Answer to Referee # 1

(Reviewer's comments in BLACK, authors' reply in BLUE, corresponding modifications in the MS (manuscript) in red)

General Comments:

This study was performed based on the intensive field measurement data during the GMAP campaign in South Korea. The vision of this campaign is clear, the measurement looks well performed, and the analysis result looks interesting. But the key message is not well transferred to authors; In other words, authors need to spend more efferts to make a main root of this manuscript. This study actually focused on the relationship between the column density and surface level of not associated with the validation of satellite data (GEMS). But authors would like to underline the 'validation' purpose of GMAP, again, which is not clearly related to the main finding of this study. This mismatching makes the whole manuscript very distracted, hurts the organization, and finally results in the poor delivery of key points of this study. Thus, the major revision is largely required.

→ GMAP campaign has been actually started for multi-purpose, including GEMS validation. During GMAP-2021, multi-perspective observations were further obtained from the ground and space; participating remote sensing instruments were MAXDOAS, Car-DOAS, GCAS, and Pandora data. Regarding the revised MS, we limited our interpretation to 2020 campaign only by removing the explanation of GMAP-2021, following the reviewer's point. We appreciate the insightful and useful comments on ways to strengthen our MS. We modified or took out, following the reviewer's suggestion, as bellow.

Specific Comments:

Whole manuscript: It seems that all results in this manuscript are related the GMAP 2020 (focus on Seosan), not GMAP 2021 (related to the Seoul Metropolitan). I strongly recommend to remove the statement of GMAP 2021 here. Namely, the result of GMAP 2020 is OK enough for this manuscript. It actually makes the key point more obvious because the data analysis using GMAP 2021 was not performed here. This problem is coming from authors' emphasis on the 'validation purpose' of

GMAP campaign. Again, the results in this manuscript are not much related to the validation of GEMS data. This work mainly discussed the similarity/difference between the column and surface NO2. This topic is solely interesting enough. If the 'validation'-related wording is frequently raised in the manuscript, however, the merit of this study (analysis for the relationship between column and surface NO2) becomes weaker and readers would like to see the 'validation result' that is not included in this manuscript. → We agree with reviewer's point; thus, removed GMAP-2021 in the Fig. 1, as well as paragraph described in the MS (Line 181-182, Line 562-565, Line 710-711, 720-723, and etc, in the original MS). Thank you for pointing to this.

Title: in this context, I would recommend to change the title, only focusing on the relationship between column and in situ NO2 during the GMAP 2020 campaign. → In the title, observation has been replaced by specific species 'NO$_2$'. Thus, we changed the title to "Evaluation of correlated Pandora column NO$_2$ and in situ surface NO$_2$ measurements during GMAP campaign"

Line 6 and whole manuscript: SI implies the 'surface in situ', but here the in-situ NO2 obtained from the aircraft measurement is rather utilized for the analysis (e.g., Figs. 7 and 8), which was collected in the same GMAP 2020. → We also agree on this reviewer's point. In our text, however, there has been a need to distinguish surface 'in situ' and in-situ aircraft measurement: thus, we concurred that aircraft measurement (without mentioning 'in-situ') would be concise to deliver our results to Journal reader.

Line 9-10: This is a repetition of previous statements in line 5-6. → Done. We removed the redundant phrase. In the revised text, Line 9-10 were now replaced by 'we explored the synergy of combined analysis of both PC-NO$_2$ and SI-NO$_2$ measurements.'

Line 16-18: High wind speed and PBL height suppress the fluctuation of NO2? Why? Usually mixing is enhanced by the high wind speed, then the fluctuation becomes larger. → The planetary boundary layer was now replaced by advected marine boundary layer. This is because our study area is coastal area where westerly winds frequently cause TIBL (Thermal Internal Boundary Layer).

Line 25-27: This is not the good conclusion after the GMAP 2020 campaign. Everybody knows the difference between column and surface NO2 value. This study can have a merit because the difference between column and surface NO2 was analysed and diagnosed in detail using dense measurement of data. Please make a better statement to underline this merit. ➔ The original paragraph 'The results of this study also indicate that when performing GEMS validation using either PC or SI observations alone, particularly under prevailing local wind meteorological conditions or transport processes', has been replaced by '*The discrepancies suggest that using either PC-NO2 or SI-NO2 observations alone involves a higher possibility of uncertainty under LD conditions or prevailing transport processes*

Line 39 and whole manuscript: 'column density' sounds much better that 'column amount'➔ Done.

