Review of: Highly resolved mapping of NO2 vertical column densities from GeoTASO measurements over a megacity and industrial area during the KORUS-AQ campaign (Choo et al., 2022)

The manuscript discusses results from the KORUS-AQ campaign and demonstrates for the first time the NO2 VCD horizontal distribution over highly urbanized / industrialised regions in South Korea, based on airborne imaging data. The scientific content of the paper fits within the scope of AMT, although the study lacks novelty. The paper is short on the actual data retrievals and focuses more on the interpretation of the NO2 spatiotemporal variability over the different regions. Regarding the VCD retrievals, a number of choices/assumptions are made without proper argumentation which are not consistent with other similar studies that included extensive explanations/analysis. This can affect the results and the made conclusions, especially regarding the absolute VCD values. I'm referring for example to the fitting interval, reference SCD, surface reflectance product. Also the error analysis has some serious flaws and the 'validation' lacks a lot of details in order to assess the validity of the comparisons. Overall, I would recommend publication in AMT. However, major revisions (detailed below) need to be conducted in the paper before publication. Please note that all questions below should not only be addressed in the author's reply, but also in the manuscript.

General comments

-Please extend section 2.1. with campaign information. Some information about the campaigns is scattered in the manuscript, but a clear campaign section shortly discussing the number of flights, time and duration of flights, ROI, SZA change during flights, environmental conditions, e.g. cloud fraction, etc. is missing while it would improve interpretation. Also a table would help indicating the different flights, their time of flight and the region of interest. This would also help to understand certain choices made for the comparisons (see later comments on that).

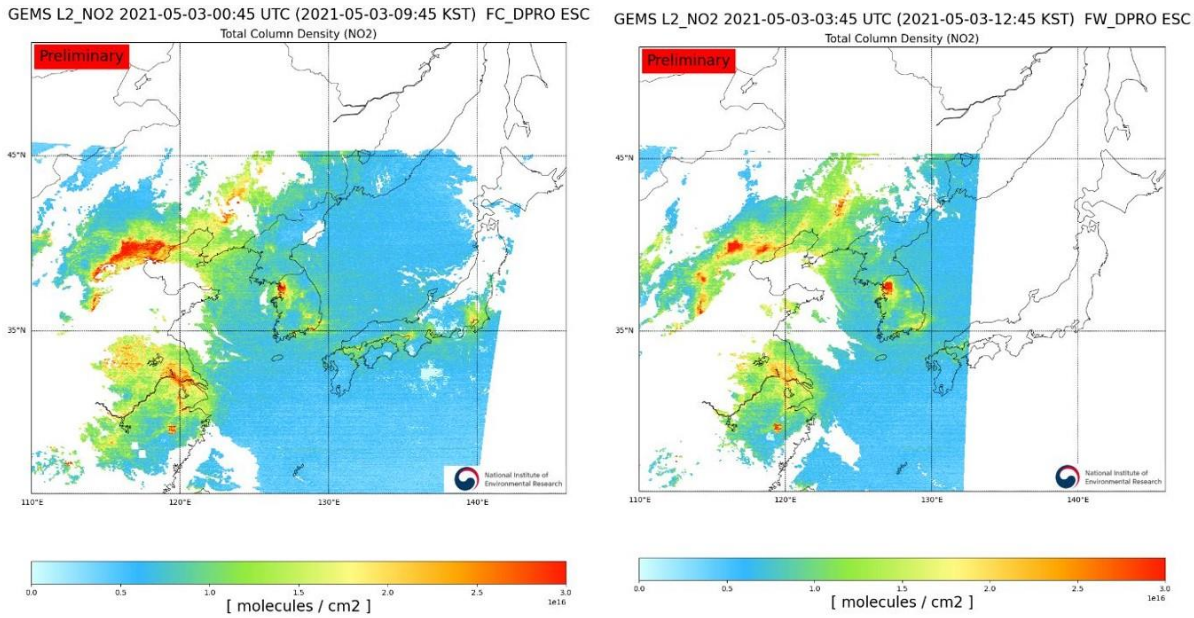
Response: First of all, we sincerely apologize for the late submission.

We thank the reviewer's kind comment and advice. Additional descriptions of the campaign and flight information (number, time, altitude, route, and etc.) have been added into the manuscript. Please, see P. 5, Line 125. The detailed information on each flight date is on the http://www-air.larc.nasa.gov/missions/korus-aq/docs/KORUS-AQ_Flight_Summaries_ID122.pdf as well as written in the manuscript.

-p.5 l.150: One reference is used to analyse the whole data set if I understand it correctly. Three questions that should be properly addressed: 1) do you use one spectrum or an average of spectra over your reference area in order to improve the SNR? ; 2) What is the residual amount in the reference spectrum used and how was this determined. Do you use the value 6.751 x 1015? This is not really clear from the manuscript. This value seems high for a considered background area over the ocean and can result in an overestimation of all retrievals shown in this work. Have you compared with the OMI value retrieved for that day or an average + st. dev. OMI value to get an idea on the NO2 variability for this area? ; 3) for airborne hyperspectral imagers often a daily reference is preferred as the spectral properties of the instrument can change resulting in an along-track and across-track drift. Please prove that this is not the case with this instrument and

the right choice was made, i.e. by showing RMS of the fit on the day when the reference was taken and RMS for flights further away from the reference date.

