Performance of AIRS ozone retrieval over the central Himalayas: Case studies of biomass burning, downward ozone transport and radiative forcing using long-term observations
By Prajjwal Rawat et al., 2022 (AMTD)

We are grateful to both the referees for their useful comments and constructive suggestions, which have improved the MS significantly. The manuscript is suitably revised by incorporating their suggestions and comments. We are also thankful to the editors for their time. We feel that the revised manuscript is suitable for publication in AMT. Please find here our responses in boldface and the referee's comments are in regular font.

Refree#2

We thank you for your constructive comments and suggestions. The point-by-point responses to the comments are given below in boldface font.

The authors have access to some 250 ozonesonde profiles from the central Himalayas. They were launched from a high altitude location just north of many heavily populated cities in the Indo-Gangetic Plain. Their objective is to use this valuable and unique dataset to evaluate the quality of ozone data from several satellite sensors, particularly the AIRS sensor on NASA's Aqua satellite. Though their objective is commendable, the paper suffers from several problems that include flaws in the analysis methodology, poor quality of the figures and captions, and lack of careful editing.

The authors rely very heavily on the use of the so-called "smoothing" formula (Equation 1) proposed by Rodgers and Connor (2003) published in JGR (vol 108, D3). Unfortunately, this formula is often misused. Equation 1 actually creates a hybrid of a high res profile and a priori (AP) profile. Its purpose is to asses if a remote sensing instrument has been properly calibrated and its retrieval algorithm has been correctly implemented. In such cases the retrieved and the hybrid profiles should agree. However, the formula does not provide a method of assessing the science value of the profiles independently provided by the low vertical resolution sensor. To asses it one needs to apply more traditional smoothing methods, such as Gaussian smoothing or computation of layer columns.

To understand the difference let us consider two simple examples. Let us say that a satellite sensor provides no information in a given atmospheric layer. In such cases the AK of the satellite sensor in that layer will be zero and eqn. 1 will yield the a priori (AP) value in that layer irrespective of what the ozonesonde measures. This is not what one means by "smoothing". A more relevant case is when a satellite sensor contains just the total ozone information with no useful profile information. In such cases it can be shown that eqn 1 will transform two high res profiles with very different shapes but containing the same total ozone amount to exactly the same profile that will look like the AP profile but scaled to provide the correct total ozone. Again, this is not what one means by “smoothing”. In such cases it is best to compare total ozone values from different sensors directly.

Given this background I find only Fig 10 of the paper useful. Unfortunately, the figure is marred by several flaws. Firstly, computation of layer amounts by itself amounts to smoothing, so equation 1 should not be applied to the ozonesonde profiles. Secondly, the figure seems to show ozone variability as error bars. It is far better to plot the standard error of the mean, which is the proper method of assigning errors bars to mean values. These two changes will make the figure less cluttered and easier to evaluate.

Unfortunately, my assessment of the results presented is that the correct smoothing of the ozonesonde profiles by applying a Gaussian filter or by comparing the layer amounts (without applying eqn 1) would not confirm the key conclusion of this paper that AIRS does well in the troposphere and the stratosphere but not in the UTLS. Still, given the uniqueness of the location, the results are worth publishing.

Thank you very much for your elaborate comments and suggestions. In general, when comparing the measurements of two different sensors, there is no perfect way to minimize the effect of different horizontal and vertical resolutions. However, to minimize the biases due to different vertical resolutions, high-resolution profiles are generally smoothed. The AIRS IR ozone retrieval utilizes the optimal estimation-based algorithm, which have limited vertical resolution, depending upon the spectral resolution of instruments or simply the weighing function.

In the comparison analysis, to account for the different vertical resolutions and to perform a meaningful comparison of two independent instruments (e.g., ozonesonde and satellite), various groups have utilized the satellite averaging kernels and a-priori information (Boynard et al., 2009; Zhang et al., 2010; Verstraeten et al., 2013; Bak et al., 2019; Zhao et al., 2020) to convolved the ozonesonde or any other high-resolution instruments for smoothing their ozone profile according to Eq. 1 of MS. For example, (1) Boynard et al. (2009) utilized the averaging kernel matrix of the IASI retrievals to smooth the ozonesonde profile before their comparisons to minimize error arising from different vertical resolutions, (2) Zhang et al., 2010, used OMI and TES AKs smoothing to compare their ozone retrieval of tropospheric ozone with ozonesonde, (3) Verstraeten et al. (2013) utilize the TES AKs to compare TES retrieved ozone profile with ozonesondes (4) Bak et al., 2019 used GEMS AKs to compare GEMS simulated tropospheric ozone profile with ozonesonde, (5) Zhao et al., 2020, utilize the TROPOMI AKs to compare TROPOMI retrieved ozone profile with ozonesondes.

