

Response to Reviewer #2

Thank you for your careful and thorough reading of this manuscript and your thoughtful comments and suggestions. Our responses follow your comments (in *Italics*).

General comments:

The manuscript by Li et al. presents an important evaluation research work of NO₂ diurnal variation using observations and modelled results from DISCOVER-AQ 2011. The research topic is important and interesting to atmospheric modelling and observation communities. The approach used is comprehensive. Some of the findings (e.g., potential spatial distribution bias in emission inventory, potential bias in ground-based remote sensing instruments) in this work are important for not just modelling groups but also observation groups. But, the presentation of this work should be improved. I would recommend publishing this work if the following concerns and comments can be addressed.

Thank you for your suggestions and comments. We have made several major revisions to the manuscript based on the suggestions by you and another reviewer.

- 1) We have updated our WRF and REAM simulations using WSM6 (WRF Single-Moment 6-class scheme) instead of WSM3, as listed in Table S2 (Line 34) in the revised supplemental table file. We also have downloaded the updated DISCOVER-AQ 2011 P-3B observations. All relevant results, including figures and tables, have been updated accordingly. The new results are almost the same as before except for some minor differences. The changes have no impact on our conclusions. The WSM3 results are now used as a sensitivity test (Lines 626 – 632) to confirm the reliability of our results and conclusions.
- 2) We have deleted the discussion on the reliability of 36-km NO_x emissions and total NO_x emission amount in the DISCOVER-AQ region but added an analysis of the 36-km NO_x emission distribution issue. So the current manuscript just focuses on the distribution issue but does not include any judgment on the total NO_x emission amount. Please see Lines 43 – 47, 125 – 126, 666, 718 – 730, 732 – 733, 753 – 789, 794, 797 – 806, 850 – 857, and 1385 – 1390 (Figure 15) in the revised main manuscript and Lines 281 – 288 (Figure S27) in the revised supplemental figure file.
- 3) We have added individual site comparisons in the supplemental figure file (Figures S19 – S23, Lines 242 – 260) to demonstrate the NO_x emission distribution issue. Please see Lines 37, 40 – 43, 668 – 678, 814, 829 – 831, and 837 – 848 in the revised main manuscript.
- 4) We have added some more detailed explanations for the Pandora issue in the late afternoon and early morning. Please see Lines 536 – 544 in the revised main manuscript and Figure S13 (Lines 185 – 190) in the revised supplemental figure file.
- 5) We have moved the evaluation of WRF meteorological fields to a new section 3.1 and added the evaluation of vertical profiles for several meteorological variables in the new Figure S6. Please see Lines 158 – 183 and 339 – 366 in the revised main manuscript and Figure S4 – S8 (Lines 108 – 146) in the revised supplemental figure file.
- 6) We have used stricter and more consistent criteria to filter out invalid satellite NO₂ TVCDs (Lines 205 – 206 and 224 – 225). The GOME-2A morning high bias is gone, as shown in Figure 10 in the revised main manuscript (Lines 1344 – 1353). Relevant changes in the text are in Lines 564 – 567, 818, and 823 – 826.

Detailed responses are as follows.

Specific comments:

L87-88. Many previous works were not properly cited. As I know, various research work has been done to convert Pandora NO₂ VCD to TVCD or surface values to study diurnal variations. The authors should update relevant knowledge on these. E.g., Kollonige et al., 2017; Spinei et al., 2014; Zhao et al., 2019. I believe some of the results in this work could be compared with previous findings and may cast some light on the research topic.

Reply:

Thank you for your suggestion. We have added three more citations using both Pandora VCD and in-situ surface observations to investigate NO₂ diurnal variations. Please see Lines 88 – 93 in the revised main manuscript. Zhao et al. (2019) had two figures showing the diurnal cycles of NO₂ surface concentrations (model, in-situ, and Pandora-derived), so we have also cited it in Lines 400 – 401. The study of Spinei et al. (2014) was not much related to tropospheric NO₂ diurnal variations. It was cited in Line 247 to show the stratospheric NO₂ VCD diurnal variations.

L151-176. These detailed discussions of the wind-filed and precipitations should not be done here, as the reader does not know anything about your trace gas simulation results/discrepancy yet. Such detailed discussions (the author used six figures in total, Figs. S2-S7) of potential causes should be included in a separate discussion section.

Reply:

Thank you for your suggestion. As mentioned above, we have moved the evaluation of WRF meteorology to section 3.1. Please see Lines 158 – 183 and 339 – 366 in the revised main manuscript.

L203 and L213. 36-km REAM profiles were used to calculate AMFs for both OMI and GOME-2A. Are these new AMFs have higher or lower (or comparable) resolution compared to the original AMFs used in the satellite data products? Please provide a brief description of how the model output has been smoothed or interpolated to OMI and GOME-2A grids.

Reply:

Thank you for your suggestion. We first regridded corresponding 36-km REAM NO₂ vertical profiles to OMI/GOME-2A pixels, then calculated AMFs by using the regridded vertical profiles. Therefore, the updated AMFs have the same resolutions as the original ones from the DOMINO/GOME-2A products. We have added a brief introduction of our retrieval method in Lines 214 – 216 in the revised main manuscript.

For the regridding approach, we would like to show the details below but not in the main manuscript, as it is not easy to explain it clearly in 1-2 sentences and not directly relevant to the topic of the study either.