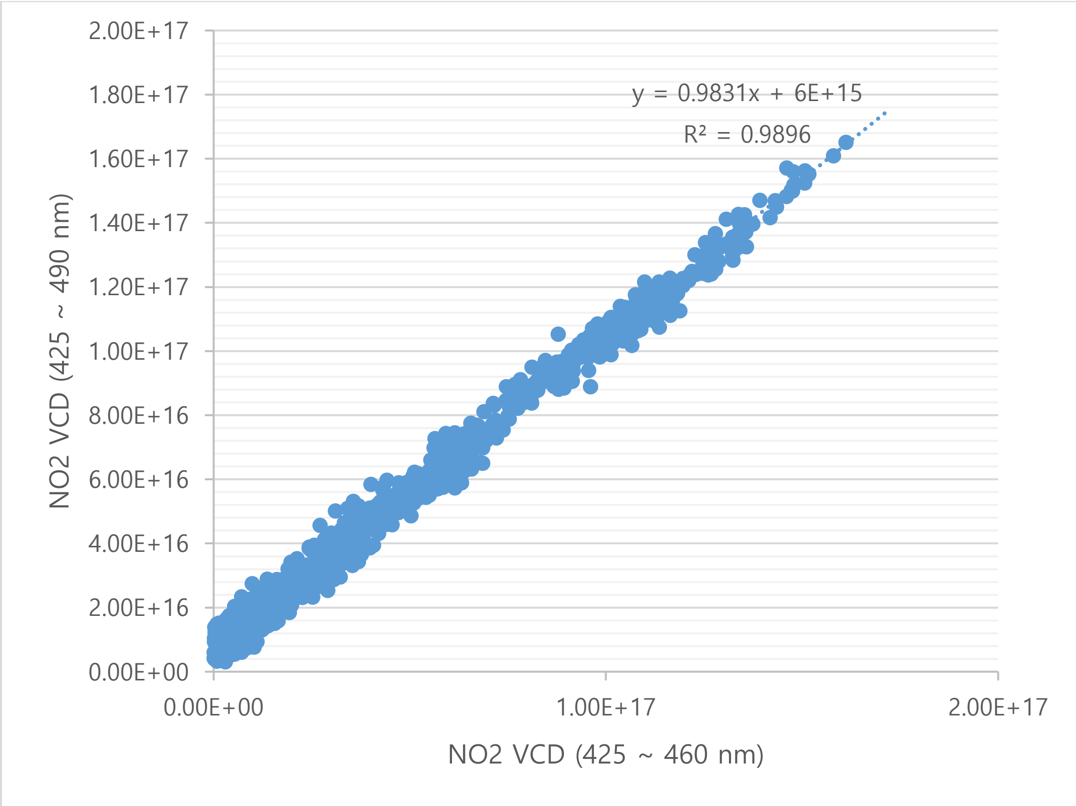
Response: In this present study, we used reference observed radiance averagedby 250 m × 250 m in each 33 across tracks over South Ocean of Jeju Island wihich is one of the most clean region in south Korea. The NO2 VCD and standard deviation in OMI in this region (SZA<85, cloud fraction < 0.5) is 4.77 × 10^{15} molec. cm^{-2} and 1.33 × 10^{15} molec. cm^{-2}, respectively. When radiance over the high NO_2 region is used as a reference, the fitting results normally are underestimated. I think you see the NO_2 VCD is overestimated in Seoul metropolitan regions especially background regions in the afternoon. But, the NO_2 VCD has a very high value in GEMS and sometime those has large values in the afternoon than morning time (Left, 3, May, 2021, 09 UTC: right, 3, May, 2021, 12 UTC).



As you mentioned, I agree that it is good to set a reference at each flight, but as you can see in figure 1, most Geo-TASO observation do not cover the clean area.

-p.5 l.155: This NO2 spectral window is not very optimal and moreover very narrow, while the instrument properties allow a more optimal and larger window for NO2 fits, e.g. like the 425-490 nm recommended in the literature. Please show some fit results with this larger window (or similar larger window) and show the VCD differences with the narrow window to clarify the impact of the choice made and also give a clear motivation for choosing this non-optimal window 425-450 nm for your retrievals.

Response: We also agree with your comment that NO_2 VCD resulted from fitting using a wide spectral window (ex, 425 ~ 490 nm) is more stable, especially space borne measurement. But in this present study, we prepared all dataset (a priori AOD, airmass factor calculation) with spectral window from 425 nm and 460 nm. I compared the retrieved NO_2 VCD of 16 cross tracks in 5, June between spectral window from 425 ~ 460 nm and those from 425 ~ 490 nm as follows:



If you think that NO_2 VCD should be retrieved by using fitting window between 425 nm and 490 nm, we will perform AOD re-calculation by using CMAQ and also we will re-calculate AMF.

p.6 l.175: It is shown in several studies, some of them cited in this work, that surface reflectance can have a very strong impact on the AMF. Is GeoTASO not absolutely calibrated in order to derive a surface reflectance product from the instrument itself? If not, why isn't the MODIS product MCD43A1 or MCD43A3 used like during the LISTOS campaign? It provides a much higher spatial resolution matching better the spatial resolution of your retrievals and it is proven to be a reliable product. I'm worried that the coarse resolution product you have chosen strongly impacts your AMF, thus also your VCD retrievals.