However, in some cases, very small improvements in biases are seen after applying the averaging kernels smoothing, as in the case of MLS (Adams et al., 2014), which is due to the delta functions nature of MLS averaging kernels. In such cases, the smoothing is acquired using various other techniques like the Gaussian or triangular smoothing with a full width at half maximum (FWHM) of the respective distribution equal to the typical vertical resolution of a low vertical resolution satellite instrument (Wang et al., 2020). Wang et al. (2020) for assessing SAGE III/ISS ozone retrieval with collocated satellite instruments (MLS, OMPS LP), ACE-FTS, and ozonesonde utilize the Gaussian smoothing to high-resolution profiles. While Nalli et al. (2017) utilized the broad-layer averages to compare CrIS ozone retrieval with ozonesonde.

Furthermore, we would like to mention that this is the first attempt in which ozonesondes launched over the central Himalayan site are utilized to evaluate the performance of AIRS, IASI, and CrIS ozone, particularly AIRS ozone retrieval. The AIRS averaging kernel is successfully calculated in all the 100 RTA layers using the trapezoid function and utilized for the first time in the evaluation study. There was very limited or no discussion on the AIRS ozone averaging kennels, which is a fundamental output of retrieval algorithm and possess the information of retrieval sensitivity in the previous studies (Bian et al., 2007; Monahan et al., 2007; Divakarla et al., 2008; Pittman et al., 2009).

Nevertheless, as suggested, we have applied the Gaussian smoothing to ozonesonde observations with a Gaussian distribution FWHM close to AIRS vertical resolution (~5 km, upper troposphere). Below Figure 1 shows the relative difference (RD) between AIRS ozone retrieval and smooth ozonesonde with Gaussian smoothing [O3sonde (GS)] and averaging kernel smoothing [O3sonde(AK)] for the 2011-2017 period. The RD looks smoother in the AK method than in the Gaussian method. Though the average RD profile in both the smoothing is more or less similar, the seasonal RDs are very different, which could be due to the low pass filter nature of Gaussian smoothing and Apriori contribution in AKs smoothing. In the revised MS, we have added the discussion on the choice of smoothing, and some discussion is added on the Gaussian smoothing in section 2.2.

Additionally, we agree with you that in the layer average mixing ratio and columnar ozone, the AKs smoothing must not be applied. Now smoothing is removed in all the layers and columns in the revised MS (Figure 6 and Figure 10 in the MS).