We first construct a latitude-longitude matrix with a resolution of 0.01° (~ 1 km), as shown by the dash lines in Figure R1. We calculate the location of each point of the matrix in the 36-km REAM domain. For example, the red, green, and purple points in Figure R1 are corresponding to (i = 50, j = 60), (i = 50, j = 60), and (i = 51, j = 59) of the 36-km REAM domain. For any

given OMI/GOME-2A pixel (determined by the corner latitudes and longitudes), as shown by the black box in Figure R1, we can obtain the matrix points located inside the satellite pixel. We know its corresponding location in the 36-km REAM domain and then the corresponding REAM NO₂ vertical profile for each point inside the pixel. We then average the corresponding 36-km REAM NO₂ vertical profiles of all points inside the pixel, which is the a priori NO₂ vertical profile for that satellite pixel. The updated a priori NO₂ vertical profile is then used to calculate AMF and NO₂ TVCDs. This differential-like approach can be used to satellite pixels at any scale and in any shape, as long as the latitude-longitude matrix resolution is high enough compared to those satellite pixels so that the computation error is ignorable. For OMI and GOME-2A pixels with nadir-resolutions of $13 \times 24 \text{ km}^2$ and $80 \times 40 \text{ km}^2$, 0.01° is enough.

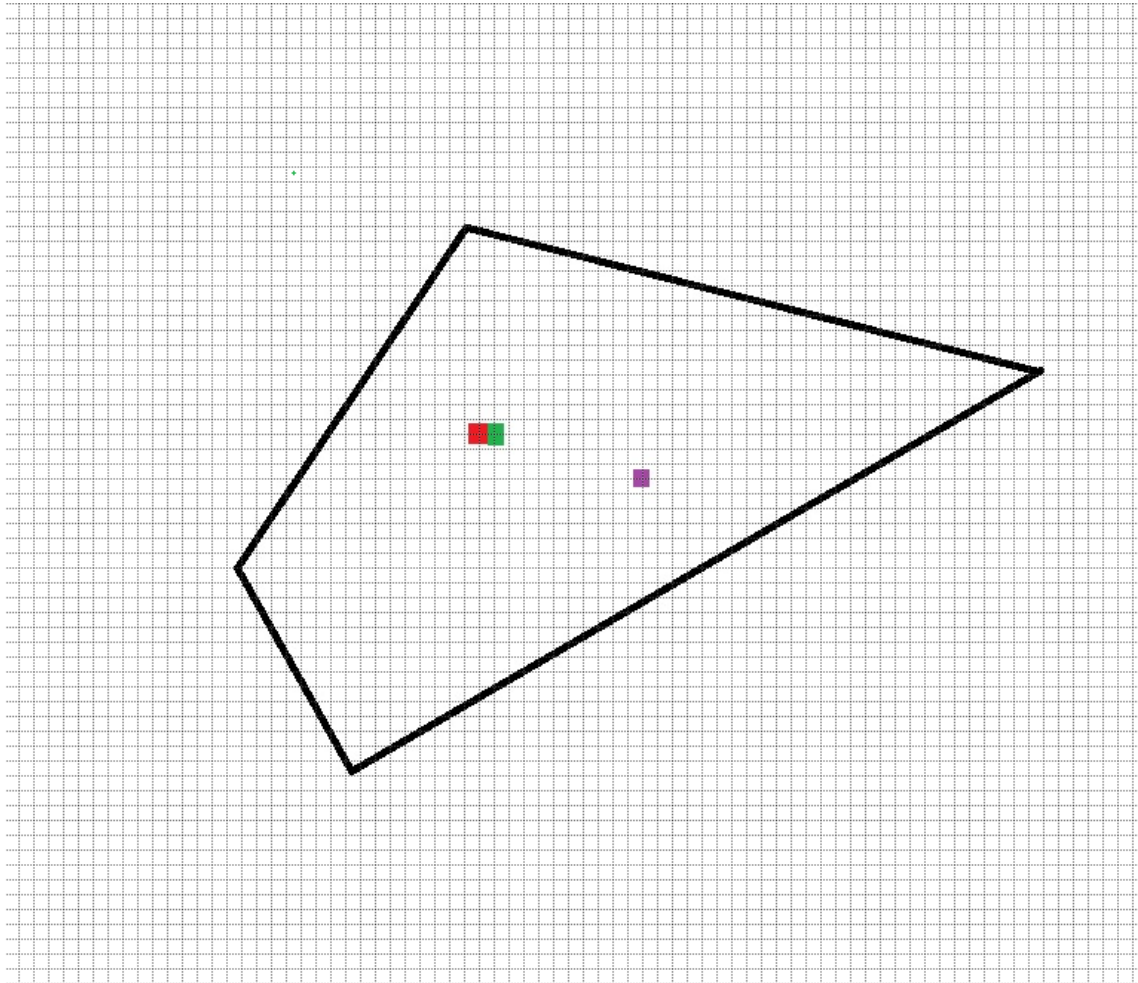


Figure R1. Schematic of the regridding approach. The dash lines denote a $0.1^\circ \times 0.1^\circ$ latitude-longitude matrix, and the black box denotes a satellite pixel. Colored points represent those $0.1^\circ \times 0.1^\circ$ points inside the satellite pixel.

L328-353. I saw at least three names for Kzz modelling, and I do have difficulty understanding which one is which. After reading this section back and forth several times, I think two Kzz modellings were used, i.e., Kzz-WRF and Kzz-modified. But, I am not sure if this Kzz-WRF is the same as Kzz-YSU. I can understand the logic of why the authors want to modify Kzz for

nighttime, but please improve the descriptions to make it easier for a reader to absorb your idea.

Reply:

Thank you for your suggestion. Yes, there are two types of k_{zz} used here: one is from the WRF simulation with the YSU scheme, and the other one is that we modify. K_{zz} -WRF is the same as k_{zz} -YSU, since our WRF simulations use the YSU scheme. We now use a consistent name — WRF-YSU — to denote the k_{zz} data simulated by WRF with the YSU scheme. Please see Lines 370, 373 – 374, 378, 382 – 383, 385, 392, 1311, and 1316 – 1317, and Figure 6 in Line 1314 for relevant modifications. We hope it is easier to understand now.

L347-351. Some justifications for the selected parameters are missing. A sensitivity test or correlation studies are needed to justify this 5 m s^{-2} . The idea of a magic number is not impressive. It is difficult to justify the selection with Figure 4, which shows even the modified results still have large discrepancy compare to observations.