Response: We agree with reviewer's comments, thus tried our best to answer your questions. First of all, as you mentioned, errors due to AOD, SSA, surface reflectance, etc., which are used as input data when calculating AMF, may affect AMF results. Therefore, in this study, the uncertainty according to each input variable was calculated and provided in Section 3.2. In our result, the variable with the largest AMF error was aerosol loading height (26.4%), which is SSA (4.2%), AOD (3.0%), and surface reflectance (2.8%). In addition, the average of the surface reflectance during

KORUS-AQ period is 0.055, and there is a study using a fixed ground reflectance of 0.05 to retrieve SO_2 VCD (Chong et al., 2020).

GeoTASO does not make any corrections to calculate surface reflectance from the device itself. However, GeoTASO data for NO_2 retrieval were spectrally, geometrically, radiometrically calibrated at the NASA Goddard Space Flight Center. The algorithms for calculating observed surface reflectance using GeoTASO data and surface reflection results were not provided. We tried to use surface reflectance (500 m resolution) of MCD43A3, but there were many pixels that were not produced within the observation area. We wanted to know the spatial distribution of NO_2 VCD. Therefore, we used MOD09CMG and MYD09CMG. Because the data exist in all pixels, the surface reflectance.

- Chong, H., Lee, S., Kim, J., Jeong, U., Li, C., Krotkov, N. A., Nowlan, C. R., ... & Koo, J.-H. (2020). High-resolution mapping of SO2 using airborne observations from the GeoTASO instrument during the KORUS-AQ field study: PCA-based vertical column retrievals. Remote Sensing of Environment, 241, 111725.

Sect. 2.5: You clarified later on that SSA and AOD are also derived from CMAQ. This should be described here as well, besides the vertical NO2 profile.

Response: As reviewer suggested, sentences and equations explaining the values (SSA, AOD, and vertical NO_2 profiles) have been added to the revised manuscript in P. 10, Lines 237-246.

After modification (P. 10, Lines 237-246):

"CMAQ AOD was calculated by integrating the aerosol extinction coefficient (Q_{ext}), which is the sum of scattering (Q_{sca}) and absorption (Q_{abs}) coefficients, over all vertical layers (z) as follows:

$$\text{AOD} = \int Q_{sca}(z)\, dz \; = \int \{Q_{sca}(z) + Q_{abs}(z)\}\, dz \tag{9}$$

$$Q_{abs}[\text{Mm}^{-1}] = \sum_i \sum_j \left\{(1-\omega_{ij}) \cdot \beta_{ij} \cdot f_{ij}(RH) \cdot [C]_{ij}\right\} \tag{10}$$

$$Q_{sca}[\text{Mm}^{-1}] = \sum_i \sum_j \left\{\omega_{ij} \cdot \beta_{ij} \cdot f_{ij}(RH) \cdot [C]_{ij}\right\} \tag{11}$$

Here, ω_{ij} indicates SSA of particulate species i for the particulate mode (or size bin) j, β_{ij} denotes the mass extinction efficiency, $f_{ij}(RH)$ is the hygroscopicity factor at the relative humidity (RH), and $[C]_{ij}$ is the concentration of particulate species. CMAQ SSA is defined as the ratio of the integrated Q_{sca} to AOD, and NO2 vertical profiles were obtained from NO2 concentrations at each vertical layers by conducting CMAQ simulations. More details of the model descriptions are shown in Lee et al. (2020) and Malm and Hand (2007)."

- Malm, W. C. and Hand J. L.: An examination of the physical and optical properties of aerosols collected in the IMPROVE program, Atmospheric Environment, 41, 3407– 3427, https://doi.org/10.1016/j.atmosenv. 2006.12.012, 2007.

Sect 3.1: The spatial binning (0.01°) has not been explained in the manuscript, but only mentioned in the relevant figure captions. Do you bin spectra prior to the retrievals to increase the SNR or do you bin the VCDs afterwards? In the latter case, why do you bin the data (native resolution of 250 m) to this coarser resolution? You are losing some spatial detail while NO2 is a species with strong spatiotemporal variability.

Response: We apologize for the confusion. We agree with the reviewer's advice and have therefore revised to the manuscript in P. 10, Line 249. We did spatial binning after calculating NO_2 VCD. In Chong et al. (2020), we have SO_2 VCD by using data observed in GeoTASO during the KORUS-AQ period, and binning was performed at a resolution of GEMS (7 x 8 km^2). As a result of previous studies, it was found that relative random uncertainty decreased as SO_2 VCD increased. This random uncertainty can be reduced even when additional spatial binning is performed, which provides a balance between random error and spatial resolution. In conclusion, the study showed that larger VCDs at 250 m resolutions do not necessarily lead to larger VCDs at 7 km x 8 km resolutions. Additionally, we showed the AMF which is native resolution by modifying Fig. 5(c). Responding to your comments.

After modification (P. 10, Line 249-252):

"We showed the finally NO_2 VCDs by binning them with 0.01° × 0.01° from 250 m spatial resolution. Although the spatial binning NO_2 VCDs were compared to those at native resolution, we noted that the spatiotemporal variability was still able to be clearly distinguished from the background at 0.01° binning resolution. Chong et al. (2020) showed that larger VCDs at 250 m resolutions do not necessarily lead to larger VCDs at wider resolutions."