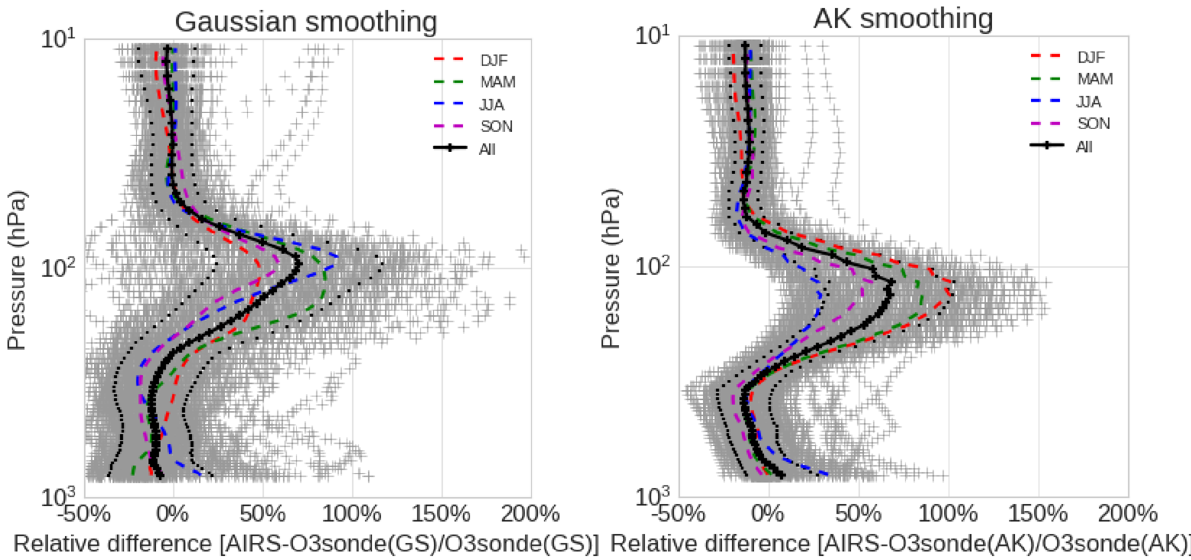

Figure 1. The relative difference of AIRS and ozonesonde with Gaussian smoothing [O3sonde (GS)] and averaging kernel smoothing [O3sonde(AK)] for 2011-2017. Individual profiles are shown by a plus sign in gray color and a dashed line for the average profile for different seasons, and a thick black line for the average of all profiles.

Boynard, A., Clerbaux, C., Coheur, P.F., Hurtmans, D., Turquety, S., George, M., Hadji-Lazaro, J., Keim, C. and Meyer-Arnek, J.: Measurements of total and tropospheric ozone from IASI: comparison with correlative satellite, ground-based and ozonesonde observations. *Atmospheric chemistry and physics*, *9*(16), pp.6255-6271, 2009.

Bak, J., Baek, K.H., Kim, J.H., Liu, X., Kim, J. and Chance, K. Cross-evaluation of GEMS tropospheric ozone retrieval performance using OMI data and the use of an ozonesonde dataset over East Asia for validation. *Atmospheric Measurement Techniques*, *12*(9), pp.5201-5215, 2019.

Zhao, F., Liu, C., Cai, Z., Liu, X., Bak, J., Kim, J., Hu, Q., Xia, C., Zhang, C., Sun, Y. and Wang, W.: Ozone profile retrievals from TROPOMI: Implication for the variation of tropospheric ozone during the outbreak of COVID-19 in China. *Science of The Total Environment*, *764*, p.142886, 2021.

Adams, C., Bourassa, A.E., Sofieva, V., Froidevaux, L., McLinden, C.A., Hubert, D., Lambert, J.C., Sioris, C.E. and Degenstein, D.A.: Assessment of Odin-OSIRIS ozone measurements from 2001 to the present using MLS, GOMOS, and ozonesondes. *Atmospheric Measurement Techniques*, *7*(1), pp.49-64, 2014.

(Here, we have listed additional references only those are used in the response part, references those are available in the MS are not listed here. Similar practice is followed further.)

Detailed Comments:

1) Short Summary: I have not seen any compelling evidence that AIRS does "well in the lower troposphere and stratosphere" at their site.

Thanks. We wanted to convey it in the relative terms, when compared with the upper troposphere and the lower stratosphere. We have now revised this sentence as "AIRS is shown to overestimate ozone in the upper troposphere and lower stratosphere, while the differences with ozonesonde are lower in the middle troposphere and middle stratosphere". This statement is from the statistical analysis (MS Figure 7), where we see relatively lower biases and lower standard deviation in the middle troposphere and middle stratosphere between ozonesonde and AIRS ozone retrieval. In addition, the relative difference at broad layer average (MS Figure S6) and relative difference profile (MS Figure S5) also shows lower differences between the two measurements in the middle troposphere and middle stratosphere.

2) Abstract: Worth mentioning the total number of sondes. These sondes, combined with sondes from other sites in India constitute a unique resource not only to evaluate satellite data but to understand the transport of ozone over north India. As I have noted above, I do not agree with the statement that "AIRS can provide quality data of ozone in the lower and middle troposphere and stratosphere" at their site. The statement "similar to AIRS, Infrared Atmospheric Sounding Interferometer (IASI) and Cross-track Infrared Sounder (CrIS) are also able to produce ozone peaks and gradients successfully" may be true at other locations, but no compelling evidence has been presented to show that it is true at their site. The statement "the monthly variations of columnar ozone (total, UTLS, and tropospheric) are captured well by AIRS, except the total columnar ozone" is confusing. It should say that monthly variation of column ozone at their site is not captured well by AIRS. The evidence that AIRS measures UTLS and tropospheric layer ozone well needs stronger justification.