Line 39-41: Is there no equation number? And what is the reference of this equation? ➔ As a reference, Zhao et al. (2019) has been added. Zhao et al. (2019) defined this simple but robust scaling method based on Lamsal et al. (2008) and McLinden et al. (2014). The equation number was not added, because only single equation is used in the MS and no needs to distinguish between multiple equations.

Line 42-57: Based on this paragraph, readers expect to see the importance of 'a priori vertical profile' for the accuracy of column data. But that is not the key point of this study right? The main finding of this study is the examination for the relationship between column and surface NO2 in terms of meteorological pattern, acquired by the clustering. Again, the background work suggested in the introduction chapter is not well connected to the key point of this study. ➔ Yes. As noted in the answer to the general comments at the top of this document, MS has been duly treated, underlining that our focus is on the exploring the discrepancies between PI and SI depending on meteorological conditions. In the original MS, our goal description in the paragraph starting at Line 59(starting with 'By contrast') is now looking relatively more emphasized.

Line 58: The meaning 'weak vertical profile correlation' is not clear. what is the profile correlation? ➔ This sentence has been now replaced by '*weak correlations between column and surface measurements*' in the revised MS.

Line 63-64: The meaning of statement "found that they originated ... from the surface layer" is not clear. How this explains the weak correlation between PC and SI NO2 during the KORUS-AQ? What is different from Wang and Christophere (2003) showing the high correlation in Alabama? →We now paraphrased this to '.. and suggested, as a possible reason, the transported non-uniform plumes originated in China, and Seoul hundreds of meters above the ground from the surface layer above the ground' (see Line 57-59) In the revised MS.

Line 85-86: 'Low-orbit' and 'geostrationary' cannot be used together. → It has been changed to *'low-orbit or geostationary'* in the revised MS.

Line 110-123: I would recommend to focus on the result of GMAP2020 campaign only. The result in GMAP2021can be a PART 2 paper in the future. → As explained earlier, we all removed this part of GMAP-2021 in the Fig. 1, as well as paragraphs described in the main text.

Line 133: What is LPS?
→ In the revised MS, full name + acronyms (in parentheses) for LPS, Met, AQM, and PA were checked and re-described in the text.

Line 113-114: Reference? → As the sentences regarding GMAP-2021 campaign information have been deleted, no relevant reference is needed any more.

Line 158-159: Cloud cover 0.6 looks a loose criteria. Is PC NO2 quality OK under the cloud cover - 0.5 (50%)? If yes, how is it justified? → Due to the lower data acquisition rate, we employed 0.6. We know that the standard for cloudiness of 0.6 is loose, but it corresponds to the first screening: however, in the subsequent DOAS fitting process, the cases with a large error have been excluded. Therefore, the effect of larger cloudiness has been further removed, and we do not expect the biases to change the major findings of the present study.

Line 162-163: I think that this data to 30 October 2021 are not part of GMAP 2021. Again, the solely usage of GMAP 2020 data looks meaningful and better to derive the obvious key message from this work. →We duly corrected and removed the sentences relevant to GMAP-2021, hereby only focusing on GMAP-2020, with the only exception with statistics of PA_4 that was still mentioned

in Table 1.

Line 179: 'NIER-GP2021-002' this format is right as the reference of ACP?
→ (Line 171) We changed to (NIER, 2018), and completed to adjust all of the references, following the acp guideline, as acp-editorial board requested.

Line 205-222: The methodology of clustering is not clear. The minimum amount of basic theoretical description is necessary. Here, the usage of XLSTAT software is the only clear part related to the conduction of clustering analysis, which does not look enough. → We shortly explained our implemented k-mean cluster analysis and agglomerative hierarchical cluster approaches, and also cited references (see Line 212-218 in the revised MS).

