Answer to Referee # 2

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

General Comments:

The authors compare in situ measurements and column measurements of NO_2. The overall purpose of this comparison is to investigate how well surface measurements can be used for the validation of (GEMS) satellite measurements of NO_2. The authors find that the correlation between both data sets depends on meteorological conditions. They define specific classes based on different slopes and correlation coefficients. They explain the different relationships making also use of additional measurements, e.g. vertical profiles from aircraft measurements.

Their main conclusion is that ,caution is required when performing GEMS validation using either PC or SI observations alone, particularly under prevailing local wind meteorological conditions or transport processes.' While I agree that caution is required if only surface observations are available, I disagree that caution is required if tropospheric column amount measurements from independent sources are available. Overall, this is a very useful study and I recommend publication in ACP after major revisions.

→ We appreciate the reviewer's points. GMAP campaign has been actually started in2020 for multi-purpose, including GEMS validation. In the next year 2021, GMAP-2021 campaign was carried out based on more multi-perspective observation instruments from the ground to the space for GEMS validation such as MAXDOAS, Car-DOAS, GCAS, and Pandora data. However, in this MS, regarding GEMS validation, we carefully checked and removed direct connection between GEMS validation and GMAP campaign. We now believe the MS was revised in accordance with reviewer's comments below, and, by reflected the reviewer's suggestions and comments, as bellow.

Specific Comments:

Major points:

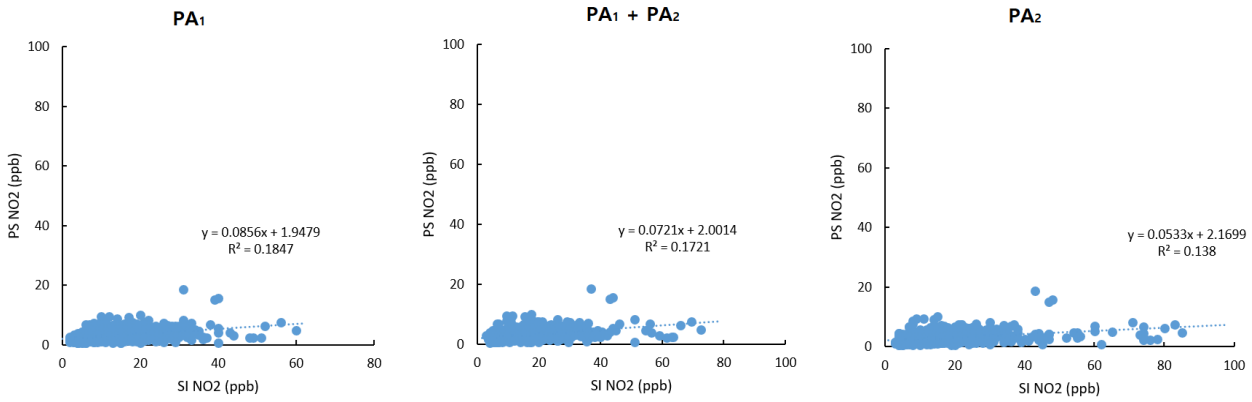
1) As mentioned above, I disagree with the authors that caution is required if tropospheric column amount measurements from independent sources are available. In contrast, such independent data sets (e.g. from MAX-DOAS observations) are a very good source for satellite observations of tropospheric NO_2 columns. I suggest to remove this statement from the abstract and from other parts of the paper. → We all agree with reviewer's points. We removed this issue In the revised MS, and highlighted the analysis of relationship between PC-NO2 and SI-NO2, rather than association directly with the GEMS satellite validation. As reviewer recommended, we have all removed the sentences relevant to GEMS validation (over the abstract and the main text as well), such as Line 25-27, 98-100, 123-124, 211-213, 391-393, and 526-528 in the original MS)

2) From Pandora, also tropospheric profiles and tropospheric VCDs are available. These quantities are much better suited for a correlation analysis than the total NO_2 columns. It is not clear to me why the authors chose the total VCDs. I suggest that the correlation analysis should be extended (or replaced) using tropospheric VCDs and surface concentrations from pandora measurements.

→We checked further on this point and discussed with co-authors, such as the differences between tropospheric-VCD vs. total-VCD, and we concurred the following points.