Reply:

Thank you for your suggestion. As emphasized in the manuscript, the assigned value of 5 m s^{-2} is arbitrary. We did not choose a magic number. You may have noticed that we also need to select a height value and a time range for the k_{zz} adjustment. In the manuscript, we used 500 m and 18:00 – 5:00 LT. Using these values is just to simplify the modification but not to best match the observations. That's why modified PBLHs are still lower than observations in the nighttime and late afternoon.

We made many sensitivity tests when finalizing the selection of 5 m s^{-2} , 500 m, and 18:00 – 5:00 LT in the manuscript since the beginning of this study. We have shown the sensitivity test results with k_{zz} adjusted to 2 m s^{-2} and 10 m s^{-2} in Figure S10 in Lines 152 – 155 in the revised supplemental figure file and cited it in Line 396 in the revised main manuscript. Nighttime surface NO_2 and O_3 concentrations are very sensitive to k_{zz} . Using 2 m s^{-2} also makes significant changes to the simulated results.

In fact, it is almost impossible to make some simple adjustments of k_{zz} to perfectly match the vertical profile of k_{zz} and diurnal variations of PBLH. As shown in Figure R2, WRF-YSU k_{zz} shows a “C” shape. From afternoon to nighttime, the k_{zz} values change, and the “C” shape height varies. In other words, k_{zz} at different heights changes differently. We previously used a very complex equation to imitate the diurnal evolution of the k_{zz} vertical profiles and try to slow down the variation rate from afternoon to nighttime, as shown below.

when $k_{zz}(t, l) \geq 0.01 \text{ m} / \text{s}^2$,

$$k_{zz}(t + \Delta t, l) = \max\left(k_{zz}(t, l) \cdot \alpha(l)^{EF \cdot \beta(t + \Delta t)}, \text{WRF} \cdot k_{zz}(t + \Delta t, l)\right) \quad (\text{R1})$$

when $k_{zz}(t, l) < 0.01 \text{ m} / \text{s}^2$,

$$k_{zz}(t + \Delta t, l) = \max(k_{zz}(t, l), WRF.k_{zz}(t + \Delta t, l)) \quad (R2)$$

where l denotes model vertical levels less than 15 (\approx boundary layer top at 15:00 LT); t is the current time, while Δt is an updating time step ($= 0.5$ hours); α is a coefficient dependent on model levels; β is a coefficient dependent on time; EF is a coefficient related to land types, and EF is 1 for urban regions and 2 for other land types; $WRF.k_{zz}$ is the original k_{zz} from the WRF simulation. Equations (R1) and (R2) calculate k_{zz} at the next time step with the current k_{zz} . The equations are only active when $t > 15:00$ LT and $t < 5:00$ LT. The updated k_{zz} values are decreasing more slowly than the original WRF-YSU values since later afternoon and satisfy the vertical characteristics shown in Figure R2. The derived PBLH can match the observations in Figure 6 very well. The equations also consider the effect of land use/land cover. However, these equations have no physical meaning and are inappropriate to be used in the manuscript.

Anyway, 5 m s^{-2} is not a magic value but just to mitigate the nighttime vertical mixing problem. The selection of 5 m s^{-2} is not intended to and neither able to completely solve the nighttime vertical mixing bias. Not to mention the site differences. Readers are free to make their own adjustments in their studies if nighttime mixing is underestimated.

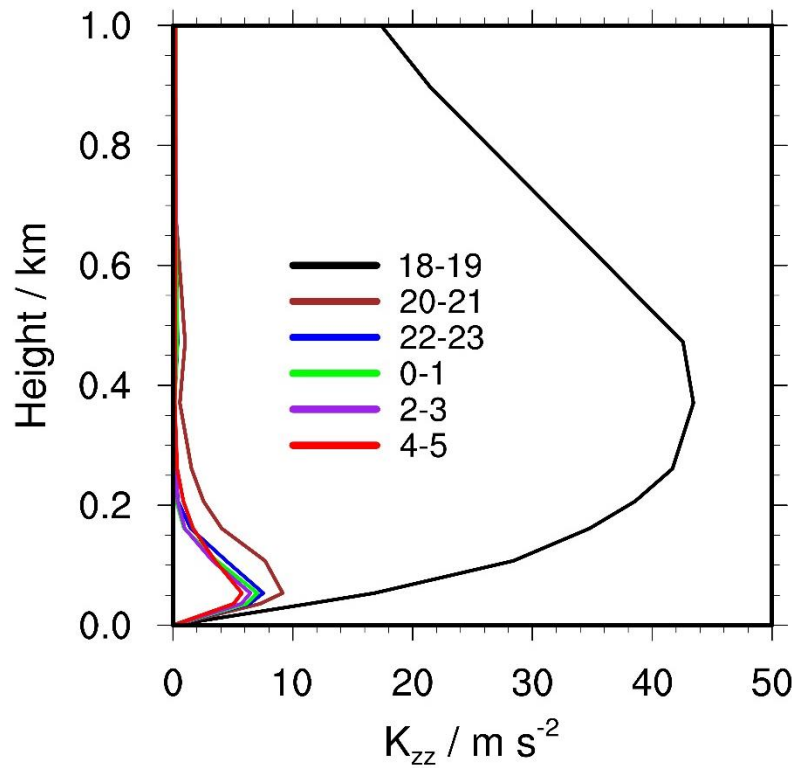


Figure R2. Vertical profiles of WRF-YSU simulated k_{zz} at different local times in July 2011 at the UMBC site.

L364-369. I am worried that the ground observations from various sites should not be studied as a single group. Different local emissions patterns should be addressed. E.g., do all 11 NO_2 sites show the same concentration peak values at 5:00-6:00 LT? Do we see any differences between rural and urban sites?