- Chong, H., Lee, S., Kim, J., Jeong, U., Li, C., Krotkov, N. A., Nowlan, C. R., ... & Koo, J.-H. (2020). High-resolution mapping of SO2 using airborne observations from the GeoTASO instrument during the KORUS-AQ field study: PCA-based vertical column retrievals. Remote Sensing of Environment, 241, 111725.

p.7 l.204 and Figure. 4: I would like to see a zoom on figure 4 on the expressways (eventually at the native spatial resolution). In Figure 4, it is difficult to judge if there is a stronger NO2 pattern downwind than upwind from the expressway line sources, but this is something that should be present in the data if the made statements are correct. You could also do it for the afternoon with a different dynamic scale for the VCD map.

Response: We have indicated the NO_2 VCD with the native resolution (250 m) in Fig. R1 responding to your comments. At this time, the VCD map (Fig. 4(b)) in the afternoon showed NO_2 VCD with a different dynamic scale than before. Additionally, (2) Seohaean expressway in Fig. 4(a) is enlarged and represented at the same color scale of the same NO_2 VCD (Fig. R2). In the enlarged Fig. R2, it was found that the NO2 VCD was relatively higher on the highway than on the surrounding mountains. We showed wind direction and wind speed at 1000 hPa level in the morning and afternoon respectively. It can be seen that the wind directions are different in the morning and afternoon on 9, June 2016, and the wind speed is relatively higher in the afternoon than in the morning.

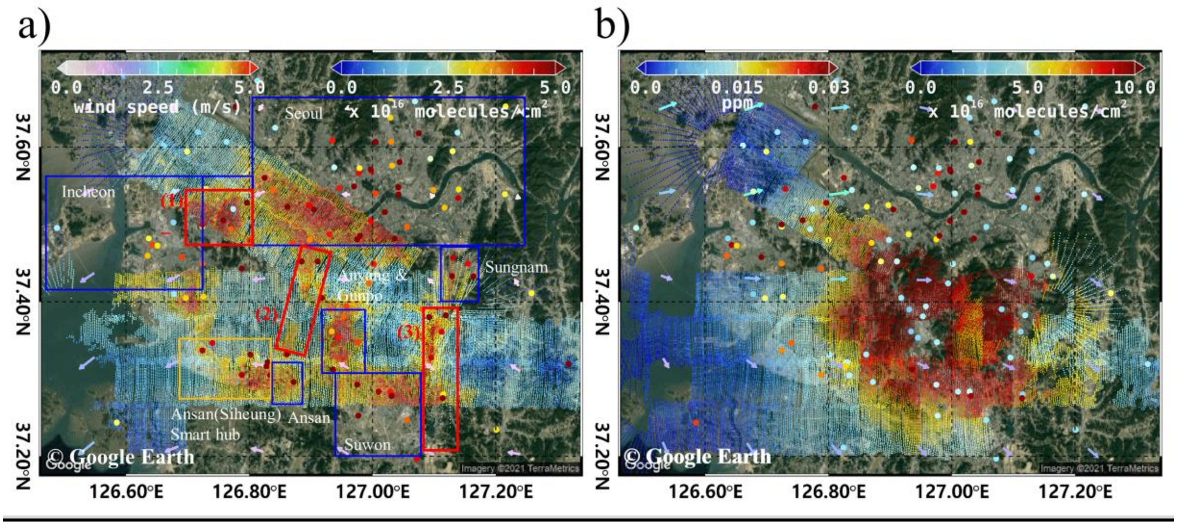


Fig. R1. Similar to Fig. 4 except that the native spatial resolution (250 m).

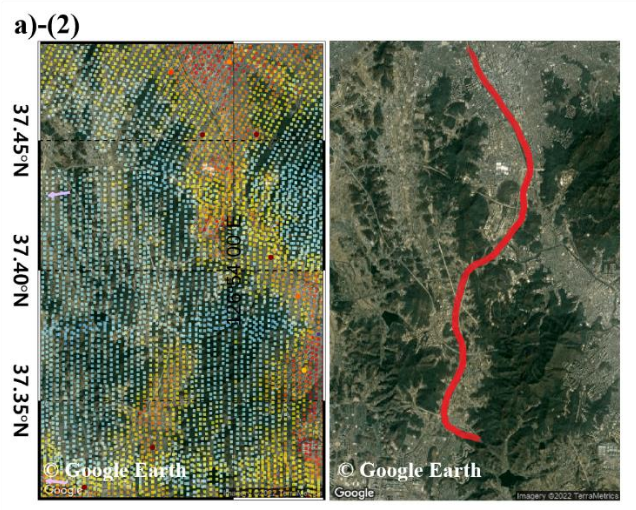


Fig. R2. A zoom in Fig. R1-(2) Seohaean expressway (the red line).

P.7 l.206: So it means the background NO2 VCD is around 1 x 1016? This seems pretty high (see also earlier comment on the SCDref). Can you demonstrate that this is somehow consistent with OMI/TROPOMI retrievals? Please do the same for statement on p. 9 l. 255

Response: I understand your opinion that the background NO_2 VCDs are very high. And unfortunately, in this period, TROPOMI was not launched yet, OMI did not provide NO_2 data except for a very few days. We re-retrieved NO_2 VCDs during review period, but NO_2 VCDs still remained high in Seoul metropolitan region. But as you can see Goldberg et al., 2019 this figure 6, the averaged NO_2 VCDs in May, 2016 retrieved OMI have large values in Seoul metropolitan region, and also they have larger values over 3 x 10^{16} molec. cm^{-2} when AMF is adjusted. It is irrelevant with this paper, but sometime GEMS NO2 values are extremely large in Seoul metropolitan region, especially in the afternoon. We also validate this large values by using ground-based measurements in GEMS Map of

Air Pollution (GMAP) and Satellite Integrated Joint Monitoring Air Quality (SIJAQ) campaign which has been carried out since 2019. I also think it is necessary to keep finding the reason why NO_2 VCDs increase in the afternoon in Seoul metropolitan region although NO_2 VMR observed from in-situ measurements decreases.