In Seosan, we confirmed that the tropospheric-VCD has a high correlation with the total-VCD, mainly because the stratospheric column changes little in space and time at the local scale. Therefore, we guess that it is highly likely that the results from tropospheric-VCD would be almost the same (or similar) as our total-VCD-employed results. However, regarding the data uncertainty, we found out that the quality of tropospheric VCD provided by PGN in Seosan is still low, and some tropospheric VCDs are found to be even larger than the total VCD in some cases. In this background, we thought that it is believed to be still too early to use tropospheric-VCD in our study. Furthermore, PGN does not provide profile data of NO2, and the atmospheric aerosol-loading is significant over the East Asia, and accordingly it is highly likely to cause considerable uncertainty in the NO2 profile calculation if it is not properly treated. As an example of uncertainty, ground concentration level calculated by Pandora shows, when compared to SI-NO2, the lower correlation (R^2<0.2) was found and the tendency to underestimate (slope<10%) is highly likely, as seen below. In this situation, total-VCD would be still useful to employ for our

study. However, as reviewer pointed out, it seems unreasonable to connect our findings to connect 'GEMS validation purpose', and therefore removed those through the MS, as mentioned above.



3) I suggest to extend the correlation analysis by looking how the correlation between PC or SI observations depends on the time of the day. This would be a very valuable addition. We can expect that the correlation changes with time, because also the vertical mixing and the photolysis rate changes with time.

→ We carried out further analysis to check this reviewer's point: tropospheric-VCD vs. total-VCD, and added the description and add a supplementary Figure (Figure S1, see below) in the revised MS.

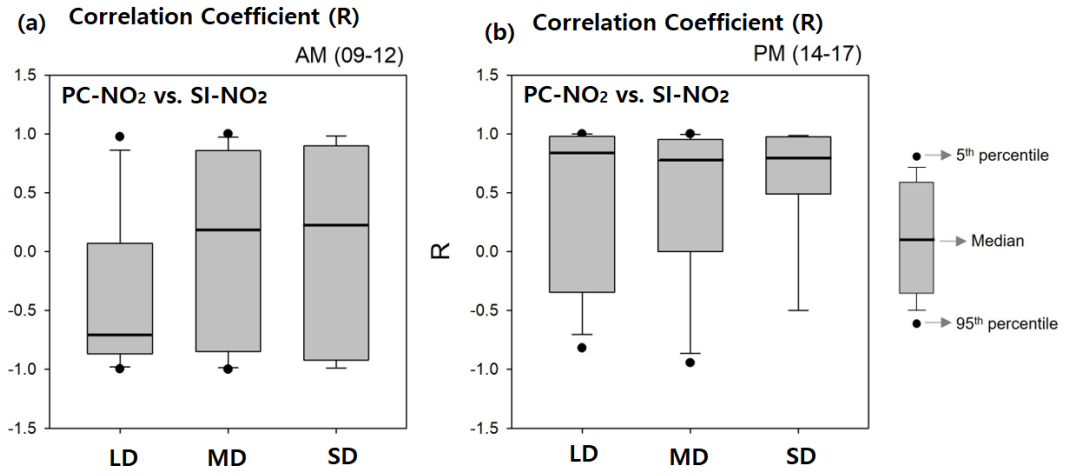


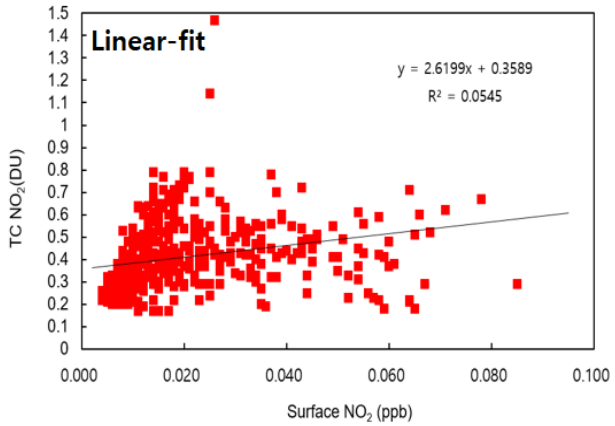
Figure S1. Correlation coefficients between PC-NO2 vs. SI-NO2 at (a)0900~1200 LST and (b)1400~1700LST, observed at Seosan during GMAP-2020 campaign.