Line 213-214: How did authors hypothesize this? Actually, I can accept this idea, but in the manuscript, the reliable logic / scientific reason is needed to have a hypothesis. → We introduced the relevant references that appears to correlate with our interpretations. (See Line 206-212 in the revised MS)

Line 239: Why the correlation is estimated in a 'log' scale? → As indicated in the MS, we presented the results of both linear- and log-scale: a fair logarithmic relationship (R = 0.45), and a relatively lower 1:1 linear relationship (R = 0.41). We guess that log-scale would be more appropriate by direct guessing from the apparent shapes (see the Figures below). See two Figures below.
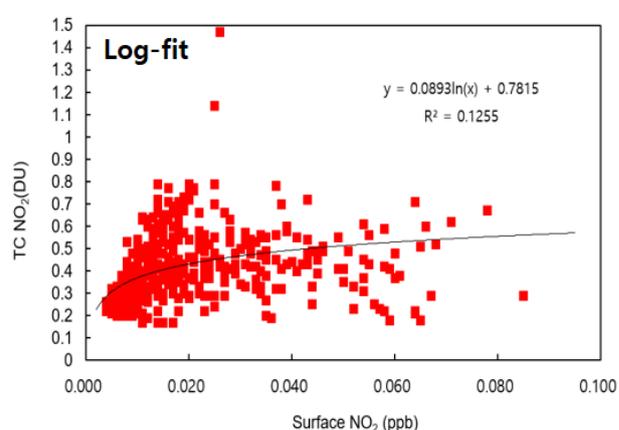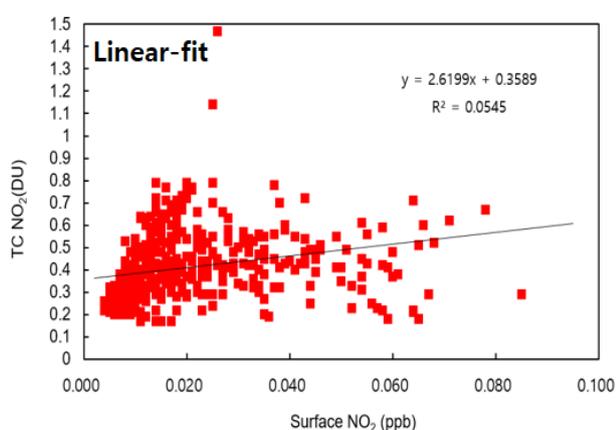
By comparatively investigating two Figures above, we concurred that Log-fitting is more appropriate, because sharp increase for lower SI-NO2 range (i.e., < ~20 ppb), and almost constant trend for higher SI-NO2 range (i.e., SI-NO2 > ~20 ppb).

Log-scale distribution (high increase for lower SI-NO2, and almost constant for higher SI-NO2) is probably because the Pandora data and SI-NO2 are both on hourly bases (not daily), and SI-NO2 will be compressed (or diluted) directly by shrinking (or developing) PBL, whereas PC-NO2, the total column amount, does not change by PBL. These effects were more pronounced at higher SI-NO2 concentrations. Of course, there should be the exceptional cases that show linear relationships, which have already been presented in Figure 3b, as an example in the current study.

Line 242-244 + Fig. 3: This different correlation looks very interesting but there is no explanation about this. Some ideas to descibe this difference should be added here. → We added the description of negatively correlated case (R<0) (Fig. 3b) in the text, such as photochemical process and/or pollutant transport, yielding nonhomogeneous NO2 profile. In the revised MS, we added 'The negative correlation on April 21 (Fig. 3b) implied that the nonhomogeneous NO2 distributions vertically were due to the photochemical process. For example, the decrease in PC-NO2 despite an increase in SI-NO2 might have occurred because NO2 is removed by photochemical loss; it can occur more severely in the upper atmosphere with high OH concentrations. Another possible reason is the occurrence of lifted layers related to pollutant transport, yielding sharp changes in vertical concentration from the surface to the upper layer: The case-specific discussion follows' (see Line 249-255)

Line 245-247: This is associated with the colume NO2 or surface NO2? It is not clear.
→ This is probably for Line 250-260. We now clarified by adding short note such as 'surface' and 'column' in the individual sentences in the revised MS (Line 256-262)

Line 266-269: How to find the PBLH using the HYSPLIT simulations? The method is unclear.
→This is probably for Line 271-274. We have now clarified that PBLH was simulated by Global Forecast System (GFS), and Lagrangian backward trajectories obtained from Hybrid Single-