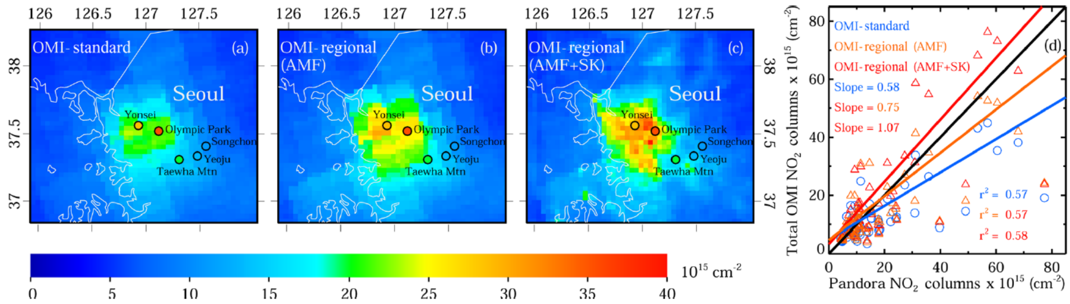


Figure 6. **(a)** Total vertical column contents from the OMI-standard NO_2 product for May 2016, **(b)** same quantities from the OMI-regional product with only the air mass factor adjustment (AMF) during the same time frame, **(c)** same quantities from the OMI-regional product with the air mass factor adjustment and spatial kernel (AMF + SK) during the same time frame, and **(d)** a comparison between total column contents from the three OMI NO_2 products and Pandora NO_2 during May 2016. An average of Pandora 2 h means co-located to valid daily OMI overpasses are overlaid in the spatial plots.

p.9 l.364: The error on the SCDref is an important source of uncertainty as well, not taken into account in this study. Please include.

Response: we added this sentence "(averaged NO_2 VCD obtained from OMI available during KOURS-AQ period is 4.77 × 10^{15} molecules cm^{-2} and standard deviation of 1.33 × 10^{15} molecules cm^{-2}, respectively)"

p.9 l.273: Please clarify how these percentages were obtained. This is unclear. An important uncertainty regarding the AMF is also the uncertainty related to the a priori NO2 profile shape. This is mentioned as well in the conclusion but it is completely ignored in this uncertainty study. Please include.

Response: The NO_2 AMF errors were recalculated based on the uncertainties of parameters (AOD, SSA, ALH, and surface reflectance) obtained from previous studies. Relevant references for the uncertainty were also inserted in the revised manuscript (Page 17-18, lines 377–379):

"The σ of AOD, SSA, surface reflectance, and ALH are assumed as 30% (Ahn et al., 2014), 0.04 (Jethva et al., 2014), 0.005+0.05×surface reflectance (EOS Land Validation; https://landval.gsfc.nasa.gov), and 1 km (Fishman et al., 2012), respectively, in this study."

Our paper more focused on the NO_2 retrieval from GeoTASO and the investigation of its spatial distribution. Therefore, for now, we considered AOD, SSA, ALH, and surface reflectance in the error analysis section. The authors fully agree with the necessity of AMF error analysis due to the uncertainty related to the a priori NO_2 profile shape and preparing it now for the next paper. Some sentences were added to point out the necessity related to NO_2 profile shape (Page 19, lines 404–406):

"A priori NO_2 profile shape also can be one of factors to cause calculation error for NO_2 AMF as reported in the previous studies (Leitao et al., 2010, Meier et al., 2016). It is necessary to calculate the effect of a priori NO_2 profile shape on airborne NO_2 AMF error in the future."

- Ahn, C., Torres, O., & Jethva, H. (2014). Assessment of OMI near-UV aerosol optical depth over land. Journal of Geophysical Research: Atmospheres, 119(5), 2457-2473.
- Jethva, H., Torres, O., & Ahn, C. (2014). Global assessment of OMI aerosol single-scattering albedo using ground-based AERONET inversion. Journal of Geophysical Research: Atmospheres, 119(14), 9020-9040.
- EOS Land Validation (https://modis-land.gsfc.nasa.gov/ValStatus.php?ProductID=MOD09)
- Fishman, J., Iraci, L. T., Al-Saadi, J., Chance, K., Chavez, F., Chin, M., ... & Wang, M. (2012). The United States' next generation of atmospheric composition and coastal ecosystem measurements: NASA's Geostationary Coastal and Air Pollution Events (GEO-CAPE) mission. Bulletin of the American Meteorological Society, 93(10), 1547-1566.
- Meier, A. C., Schönhardt, A., Bösch, T., Richter, A., Seyler, A., Ruhtz, T., ... & Burrows, J. P. (2017). High-resolution airborne imaging DOAS measurements of NO 2 above Bucharest during AROMAT. Atmospheric Measurement Techniques, 10(5), 1831-1857.

p.10 l.283: 7.3% uncertainty on the AMF is really on the optimistic side and not in line with other studies that report around 15%-20%, or larger. Please revise your calculations and/or explain. Just to clarify: it is perfectly fine if you obtain other numbers than in other studies but this should be clarified based on an in-depth analysis. Like mentioned in the previous comment it is unclear where the percentages are coming from, while in other studies these are based on discussed sensitivity tests. Are the numbers provided in Table 6 an average for the flight on 9 June 2016 or what does it represent exactly?