We thank you for appreciating our efforts towards continuous balloon-borne ozone soundings over the complex Himalayan terrain. Following your suggestion, the total number of sondes is mentioned in the abstract.

Similar to the previous response, the sentence "AIRS can provide quality data of ozone in the lower and middle troposphere and stratosphere" is revised to "AIRS has lower difference with ozonesonde ozone in the lower and middle troposphere and stratosphere with nominal underestimations of less than 20%".

Regarding the ozone peak and gradient, we have estimated the ozone gradient and the below figure 2 shows the vertical distribution of the running ozone gradient. The gradient profiles

are more-or-less similar during four seasons. The estimated annual average ozone gradient in regions between tropopause to gradient peak are 231.5 ppbv/hPa, 199.0 ppbv/hPa, 193.2 ppbv/hPa and 199.1 ppbv/hPa for ozonesonde, AIRS, CrIS, and IASI, respectively. Similarly the ozone peak altitudes are 11.35 hPa, 10 hPa, 9.11 hPa, and 7.78 hPa for ozonesonde, AIRS, IASI, and CrIS, respectively. We have now added these information in the revised MS (section 3.4).

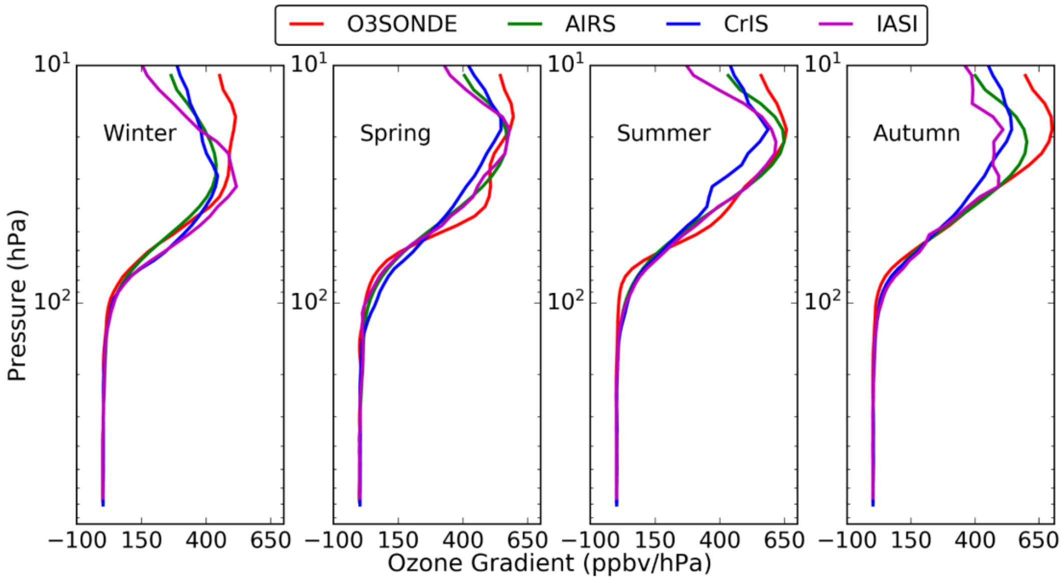

Figure 2. Ozone gradient profile along the AIRS RTA pressure levels from ozonesonde, AIRS, CrIS, and IASI.

About the monthly variations of columnar ozone, we have now revised this sentence. The revised sentence is "Furthermore, AIRS fail to capture the monthly variation of the total ozone column, with a strong bimodal variation, unlike unimodal variation seen in ozonesonde and Ozone Monitoring Instrument (OMI). In contrast, the UTLS and tropospheric ozone column are in reasonable agreement."

In addition, though there are persistence biases, particularly for the UTLS column, the correlation of UTLS (between AIRS and ozonesonde) is very strong (0.75) (below figure 3). In addition, we have performed an additional estimate for the correlation at each pressure levels in UTLS region and r2 is 0.82.