Reply:

We understand your concern about different local emissions patterns at different sites. EPA indeed shows that in some rural regions, the NO_x emissions show a unimodal diurnal pattern with a peak around noontime. However, the 11 NO_2 sites in Figure 5 are based on 36-km REAM. On the one hand, the 36-km REAM cannot resolve urban and rural well. On the other hand, all the 11 sites were not so rural, as they are all located around the Baltimore-Washington urban regions. Their emissions are still high and have similar emission diurnal variations as urban regions. Figure R3 shows the 36-km NO_x emission diurnal cycles for each of the 11 sites in Figure 5. All the sites have similar diurnal patterns and show a sharp increase of NO_x emissions in the early morning (NO_x emissions may be biased due to the distribution issue of NEI2011 at 36-km resolution).

Figure R4 shows that the monthly weekday observations at all the 11 sites peak around 6:00 LT. We do not find significant differences among these sites. Our 36-km REAM with updated k_{zz} cannot reproduce all the observed peaks at different sites mainly due to the remaining biased nighttime vertical mixing. However, the 36-km REAM can still somewhat capture the increase of NO_2 surface concentration around 6:00 LT at each site.

It is possible that NO_2 surface concentrations peak at other hours of the day but not around 5:00 – 6:00 LT. Figure 5 in the revised main manuscript shows two general conditions. 1) Nighttime vertical mixing is very weak, then NO_2 accumulates in the surface layer, possibly producing a peak around midnight, as shown by the REAM simulation with the original WRF-YSU k_{zz} . 2) The early morning increase of NO_x emissions is mitigated or removed entirely (different NO_x emission diurnal variations), leading to a much weaker or complete missed surface NO_2 morning peak as shown in Figure 5b. If one is concerned with a specific day at a particular site, anything can happen depending on the local conditions in the day, e.g., Thompson et al. (2019) showed that NO_2 surface concentration suddenly peaked around 13:00 LT in one day at an observation site in Korea (Figure 2 in their paper).

As mentioned above, we have added individual-site (the 11 sites in Table S1) comparison results in Figure S19 – S23 (Lines 242 – 260 in the revised supplemental figure file) to emphasize the NO_x emission distribution issue at both 36- and 4-km resolutions.

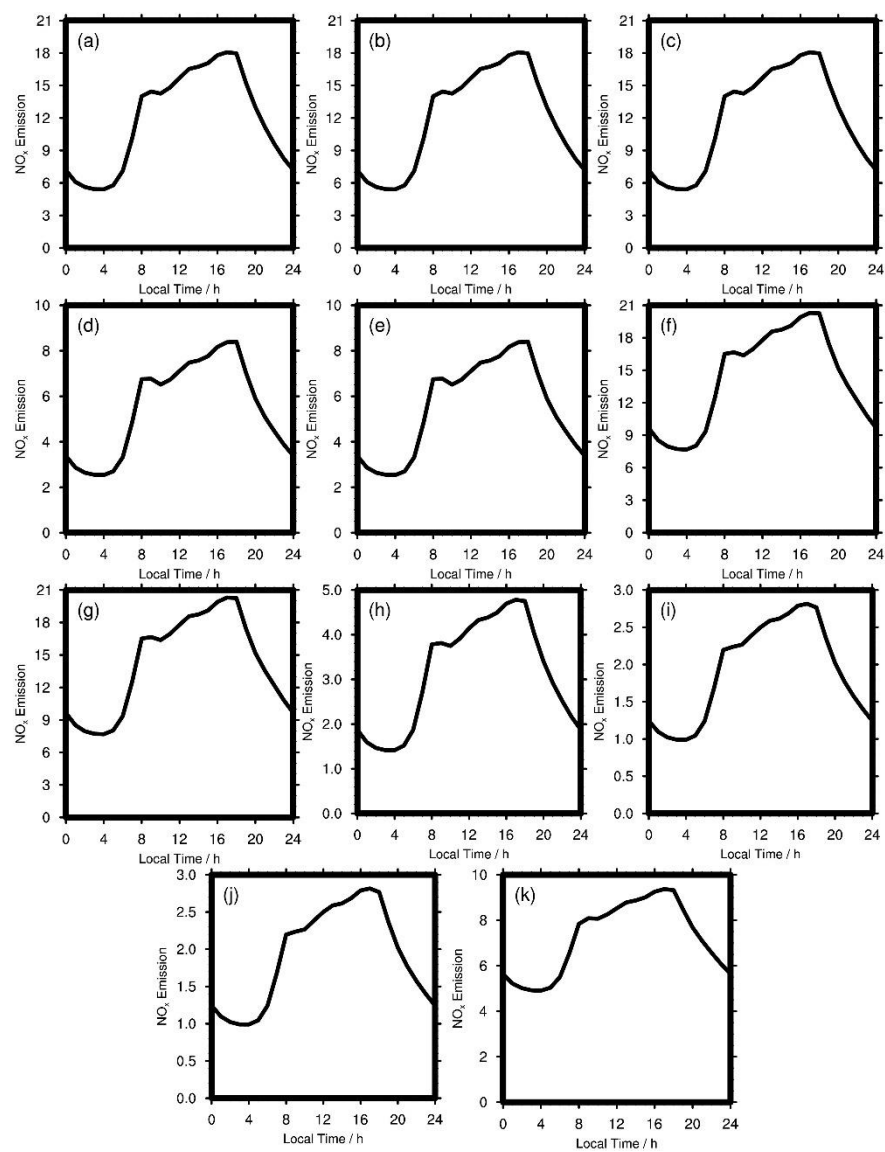


Figure R3. 36-km NO_x emission diurnal variations for the 11 sites in Figure 5 in the revised main manuscript. The unit is 10^{21} molecules $\text{km}^{-2} \text{s}^{-1}$.

