained in detail agreement and disagreement of experimental and modelled structure functions for total and stratospheric ozone columns. The text in the second revision is:

"The overall morphology - latitudinal dependence, latitude-longitude anisotropy, seasonal cycle - is similar for OMI and SILAM, for both total and stratospheric ozone column. . For total ozone column, the experimental and modelled structure functions are very similar for almost all latitudinal zones. Some disagreement in seasonal cycle is observed for polar winter conditions (for example 60-90S in June-Aug, Figs S7 and S8). This disagreement is quite expected: OMI cannot measure in polar night conditions. The shape of structure functions and the growth with separation distance are similar in Figs S7 and S8, but some difference in absolute values exists and is expected; it comes from biases between model and observations (note that the structure functions are presented in absolute values).

For stratospheric ozone column, the comparison is more complicated, because we could use only cloudy pixels of OMI, which have limited coverage. This results in less reliable estimates of structure function from the OMI data. For example, limited amount of data at large separations (> 500-1000 km), resulted in different shapes of experimental and model-based the structure functions in the equatorial zone. Although the seasonal cycle and latitude-longitude anisotropy are qualitatively similar in Figures S9 and S10, mid- and high-latitude structure functions tend to

group somewhat differently. Therefore, comparison of the stratospheric ozone column structure functions in Figs S9 and S10 should be considered as indicative only.

Since the stratospheric ozone has a bulk contribution to the total ozone (for which observational and modelled structure functions are similar), and with the above notes, we conclude that the ozone small-scale variability is realistically represented by SILAM."

Minor comments

"Page 2, lines 50-52: this is only true for the along line of sight direction. The resolution can be much higher in the across direction, e.g. ALTIUS, CAIRT."

"We added 'along line of sight' in the revised version."

This must be a misunderstanding. I meant that your sentence is only true for the along track horizontal resolution while the across track horizontal resolution can be much higher (forward/backward view is assumed). With that the correction in the revised version of the manuscript is not appropriate. Suggestion: "The measurements in the limb-viewing geometry have usually a good vertical resolution but their horizontal resolution is limited by the spatial sampling. In particular, the horizontal resolution in the along line of sight direction is limited by the effective horizontal length of interaction with the atmosphere (a few hundreds of kilometers)."

Corrected as suggested.

Page 3, line 71: The data calibration is not a serious issue then combining total/stratospheric ozone columns retrieved with DOAS-like methods

"We agree the calibration is not a serious issue, but still an issue (Fishman and Larsen, 1987)."

As far as I know both TOMS and SAGE algorithms discussed by (Fishman and Larsen, 1987) are not DOAS-like methods. Suggestion: " Aside with calibration issues when using a combination of TOMS and SAGE instruments, there was also..."

Corrected as suggested.

"Page 10, third item: this conclusion depends certainly on the sampling of the considered instruments and should not be stated in general. By the way, are the authors aware of any more or less recent publication where the residual method was applied to the monthly mean values? Isn't the recommendation not to combine the monthly mean values too obvious for the scientific community for now? Another point to this topic, as shown in Fig. S4 of the paper, there is quite a strong difference between the tropospheric ozone values calculated from daily means and from the collocated data. Thus, the recommendation given by authors to use the daily measurements can be confusing for the readers forcing them to prefer daily means to the fully collocated measurements."

"Although we are now aware about recent publications on the residual method applied to monthly mean values, we think it is worth to keep this statement. We agree that it is rather obvious, and added 'obviously' to this sentence. It is noted in the paper that SUNLIT tropospheric ozone column correspond to the local time of OMI and TROPOMI measurements, not daily mean. We stress this more in the revised version."

I do not mind to keep the recommendation not to use monthly mean data to calculate the tropospheric columns. However, I am still against a recommendation to use daily values. The usage of collocated measurements should still be preferred if feasible.

This seems to be a misunderstanding. We do not state that the daily values are preferable, compared to fully collocated data. To avoid the potential confusion here in the text, we changed the statement:

"Due to large variability of ozone field and limited sampling by satellite instruments, nadir and limb measurements should be collocated in time and space, if feasible. "

Page 11, lines 274: "region between the two ozone jumps is removed" - from the text above it is unclear which two ozone jumps are meant.

We changes "ozone jumps" to "pixels with huge ozone gradient".

It is still unclear what the authors mean. Are the pixels with values over 100 DU and their neighboring pixels are removed? Does it apply to one or more neighboring pixel at each side?

We rephrased: "The presence of row anomaly was also checked by evaluating the ozone difference in neighbouring rows. Along the swath direction, the anomaly is visible as a sudden drop and rise of the retrieved ozone column. The procedure was checking a difference in neighbouring pixels; if larger than 100 DU drop and rise are detected, all pixels between these two points were removed."