About the tropospheric column comparison, Figure 10c in the original MS shows monthly variation in ozonesonde based tropospheric column ozone using "two" tropopause (i) sonde based tropopause (ii) AIRS based tropopause. It is clear that monthly variation in tropospheric ozone column with AIRS based tropopause shows much better agreement in comparison of sonde based tropopause. The correlation (below figure 3) between AIRS and

ozonsonde is much better (0.72) when used AIRS tropopause. We have added this information in the revised MS (section 3.5.3 and Table S4)

Hence we feel that the AIRS UTLS and tropospheric ozone column information are reasonably agreeing with ozonesonde. Nevertheless, we have now revised the sentence in the abstract as mentioned above.

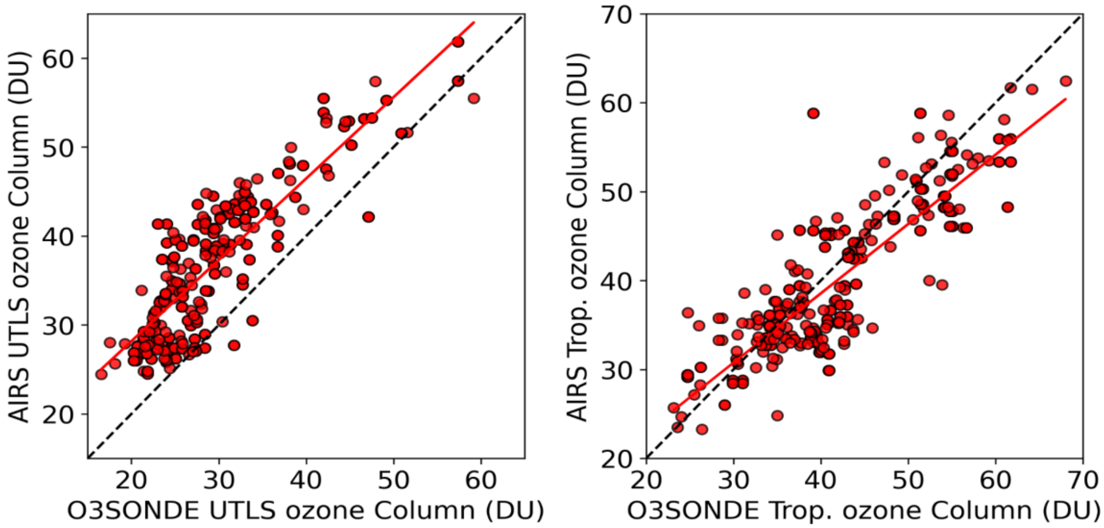

Figure 3. UTLS ozone column correlation (left) and the tropospheric ozone column (right) between AIRS and ozonesonde.

3) Table 1: The caption needs to indicate what is mean by the numbers following ± sign. I assume they are standard deviations, not standard error of the mean. In that case the standard error would be much smaller and even small differences would become statistically significant. As discussed above, the agreement in the lower layers does not necessarily imply that AIRS is doing a good job. It may only imply that AIRS AP is consistent with ozonesonde. Large differences near 100 hPa is a concern, since it implies some sort of problem with the AIRS retrieval algorithm.

We again thank you. Yes, the numbers following the ± sign are standard deviations. Now as suggested, we have estimated the standard errors and used in the revised MS.

We agree with you and this has also been described in above few comments. We have also modified the sentences (section 3.3) and abstract accordingly.

4) Table 2: If these values were derived after applying AK to the ozonesonde data, then it would be very useful to provide the values with and without applying AK, since the latter values are what a user of AIRS data would actually care about. It makes no sense to me to average the MR in the 10-100 hPa layer. Since the MR drops by nearly two orders of magnitude between 10 and 100 hPa, the average would essentially be the value near 10 hPa. It is much better to compare the ozone column in this layer (without applying AK).

Thank you very much. Table 2 shows the R^2 values of ozonesonde with AIRS, IASI, and CrIS, respectively, without applying AKs. We have now mentioned this in the caption of table 2. As indicated, we have now divided the 100 - 10 hPa region into the two layers (lower stratosphere (100 - 50 hPa) and middle stratosphere (50 - 10 hPa)). We have also revised other figures (Figure 6, Figure 8, and Figure S6) and added this information in sections 3.2 and 3.3) in the revised MS.