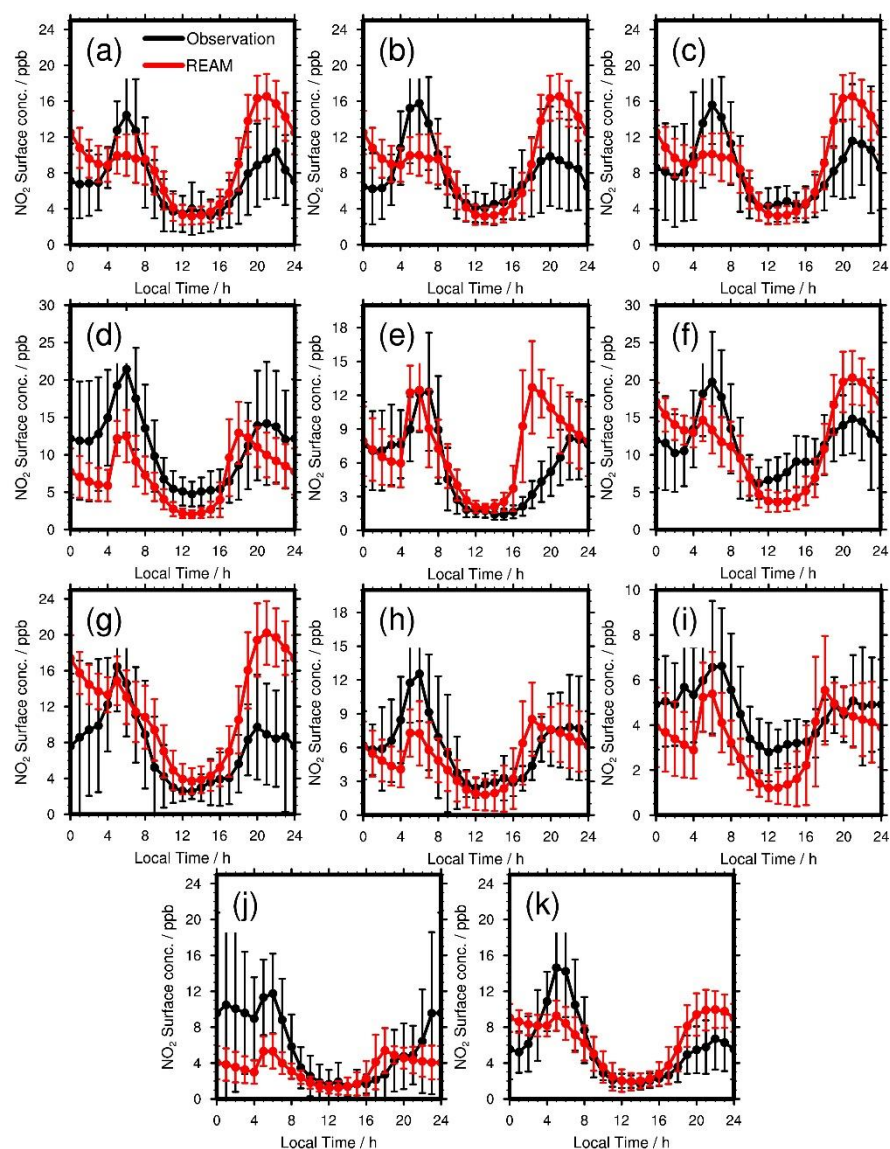


Figure R4. Diurnal variations of observed and 36-km REAM simulated NO_2 surface concentrations at different sites for weekdays in July 2011. The subplot order is corresponding to the site order in Table S1 in the revised supplemental table file.

L388-393. The general impression from Figure 5 is the REAM-4km shows a higher bias than REAM-36km compared to observations. But, this might be misleading. For example, if one looks at Figure 5b from 00:00 to 5:00 LT, the green line shows a better agreement with observations. Please provide some comments on this. The study sites should be grouped into at least two categories, e.g., rural and urban.

Reply:

Thank you for your suggestion. It should be in the daytime but not at night in Line 436 in the revised main manuscript. We have corrected it. Since nighttime vertical mixing is weak, NO_2 is primarily concentrated in lower layers, leading to large horizontal gradients. Therefore, horizontal transport plays a crucial role in nighttime NO_2 concentrations and VCDs. If nighttime

vertical diffusion is not simulated well, horizontal transport can be much different. As mentioned above, since our adjustment of nighttime k_{zz} is not perfect, we did not use nighttime comparisons in our evaluations of NO_x emissions. The nighttime vertical mixing uncertainties have little impact on daytime NO_2 surface concentrations and TVCDs (Figures 5 and R5); therefore, we mainly used the daytime simulation results and observations in our analysis and discussion. The nighttime issue is mainly discussed in section 3.2 in the revised main manuscript to describe the underestimation of nighttime vertical mixing. We have added the emphasis of the “daytime” in Lines 436, 503 – 504, and 605 in the revised main manuscript.

In addition, Figure 7b (the old Figure 5b) in the revised main manuscript is for the weekend. It is noteworthy that weekend NO_x emissions are scaled to two-thirds of weekday NO_x emissions for all sites and have the same diurnal variations, as mentioned in section 2.1 (Lines 142 – 149) in the revised main manuscript. Therefore, potential uncertainties exist in the weekend NO_x emissions. It is possible that 4-km REAM provides a more reasonable estimate of NO_x emissions at night on weekends. We have no evidence showing that the rural-urban issue contributes to the comparison results during 0:00 – 5:00 LT in Figure 7b.

As mentioned above, we have added individual-site comparison results in Figure S19 – S23 (Lines 242 – 260 in the revised supplemental figure file) to emphasize the NO_x emission distribution issue at both 36- and 4-km resolutions.