Response: As mentioned above, AMF uncertainties were recalculated based on the uncertainties of input parameters from previous studies. Higher AMF uncertainty (27.8%) has been gained. The numbers in Table 6 are the average for the flight on 9 June 2016. To clarify it, a sentence and the caption for Table 6 have been revised.

Caption for Table 6: "Total errors of NO_2 VCD caused by uncertainties in NO_2 SCD and NO_2 AMF (the average for the flight on 9 June 2016)."

Page 18, line 390: "Table 6 lists the estimated NO_2 VCD error on 9 June 2016 for each source based on the error propagation method."

Averages and standard deviations of the parameters (AOD, SSA, aerosol loading height, and surface reflectance) have been added.

Page 18, line 384: "On the flight day, average (standard deviation) values of AOD, SSA, ALH, and surface reflectance were 0.39 (0.10), 0.98 (0.001), 0.27 km (0.10 km), and 0.09 (0.04), respectively."

Sect. 3.3. all the performed comparisons should report how many measurements/data points are compared (eventually in the plots), in order to assess how statistically relevant the comparisons are.

Response: We revised section 3.3 as follow:

After modification (P. 21, Line 432-494):

"Tropospheric NO_2 VCDs retrieved from GeoTASO L1B data ($NO_{2,G}$) were compared with those obtained from OMI NO_2 VCDs ($NO_{2,O}$) and Pandora ($NO_{2,P}$). The $NO_{2,O}$ were only available for 10 June during the campaign period. Therefore, we only compared 37 $NO_{2,G}$ and $NO_{2,O}$ data points within a radius of 20 km and 30 min, which yielded a correlation coefficient of 0.70 with a slope of 0.41 (Fig. 9 a)). In order to validate, All $NO_{2,G}$ within a radius 20 km of the OMI center coordinate were averaged.

The NO_2 values are relatively low since GeoTAOS observation is carried out in a region with low NO_2 compared to Seoul metropolitan and the overpass time of OMI is about 13:30 LT when NO_2 decreased. It is thought that the reson the low slope value is because the OMI with low spatial resolution does not reflect the spatial NO_2 inhomogeneity in the pixel.

To validate the accuracy of $NO_{2,G}$ data, we made a comparison with NO_2 VCD obtained from the Pandora system ($NO_{2,P}$) during the KORUS-AQ campaign period. $NO_{2,P}$ obtained from Busan University, Olympic Park, Songchon, Yeoju, and Yonsei University Pandora sites on June 5, 9, and 10 were used for the GeoTASO validation (Fig. 1). $NO_{2,G}$ and $NO_{2,P}$ columns at these sites are compared in Fig. 10. In order to compare $NO_{2,G}$ and $NO_{2,P}$, we used averaged $NO_{2,G}$ retrieved from 16 across track with smallest viewing zenith angle and averaged 30 min NO_2 obtained from pandora measurement within a radius of 0.05 degree. $NO_{2,G}$ and $NO_{2,P}$ were correlated (R = 0.79, with a slope of 1.15), however, when $NO_{2,P}$ was lower than 1×10^{16} molecules cm^{-2}, the correlation coefficient between $NO_{2,G}$ and $NO_{2,P}$ was < 0.1. The weak correlation at low NO_2 levels are most likely to reflect the differences in viewing geometries and the horizontal inhomogeneity of the measured NO_2 between Pandora and GeoTASO. Also, from this result, it is thought that it can be used for NO_2 validation of geostationary satellite such as GEMS using Pandora and GeoTASO. However, since the number of pandora is limited in this campaign, we had difficulties to validate NO_2 retrieved from GeoTASO under various conditions. I believe that many ground-based remote sensing measurements are needed to validate GEMS under various conditions.

To evaluate the spatiotemporal distribution of NO_2 VCDs retrieved from GeoTASO, $NO_{2,G}$ in comparisons to surface spatial patterns, $NO_{2,G}$ was compared with $NO_{2,A}$ for GeoTASO data within a radius of approximately 0.05 km and 30 min (Fig. 9). In order to compare $NO_{2,G}$ and $NO_{2,A}$, we used averaged $NO_{2,G}$ retrieved from 16 across track and averaged 30 min within a radius of 0.05 degree. Since in-situ measurements provides NO_2 VMR ($NO_{2,A}$)(ppmv) once per hour, $NO_{2,A}$ of the nearest time is used to compare with $NO_{2,G}$.The correlation coefficient (R) between $NO_{2,G}$ (molecules cm^{-2}) and $NO_{2,A}$ at 9 AM and 3 PM LT in the Seoul metropolitan region was 0.45 and 0.81, respectively. When using only roadside station data from Air-Korea, the R-value for the morning increased to 0.83, which implies GeoTASO is more sensitive to emissions from NO_2 source areas, such as roadsides. As a result of the comparison, there were large differences in the morning and afternoon. These results were identified because synoptic meteorology played an important role from June 1 to June 10, 2016 (Choi et al., 2019). As described by Judd et al. (2018), the spatial distribution for NO_2 VCDs appears that reflects the emission source in local industrialized regions and transportations in the morning with relatively weak winds. In general, NO_2 concentration increases to late morning, indicating that the emissions process proceeds faster than the NO_2 removal process. As the planetary boundary layer heights (PBLH) in early afternoon increase and surface NO_2 is mixed through a deeper PBLH, the NO_2 VCDs distribution showed a wider increase in most of the Seoul metropolitan area and the overall column amounts continue to increase (Judd et al., 2018).