5) Table 3: In comparing columns one should not apply AK.

We thank and appreciate this suggestion. We agree with the reviewer that the smoothing should not be applied in comparing columns. We have now revised the table 3 and similar changes are also done in supplementary tables (Table S4 and S5).

6) Figure 2c: This figure very clearly shows the problem one has in interpreting AIRS ozone profile data. Since the AKs peak near the ozone density peak, the primary information contained in AIRS measurement is the column ozone amount. The profile information is extremely limited. However, if the variability of (log of) ozone near the peak is small, the secondary peak at 200 hPa may help capture some of the variability near that level. While the short-term variability of O3 near the density peak is probably quite small (this needs to be checked using sonde data), it is important to note that QBO in O3 occurs near the ozone density peak. So, the peak in the AK near the peak may introduce QBO like signals at the lower levels.

The AIRS ozone averaging kernels are calculated and utilized for AIRS ozone evaluation over the central Himalayan region. To our knowledge, the AIRS ozone AKs at all 100 RTA layers are constructed and discussed here for the first time. Generally, Averaging Kernel (AK), a measure of information contents of retrieval, is calculated using multiplication between error covariance matrics and radiance jacobians, i.e., $[S_x \cdot K_n^T \cdot (K_n \cdot S_x \cdot K_n^T + S_\varepsilon)^{-1} \cdot K_n]$. In each AIRS profile retrievals, the error covariance matrices will be nearly same depending on apriori informations, while the radiance jacobians will be slightly different. Hence for each retrieval, a little different shape of AKs is expected, with nearly similar information contents. We agree with the reviewer that the AIRS ozone retrieval is more

sensitive to stratospheric ozone still, the second peak in the upper troposphere has the capability to capture ozone features. AIRS tropospheric ozone retrieval is utilized by various studies to see the events-based ozone enhancements, i.e., Phanikumar et al. (2017) over the balloon launch site (Nainital) utilizes the AIRS ozone measurments to confirm the two folds enhancements of tropospheric ozone due orography induced gravity waves. Li et al. (2018) also utilizes the AIRS middle tropospheric ozone to study the high tropospheric ozone in Lhasa due to convective transport and stratospheric intrusion, etc. Additionally, we studied the ozone variability near 50 hPa (AKs peak altitude) from ozonesonde, which is about 342 ppbv (standard deviation with a mean of about 1630 ppbv), while with logarithmic values, it is 0.2. A typical variability of 20% is seen around the mean ozone mixing ratio at 50hPa.

Phanikumar, D.V., Kumar, K.N., Bhattacharjee, S., Naja, M., Girach, I.A., Nair, P.R. and Kumari, S.: Unusual enhancement in tropospheric and surface ozone due to orography induced gravity waves. *Remote Sensing of Environment*, *199*, pp.256-264, 2017.

Li, D., Vogel, B., Müller, R., Bian, J., Günther, G., Li, Q., Zhang, J., Bai, Z., Vömel, H. and Riese, M.: High tropospheric ozone in Lhasa within the Asian summer monsoon anticyclone in 2013: influence of convective transport and stratospheric intrusions. *Atmospheric chemistry and physics*, *18*(24), pp.17979-17994, 2018.

7) Figure 5: The caption should clarify what do the error bars mean. They should show standard error of the mean not standard deviation. It appears that AIRS provides just the AP value in the troposphere, as one expects from the AKs.

Thanks. Following the reviewer's suggestion, we have now changed the standard deviation by the standard error in the revised Figure 5. In the optimal estimation method, the apriori ozone profiles are modified to match the true atmospheric ozone by minimizing the cost function. Based on the weighting function of particular satellite instruments the apriori is modified at various altitude levels. In general, the ozone weighting function is low in lower troposphere, even the present hyperspectral satellite instruments cannot provide lower tropospheric ozone information. However, there have been some attempts to utilize the synergic observations of infrared and UV-VIS satellites to maximize the retrieval sensitivity to lower tropospheric ozone , as in the case of the synergic ozone retrieval from IASI + GOME-2 (Causta et al., 2013; Rawat et al., 2021) and AIRS + OMI (Fu et al., 2018). In the revised MS (section 3.2) we have briefly described the contribution of apriori in the lower troposphere and the constraint with hyperspectral retrieval.