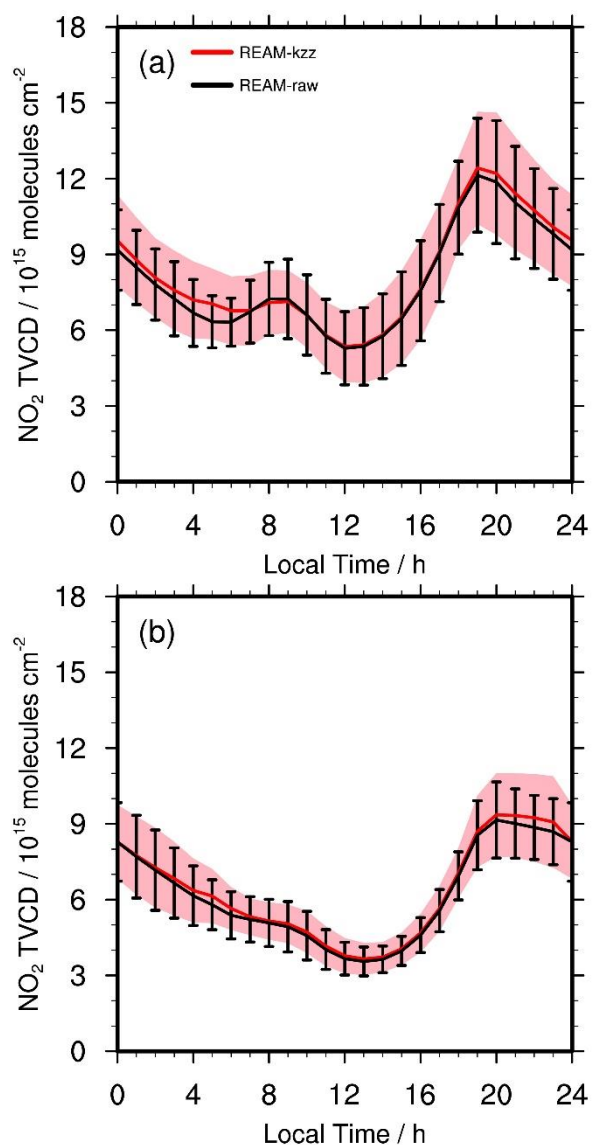


Figure R5. Comparisons of NO_2 TVCD diurnal variations between two 36-km REAM simulations on (a) weekdays and (b) weekends for July 2011. “REAM-raw” denotes the 36-km REAM simulation results with WRF-YSU simulated k_{zz} data, and “REAM-kzz” is the 36-km REAM simulation results with updated k_{zz} data.

L435-442 and Fig. S14. I guess the authors want to show the Pandora TVCD should be corrected; otherwise, the results could be biased low due to a missing surface layer. I agree with the assumption, but it needs to be studied carefully (Fig. S14 shows some indication but not good enough). Fig. S14a shows that for some sites (e.g., SERC), one can expect Pandora to miss up to 20% of NO_2 columns. However, this is not reflected by Fig. S14b at all. If this 20% difference is true, it can be verified relatively easier than other sites. Could you plot Fig. S14b for each Pandora site separately?

Reply:

Thank you for your comments. We used the averages of all 11 sites in Figure S12b (the old Figure S14b) in the revised supplemental figure file, so the differences between scaled and unscaled Pandora TVCDs are not so large (the relative difference can reach up to 6% around 6:00 LT). Figure R6 shows the difference for each Pandora site. Except for the four sites (UMCP, UMBC, SERC, and GSFC) significantly above the ground surface, all other sites have almost the same scaled and unscaled NO₂ TVCDs. For SERC, at 6:00 LT, the scaled NO₂ TVCDs is $\sim 5 \times 10^{15}$ molecules cm⁻², about 25% higher than the unscaled value ($\sim 4 \times 10^{15}$ molecules cm⁻²), consistent with Figure S12a. Since Figure S12a shows the same result as Figure R6 (the scaling ratios are the same), we don't think it is necessary to include Figure R6 in the manuscript.

The scaling may be useful for individual site comparison if the site is significantly above the ground surface. At the beginning of this study, we hoped that the scaling effect could be used to explain the Pandora's distinct diurnal variations from other datasets in the early morning and late afternoon. However, it cannot do that.