In addition, when comparing NO_2 VCDs with surface NO_2 concentrations, it should be interpreted carefully that it is a non-linear relationship between $NO_{2,G}$ and $NO_{2,A}$. Although it may vary depending on weather conditions, high NO_2 VCDs from airborne observations may sometimes be detected with low surface NO_2 concentrations. In particular, when exhaust gases emitted from industrial facilities are happen at a certain altitude (stacks/chimneys), $NO_{2,G}$ show high NO_2 VCDs, but $NO_{2,A}$ may be observed to have a low concentration. Unfortunately, in Anmyeon industrial region, $NO_{2,G}$ and $NO_{2,A}$ could not be compared due to spatial restrictions because the distribution of ground observation stations is concentrated in metropolitan areas.

In the Busan metropolitan area, the R-value of the $NO_{2,G}$ and $NO_{2,A}$ data had a correlation coefficient greater than 0.78. This reflects the more even horizontal distribution of NO_2 in the afternoon, when diffusion from the source areas had taken place. However, for a more accurate comparison, NO_2 VCD data should be converted to NO_2 MR based on mixing layer height, temperature, and pressure profile data (Kim et al., 2017; Qin et al., 2017; Jeong and Hong, 2021a). However, since the number of pandora and satellite data is limited in this campaign, we had difficulties to validate NO_2 retrieved from GeoTASO under various conditions. Since ground-based, airborne and space borne remote sensing measurements has their own advantage and disadvantage, I believe that a comprehensive observation campaign involving all of groud-based, airborne and space borne measurements should be carried out continuously for upcoming new era of geostationary environmental satellite.”

p.10 l.308: Elaborate on this part, eventually with a plot, as many things are unclear! Why are OMI NO2 VCDs only available on 10 June? How did you perform the comparison exactly? Setting a radius is a valid strategy when comparing ground based measurements with airborne or spaceborne data, but not when comparing airborne with satellite. In the latter case you should average the airborne pixels within the larger satellite pixel footprints in order to perform a fair comparison. The time constraint can be kept but could be extended to 1 hour to increase amount of compared measurements. Slope is 0.43: not clear without plot what is under- / overestimating what.

Response: OMI L2 data (V3 OMNO2) used in this research do not provide some data (30 May to 9 Jun) during the KORUS-AQ period. Therefore, we only have data on 10 June, 2016. Please, see P. 5, Lines 134-136. Also, We revised section 3.3

Sect. 3.3.1: first sentence should be rewritten as it is unclear. Again not clear: do you compare all pixels within 0.5 km and 30 min with the station measurement or do you average the airborne pixels and do you compare the averaged value? You have hourly station measurements and many flights over the whole campaign period (I think? Much details are missing on flights), but you don't seem to have many data points. Please clarify this. Why not focusing on all possible comparisons between airborne and ground-based measurements instead of restricting to AM and PM in order to improve your statistics. Like it is presented now: not really statistically relevant and without any proper discussion on the non-linear relation between columns and surface concentration, I suggest to leave this section out of the paper.

Response: We revised section 3.3

p.11 l.325: Why do you restrict to 5, 9, 10 June? You should clarify the reasons. If you only had flights over this region on these dates that is a proper explanation, but these details have not been provided.

Response: We thank the reviewer's advice and agree with you for pointing out. We conducted the research focusing on NO_2 VCD in the region of interest (megacity and industrial area). The dates (5, 9, 10 June) for calculating NO_2 VCDs were determined by performing flight observations in the morning and afternoon because the route are the same only on this three date. To help you understand, we added the content in the manuscript. The flight information is written in P. 5, Lines 123-125

(http://www-air.larc.nasa.gov/missions/korus-aq/docs/KORUS-AQ_Flight_Summaries_ID122.pdf).

Figure 5: I have a hard time to understand the AMF you obtained. Normally the AMF should be highly correlated with the surface reflectance which isn't really the case here. One (partly) explanation is maybe the rough albedo product used. But this doesn't explain the strong striping in the AMF, especially in the western part. Can you elaborate on this in the manuscript to explain the reasons?

Response: We apologize for the confusion. We agree with the reviewer's advice. The reason why the binned AMF (Fig. 5(c)) caused confusion was that there was a technical problem in the plotting process for duplicate pixels when expressing them as a mapping. Therefore, we showed the AMF which is native resolution (250 m) of GeoTASO by modifying Fig. 5(c) and we revised to the contents in the manuscript (P. 13, Line 300).

Figure 10: I would expect that PANDORA retrievals are higher than the GeoTASO retrievals as the ground-based measurements are more sensitive to the bulk of NO2 close to the surface, but it is the other way around. This could be related to the choices made in the VCD retrievals like mentioned in earlier comments which can strongly affect the absolute values. Are these findings consistent with other GeoTASO-PANDORA comparisons like from the DISCOVER-AQ or LISTOS campaign?