Cuesta, J., Eremenko, M., Liu, X., Dufour, G., Cai, Z., Höpfner, M., von Clarmann, T., Sellitto, P., Forêt, G., Gaubert, B. and Beekmann, M., 2013. Satellite observation of lowermost tropospheric ozone by multispectral synergism of IASI thermal infrared and GOME-2 ultraviolet measurements over Europe. *Atmospheric Chemistry and Physics*, *13*(19), pp.9675-9693.

Fu, D., Kulawik, S.S., Miyazaki, K., Bowman, K.W., Worden, J.R., Eldering, A., Livesey, N.J., Teixeira, J., Irion, F.W., Herman, R.L. and Osterman, G.B., 2018. Retrievals of tropospheric ozone profiles from the synergism of AIRS and OMI: methodology and validation. *Atmospheric Measurement Techniques, 11*(10), pp.5587-5605.

8) Figure 6: Delete the top panel. (See comment no 4.) The results plotted in the second and 3rd panels are hard to see. To make it clearer remove the error bars (they are not errors anyhow) and the dashed vertical Iines. It is not clear why the data are doubly averaged. If one wants to shows the mean MR in a layer, show just the mean MR without applying AK to the sonde data. This is what a user cares. But if the purpose is to evaluate the AIRS algorithm and calibration, show the MR at a single pressure level after applying AK. (See overall comments.)

Thank you very much. Following your suggestion, we have now removed the smoothed ozonesonde values, error bars, and vertical lines in figure 6. We agree with the reviewer that the average ozone mixing ratio between the 100-10 hPa region will be dominated by the ozone mixing ratio near the 10hPa region. In the revised MS, we have divided the region into the lower stratosphere (100 - 50 hPa) and middle stratosphere (50 - 10 hPa). In addition, the correlation is now in between the relative difference of ozone mixing ratio (ozonesonde and AIRS) and MI/TWV. Similarly, figure S6 is also revised.

9) Figure 7: It would be useful to plot the mean difference between sondes and MLS on the left panel. This will tell us if the sondes agree with a much higher vertical resolution satellite instrument. If not this will either imply problems with sonde data or more likely the complexity of doing satellite retrievals near their site. Recommend deleting the middle panel. In the right panel show the std devs desperately from sondes, sonde AK and AIRS to see if AIRS is at least capturing the variability irrespective of the bias. A figure showing r2 would also be useful.

Thank you for recommending this. The mean biases between ozonesonde and MLS are added in Figure 7, and the middle panel of RMSE is removed in the revised MS. The R2 profile is discussed on the right of figure 8. The mean biases between ozonesonde and MLS, a high vertical resolution satellite instrument, are smaller and MLS agrees well with ozonesonde. We have added the MLS differences with ozonsende in section 3.3. In addition, we find the statistical analysis in the previous MS was by fault, selected for a shorter time. Now, in the revised MS, the complete period from Jan 2011 to Dec 2017 is included in calculating the bias and STD.

10) Figure 8: Same comment as for Figure 6.

Thanks. In the revised MS, we have now limited the layer up to 50 hPa, instead of up to 10 hPa. Now it shows the lower stratospheric region (100 - 50 hPa).

11) Figure 9: same comment as for Figure 7.

Thanks. We have now removed the middle panel of RMSE and the figure is revised as suggested by the reviewer.

12) Figure 10: This is arguably the most important figure of the paper. Please try to improve the figure so it is easier to evaluate. See discussion in overall comments.

Thanks. We have now revised the figure 10. In the revised figure, we have now removed monthly variation with AK as suggested by you in previous comments. We agree that this has also improve its visibility significantly. We have now revised the caption also.

13) The figure captions should be self-explanatory. One shouldn't be required to hunt in the text to understand the figures.

Thank you for your suggestions. We have now revised the caption and added the needed information.

14) The paper requires careful editing. I see citations with no references and references with no citations.

Thank you for your careful reading and for pointing out the mismatched citation and references. We are sorry for the same and suitable revision has been done in the revised MS.