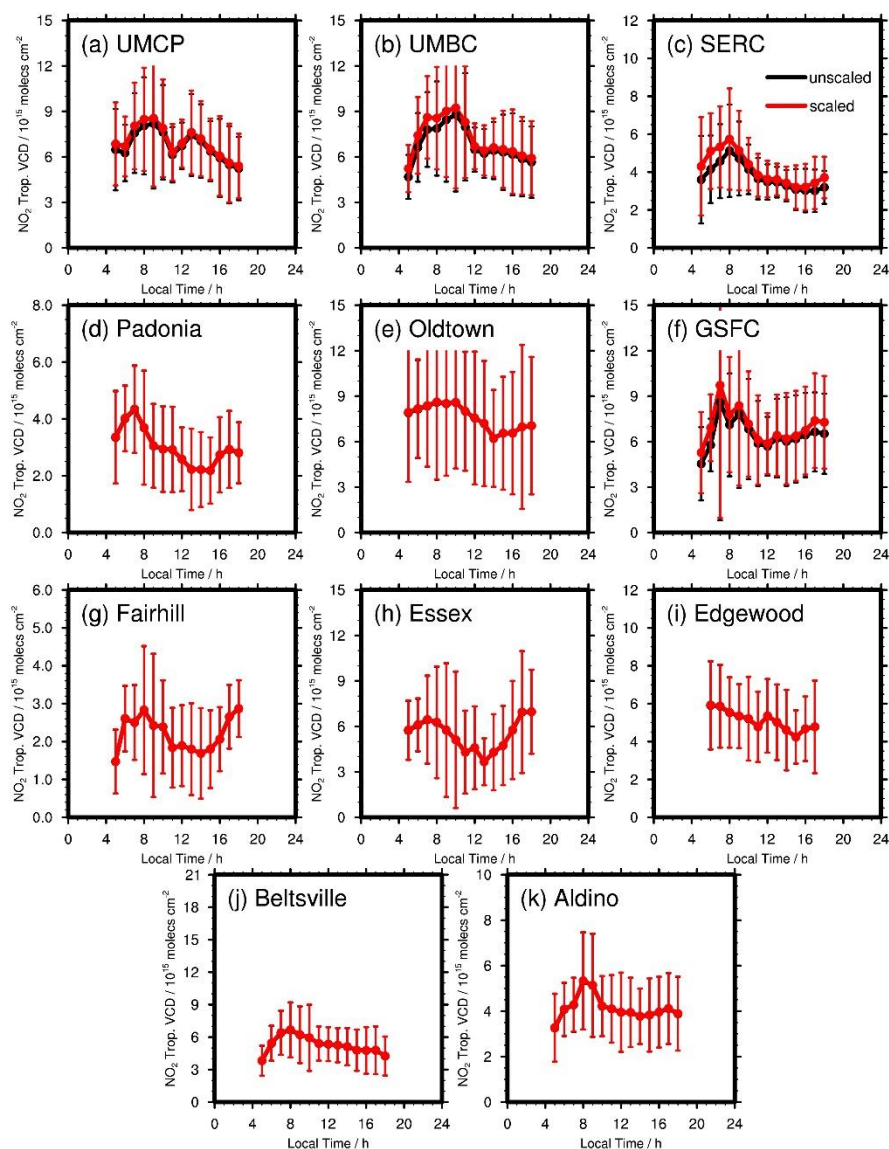


Figure R6. Same as Figure S12b in the revised supplemental figure file but for individual Pandora sites.

L469-488. The findings here are critical for the research community to understand the discrepancy between aircraft, ground-based in situ, ground-based remote sensing, and models. The synthetic aircraft TVCDs have better agreement with REAM especially for 15:00 to 17:00 LT. The agreements between REAM and aircraft profiles (Figure 6) are very nice. So, for me, it looks like Pandora TVCDs are the one that has a major low bias. But, Figure 5 also shows that the REAM has a large positive bias compared to ground-based in situ observations from 15:00 to 17:00 LT (especially for REAM-4km). Can authors conclude if Pandora TVCDs are not accurate in this period? These results may affect the claim of accuracy of Pandora NO₂ VCD is 2.7×10^{15} molecules cm⁻² in L218. Also, from Fig. S9, it is clear that the observed diurnal variations at different sites could be very different. This matched with the large error bars on the REAM modelled results in Fig. 7. But why Pandora TVCDs from 11 sites show very stable results (small error bars) in Figure 7? The current explanations are not good enough to convince me. Besides understanding the model resolutions, this could be another highlight of this research work. So, I would suggest the authors provide more investigation, explanations, or discussions.

Reply:

Thank you for your suggestions. The surface layer (1st layer of the REAM model) only contributes a small part of NO₂ TVCDs due to its shallow depth. The positive biases of NO₂ surface concentrations in REAM may be related to still underestimated vertical mixing in the afternoon (Figure 6 in the revised main manuscript). However, it is noteworthy that vertical mixing only affects the vertical distribution of NO₂ but not NO₂ TVCDs directly (vertical mixing can slightly affect NO₂ TVCDs indirectly as NO₂ lifetime is somewhat different at different heights). Therefore, the positive bias of NO₂ surface concentrations in REAM cannot provide any significant information for NO₂ TVCDs.

According to the model diagnostics, the sharp increase of NO₂ TVCDs in the late afternoon is mainly due to the sharp decrease of chemical loss (Figure 9 in the revised main manuscript). We think the model diagnostic result is reasonable. However, we cannot conclude that Pandora is inaccurate in the late afternoon. We have added more detailed explanations in Lines 536 – 544 in the revised main manuscript and Figure S13 (Lines 185 – 190) in the revised supplemental figure file. In our opinion, the most crucial point is that Pandora FOV is so small, and the instrument is located on the ground surface. Therefore, Pandora only covers a small area of air mass and can measure different air columns in the early morning, noontime, and late afternoon. Considering the significant spatial heterogeneity of NO₂, the measured NO₂ TVCDs can differ from each other significantly. In summary, Pandora measured NO₂ TVCDs are very different from those measured by satellite and simulated by models, especially in the early morning and late afternoon. Whether Pandora measurement can represent the average of a 36×36 km² column depends on the heterogeneity of NO₂ in that column. To evaluate the accuracy of Pandora, we need similar high-resolution instruments.

Figure 10 (the old Figure 7) only considers the temporal standard deviations of 21 weekdays and 10 weekend days in July 2011. We first calculated the mean hourly NO₂ TVCDs of the 11 Pandora sites. Considering the significant spatial heterogeneity of NO₂, we hoped that the average of 11 Pandora sites could represent the regional characteristics. Then, we calculated the monthly mean NO₂ TVCDs and corresponding standard deviations at each hour for weekdays

and weekends in July 2011. REAM results were processed in the same way. Therefore, discrepancies among different sites have not been considered in Figure 10 in the revised main manuscript. We have added individual site comparisons in Figure S23 in the revised supplemental figure file, showing the discrepancies among different Pandora sites.

Technical corrections:

L194. Please modify the description of estimated uncertainty. “molecules cm⁻² + 25%” does not make sense.

Reply:

Thanks. We have changed it to “an absolute component of 1.0×10^{15} molecules cm⁻² and a relative AMF component of 25%”. Please see Lines 201 – 202 in the revised main manuscript.

L217-218. The description of the precision of Pandora NO₂ VCD is not correct. In Herman et al. 2009, the 0.01 DU (or 2.7×10^{15} molecules cm⁻²) precision is for slant column (not VCD). For Pandora NO₂ VCD, the estimated precision is about 0.02 DU (e.g., Zhao et al., 2020).