Response: We revised section 3.3

Minor comments

p.3 l.79: SWING has not been operated over Antwerp. The example of Antwerp shown is a simulation. SWING has been operated on a UAV over Romania. It is also better to refer to https://amt.copernicus.org/articles/11/551/2018/amt-11-551-2018.pdf

Response: After revising the manuscript as the comments of reviewers, we have corrected the “Antwerp” mistakes. Please, see P. 3, Lines 67-68.

p.3 l.81: “regional radiative transfer models” → I think you are referring to regional air quality models here?

Response: “regional radiative transfer models” has been replaced with “regional air quality models” at P. 3, Line 82.

p.3 l.89: How is transboundary pollution defined here?

Response: Transboundary pollution is pollution that originates from a country but, by crossing the border through pathways of air, is able to cause damage to the atmospheric environment in another country.

- Glossary of Environment Statistics, Studies in Methods, Series F, No. 67, United Nations, New York, 1997

p.4 l.120: Be more clear on how exactly you do the comparison: if an airborne overpass matches the 30 min constraint, do you average spatially all the pixels within 1 km radius? Or do you compare the different individual pixels with the Pandora retrieval? Same comment for p. 5 l. 131.

Response: We revised section 3.3

p.5 l.143: What is the exact reason of flying so high to measure tropospheric species? At 3 km you

would already be well above the PBL. And while indeed the higher flight altitude allows to measure the NO2 in the free troposphere as well, it has the drawback that you are losing some sensitivity towards the NO2 in the PBL and especially at the surface due to larger scattering and absorption probability.

Response: I do not know why NASA observed altitude about 9 km by using GeoTASO. However, I think as follows:

1. Although it is not available to use, but originally GeoTASO also carried out zenith measurement for reference spectrum. For zenith measurement, high altitude flights would be more effective. 2. NASA want to observe a wide area. But it is just my guess. In SIJAQ campaign, NIER have carrying out air borne observation using GCAS measurements on 3~4 km. I hope to use this data to validate GEMS.

p.7 l.210: Also the PBL plays a role and it's rise in the afternoon + transport/accumulation of emitted NO2 in the PBL during the day. But the latter is already stated later on in this paragraph.

Response: Sorry, we only stated "PBL" in P. 13, Lines 297-298.

p.8 l.240: I agree with the statement that they can be highly complementary, but it should be clarified to the reader that the relation VCD-surface concentration is highly non-linear. Depending on meteorology it is possible that strong columns are detected while surface concentrations are low. Especially in case of industrial emissions, where the emissions happen at a certain altitude (stacks/chimneys).

Response: We agree with the reviewer's advice and have therefore revised the manuscript as you suggested. We have further explained in P. 24, Lines 480-486.

p.13 l.376: Link doesn't work

Response: The address of the link is correct. Note the "_" in "~pMENU_NO=125"

Table 2 and 5: Average mileage is average mileage per car, per day?

Response: We checked average mileage per car per day.

Technical corrections

p.1 l.24: please replace 'data' by 'VCDs'

Response: As you suggested, the sentence was modified on P. 1, Line 24.

p.2 l.31: you might add 'domestic heating' as well

Response: "domestic heating" has been added to P. 2, Line 32.

p.2 l.38: cites --> cities

Response: "cites" has been replaced with "cities" on P. 2, Line 39.

p.2 l.45: what is the point of mentioning UTC here if you don't specify a time?

Response: We have deleted ″(UTC)″ on P. 2, Line 46.

p.2 l.53: is there a need to repeat here the spaceborne sensors? This was done in the previous paragraph

Response: We deleted the sentence.

p.3 l.90: replace "from May to June 2016." By "…, organized from May to June 2016."

Response: "from May to June 2016" has been replaced with "organized from May to June 2016" on P. 3, Lines 91-92.

p.3 l.94: In this study, NO2 VCD retrieval was conducted using solar backscattered radiance observed from GeoTASO over South Korea during the KORUS-AQ campaign --> This sentence is redundant, please remove.

Response: These sentence has been removed from the manuscript.

p.4 l.106: Amnyeon --> Amnyeon region?

Response: "Amnyeon" has been replaced with "Amnyeon region" on P. 5, Line 129.

p.4 l.118: O3 → O_3

Response: P. 5, Line 413 have been modified.

p.4 l.122: ". " after Notably

Response: We have added "." to P. 6, Line 149.

p.5 l.132: "." at end of sentence.

Response: We have added "." to P. 6, Line 159.

p. 6 l.184: stratospheric NO2 → stratospheric and free tropospheric NO2

Response: As you suggested, we have modified the manuscript on P. 9, Line 226.

p. 8 l.219: decrease --> decreases

Response: "decrease" has been replaced with "decreases" on P. 12, Line 292.

p. 9 l.268: Boersma et al. 2004 --> Boersma et al., 2004

Response: The citation (Boersma et al. 2004) has been corrected to "Boersma et al., 2004"

Figure 1 caption: please repeat the campaign period here and the number of flights.

Response: Figure 1 has modified following the reviewer's suggestion.

Figure 3 caption: "." at end of sentence.

Response: We have added "." to P. 8, Line 197.

Figure 4 caption: remove 'the' in 'the each panel'

Response: We have deleted "the" on P. 11, Line 261.

Figue 6 caption "into to" --> please correct + what is the meaning of the color-coding of the arrows?

Response: P. 15, Line 327 have been modified.