Particle Lagrangian Integrated Trajectory (HYSPLIT), installed in GFS system, in the revised MS (Line 279-282)

Line 270: Which region relates to this PBLH information? Region is unclear.
→We also clarified the region for the PBLH, by adding a short note 'in Seasan, our study area,' in the revised MS. (Line 282-283)

Line 278: What is SBI?→ SBI (Sea Breeze Index) was now mentioned in the caption of Figure 4, and we now defined SBI (with equation number) also in the main text (Line 289-300)

Line 284-288: The statement is not well connected to the previous sentence. Please improve.
→We improved the sentences and clarified in the text, such as,

'. Most SBIs in group 1 ranged from 0 to 3, indicating that group 1 corresponded to the dominant local circulation (LD), whereas the SBIs in group 3 had the lowest frequencies comparing 1 and 3, which corresponded to a dominant synoptic-scale circulation (SD). Group 2 can be considered a mixture of local and synoptic-scale circulation (MD)..' (Line 296-300 in the revised MS)

Line 289-294: I do not the function of this paragraph. Why do readers think this information related to the result of this study? What can readers know better based on this paragraph?
→ This comment seems to be for Line 294-299. We agree with reviewer's opinion, and removed the whole sentences. We originally tried to understand the vertical profile, fitting to the surface levels, but it does not appear that this is directly relevant to the main point of the current study.

Line 298-310: Based on my understanding, authors addressed that the PC-NO2 and SI-NO2 shows good correlation if PBL inside is well-mixed, and this mixing condition is determined by the meterology pattern, therefore we need to consider the meteorological pattern more significantly for the analysis of relationship between column and surface NO2. Am I right? If right, authors need to put the weight more on the role of homogeneity in the PBL to the correlation between column and surface NO2. Was it found before? If yes, discussion with some previous reference is necessary. If not (i.e. this work is the first to show the importance of PBL homegeneity related to the correlation between column and surface NO2), this should be more underlined.

→ This is probably for Line 303-315. We agree and underlined the role of PBL homogeneity in evaluating PC-NO2 and SI-NO2 in more detail, in the revised MS. (see Line 249-255, Line 355-365)

Line 339-341: Reference?
→ We added Chong et al. (2018) for MAPS-Seoul, and Herman et. al. (2018) for KORUS-AQ campaign in the revised MS. (Line 349, 352)

Line 368-370: How can be the NO2 (short lifetime) transported across the Yellow Sea? It is very debatable. Please add some discussions if authors would state the possibility of NO2 long-range transport with several citations.
→This sentence is for Line 373-378. We added some relevant references and short note for this issue in the revised MS. "This transport of NOx across the region was also discussed and might be particularly high during the winter when the NOx lifetime is relatively longer (Lee et al., 2013; Stohl et al., 2002; Wenig et al., 2003)" (Line 390-392).

Fig. 1: Recommend to have the GMAP 2020 information only.
→As described earlier, we have removed the information on GMAP-2021, focusing on GMAP-2020 only.

Fig. 3b: This contrast is frequently found or one of correlation is a really irregular one? It requires more and deeper statements. →As stated above, we added the possible reasons chemically (non-linear photochemical reactions) and meteorologically (existence of lifted layer due to transport process) in the revised MS. (see Line 249-255)

Fig. 5: Left figure is for the surface NO2, but right figure is for the surface 'delta' NO2, which are different from the absolute value. Please improve the figure caption for better explanation of figures.→We now corrected the Figure, and captions. In the new caption $\triangle NO_2$ was newly defined, such as,

'(b) PC-NO2 vs. Surface $\triangle NO2$ in each meteorological condition over a 1-year period (November 12, 2020–October 30, 2021). Here Surface $\triangle NO2$ = SI-NO2–(30-day moving average) SI-NO2' in the revised MS.

Fig. 7 and 8: Figure caption should be corrected. Fig 7 is for flight 5 and 7, but Fig 8 is for flight 1 and 3, so the date and information in detail is different.

→Caption of Fig. 7 has been corrected: (a, b) FL-5 (December 6) and (d, e) FL-7 (December 6 and 9) is correct information. *We thank you for reviewer's thorough comments, and we believe that the comments raised by reviewer has been much more strengthened our MS.*

Reply