Reply:

Thank you for your suggestion. Yes, according to Herman et al. (2009), the 0.01 DU precision is indeed for SCD. We have corrected it. Please see Lines 230 – 231.

L288-294. The scale ratios look consistent between ECO and C42. The one that needs extra caution is CY42 (Thermo Model 4211-Y). But, if the Thermo Model 42I-Y NO_y analyzer measurements are not used in this study at all (see L1175-1176), there is no need to include such detailed discussions (it will only confuse the reader). Or, at least, this information should be moved to supplement. I would suggest authors move other figures such as Fig. S1 to here, which should be more important (for the reader to understand the model scales/grids and locations of observations used in this study).

Reply:

Thanks. We have deleted the discussion related to the Thermo Model 42I-Y NO_y analyzer. Please see Lines 1302 – 1303 in the revised main manuscript. And we have moved the old Figures S1 and S2 to Figures 1 and 2 in the revised main manuscript.

L377. Figure 6 is used before Figure 5. Please swap the order of the figures.

Reply:

Thanks. We have deleted the references of old Figures 6 and S13 here. We have mentioned sections 3.3 and 3.4 in the revised main manuscript, so it is unnecessary to refer to the figures again. Please see Lines 422 – 423 in the revised main manuscript.

Figure 8. Please use different symbols for >400m and <400m lines. Also, the caption said there are three bins, but I did not see proper labels for the “400m – 3.63 km”. Are those >400m lines represents “400m – 3.63km” results? Please make sure the legends match with the caption.

Reply:

Thank you for your suggestion. Yes, “> 400 m” means “400 m – 3.63 km”. We have corrected it. Please see Figure 11 (Lines 1355 – 1361) in the revised main manuscript. And we have used 300 m to separate different height bins to match the newly downloaded P-3B observations, which can go down as low as about 300 m, as shown in Figure 8 in the revised main manuscript. The modification doesn’t change the results or conclusions.

Figure 10. Description of the purple circles on panels a-c is needed.

Reply:

Thanks. We have added a sentence describing the purple circles. Please see Lines 1380 – 1381 in the revised main manuscript.

Fig. S1 should be modified. The symbols for different observations jams together and very difficult to see. One should use other means to show instruments at a single site, e.g., a pie chart.

Reply:

Thank you for your suggestion. We have changed the marker patterns for different instruments so that the figure is clearer. Please see Figure 1 in Line 1277 in the revised main manuscript.

Reference

Kollonige, D. E., Thompson, A. M., Josipovic, M., Tzortziou, M., Beukes, J. P., Burger, R., Martins, D. K., Zyl, P. G. van, Vakkari, V. and Laakso, L.: OMI Satellite and Ground-Based Pandora Observations and Their Application to Surface NO₂ Estimations at Terrestrial and Marine Sites, *J. Geophys. Res.*, 123(2), 1441–1459, <https://doi.org/10.1002/2017JD026518>, 2017.

Spinei, E., Cede, A., Swartz, W. H., Herman, J. and Mount, G. H.: The use of NO₂ absorption cross section temperature sensitivity to derive NO₂ profile temperature and stratospheric–tropospheric column partitioning from visible direct-sun DOAS measurements, *Atmospheric Measurement Techniques*, 7(12), 4299–4316, <https://doi.org/10.5194/amt-7-4299-2014>, 2014.

Zhao, X., Griffin, D., Fioletov, V., McLinden, C., Davies, J., Ogyu, A., Lee, S. C., Lupu, A., Moran, M. D., Cede, A., Tiefengraber, M. and Müller, M.: Retrieval of total column and surface NO₂ from Pandora zenith-sky measurements, *Atmos. Chem. Phys.*, 19(16), 10619–10642, <https://doi.org/10.5194/acp-19-10619-2019>, 2019.

Zhao, X., Griffin, D., Fioletov, V., McLinden, C., Cede, A., Tiefengraber, M., Müller, M., Bognar, K., Strong, K., Boersma, F., Eskes, H., Davies, J., Ogyu, A. and Lee, S. C.: Assessment of the quality of TROPOMI high-spatial-resolution NO₂ data products in the Greater Toronto Area, *Atmos. Meas. Tech.*, 13(4), 2131–2159, <https://doi.org/10.5194/amt-13-2131-2020>, 2020.

Reply:

Thank you for providing the references.

References

- Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M., and Abuhassan, N.: NO₂ column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-Sun DOAS technique: Intercomparisons and application to OMI validation, *J. Geophys. Res.-Atmos.*, 114, <https://doi.org/10.1029/2009JD011848>, 2009.
- Spinei, E., Cede, A., Swartz, W. H., Herman, J., and Mount, G. H.: The use of NO₂ absorption cross section temperature sensitivity to derive NO₂ profile temperature and stratospheric–tropospheric column partitioning from visible direct-sun DOAS measurements, *Atmos. Meas. Tech.*, 7, 4299–4316, <https://doi.org/10.5194/amt-7-4299-2014>, 2014.
- Thompson, A. M., Stauffer, R. M., Boyle, T. P., Kollonige, D. E., Miyazaki, K., Tzortziou, M., Herman, J. R., Abuhassan, N., Jordan, C. E., and Lamb, B. T.: Comparison of Near - Surface NO₂ Pollution With Pandora Total Column NO₂ During the Korea - United States Ocean Color (KORUS OC) Campaign, *J. Geophys. Res.-Atmos.*, 124, 13560–13575, <https://doi.org/10.1029/2019JD030765>, 2019.
- Zhao, X., Griffin, D., Fioletov, V., McLinden, C., Davies, J., Ogyu, A., Lee, S. C., Lupu, A., Moran, M. D., Cede, A., Tiefengraber, M., and Müller, M.: Retrieval of total column and surface NO₂ from Pandora zenith-sky measurements, *Atmos. Chem. Phys.*, 19, 10619–10642, <https://doi.org/10.5194/acp-19-10619-2019>, 2019.