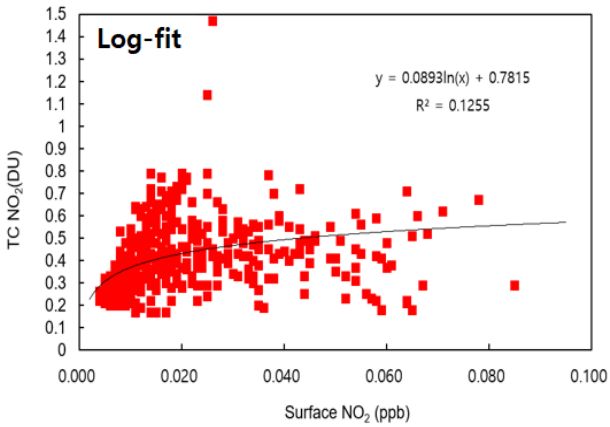
(In the main text)

"In this study, we extended the correlation analysis, and investigated the correlation between PC-NO2 and SI-NO2 measurements on an hourly basis. The results showed the correlation between PC-NO2 and SI-NO2 have a lower correlation in the morning, and a higher correlation in the afternoon (Fig. S1). The median correlation coefficients for three LD, MD, and SD meteorological conditions were -0.71, 0.18, and 0.22 in the morning (0900~1200LST), and 0.84, 0.77, and 0.79 in the afternoon (1200~1400LST), respectively. These can be interpreted from the PBL development. SI-NO2 decreases in the morning due to the rapid PBL growth, while PC-NO2 increases due to the accumulation of NO2 in the atmosphere. However, there is very little change in PBL in the afternoon, and PC-NO2 and SI-NO2 show similar changes, yielding a positive correlation each other. Our study is limited to GMAP campaign period; thus more detailed interpretations would be needed to infer more plausible causes." (see Line 355-365 in the revised MS).

4) line 241: You write: ,These hourly data exhibited a fair logarithmic relationship (R = 0.45),'

(Why) do you calculate the logartithmic relationship? Please clarify→ This is also pointed out by another reviewer. We guess that log-scale would be more appropriate by direct guessing from the apparent shapes (see the Figures below). See two Figures below.





Simply by investigating two Figures above, it appears that Log-fitting is more appropriate, because sharp increase for lower SI-NO2 range (i.e., < ~20ppb), and almost constant trend for higher SI-NO2 range (i.e., SI-NO2 > ~20ppb). 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.

Minor points:

Line 159: It is not clear how direct light measurements can be performed under partly cloud-covered skies. Please clarify.

→ Clouds increase the PC-NO2 noise and lower precision of measurement. In the case that cloud cover is small (thin or moderate), PC-NO2 can be well estimated (Herman et al., 2009, JGR). This is because, when direct light passes through a thin cloud, the sunlight reaching Pandora is constantly reduced in proportion to the wavelength; therefore, this linearity makes it possible to retrieve the PC-NO2 density. However, in the case of thick cloud, however, the proportion of scattered light becomes greater compared to direct light, and the retrieval error increases significantly. Therefore, we limited the data used in this study to that with an observed uncertainty of 0.01 DU (or less). Probably reviewers (and Journal readers also) would know the relation of clouds vs. PC-NO2-uncertainty, and thus we concluded that it is not needed to put details in the main text; only added the reference in the revised MS.

Line 254: how can you expect a constant value? The stratospheric NO_2 amount varies with season (and time of the day). Please clarify. → We then clarified the period of our estimation by adding 'during the GMAP campaign period', and specified the application period. As a reference, we estimated monthly variations of the stratospheric NO2 amount from both TROPOMI through

2020/Nov. ~ 2021/Oct. (see the Table below). The results showed that stratospheric NO2 is low in winter and high in late spring. During the GMAP-2020 (Nov. Dec. and Jan), stratospheric NO2 was found to be 0.086 DU(=[0.084+0.09+0.085]/3DU) within the analysis period, well comparable with our results. The result also showed that the average is 0.1 DU, while the standard deviation is 0.02 DU, and the sigma/mean is within about 17.4%, implying that the fluctuation range is not large. Therefore, we did not mention further or add further information about this in the revised MS.

Month	Stratospheric NO2 (DU)
11	0.084
12	0.09
1	0.085
2	0.095
3	0.111
4	0.119
5	0.137
6	0.134
7	0.132
8	0.127
9	0.118
10	0.107
Mean	0.111583
Standard Deviation	0.019374
Sigma/Mean	17.3629

Line 361: You write: ‚All observed NO2 profiles shown in Fig. 7 appeared to have generally exponential curves...'I think you cannot conclude this because there is not measurement between the surface and 500m. Please clarify.

→ This was also pointed out by another reviewer, and we deleted all this sentence in the revised MS.

We appreciate the reviewer's insightful comments that are believed to have much more strengthened our MS.