Authors' response to comments from Anonymous Referee #2

General comments:

This paper presents NO_2 results from the GEMS instrument for June-August 2021. As the NO_2 slant columns are biased, the authors present a correction at S5P overpass time based on the TROPOMI NO_2 SCD. The stratospheric correction is also based on the TROPOMI NO_2 product, and on the GEOS-Chem model, including its stratospheric diurnal variation. the POMINO algorithm is then applied to derive the AMF and the final NO_2 tropospheric columns. The POMINO-GEMS NO_2 columns are finally compared with the POMINO-TROPOMI product, as well as MAX-DOAS columns and NO_2 surface concentrations.

The paper is well-written and clear. I recommend publication after addressing the above major comments.

We thank the Referee #2 for taking time to review our paper and provide constructive suggestions and comments for improvement. We updated our POMINO-GEMS algorithm by replacing nested GEOS-Chem v9-02 derived stratospheric NO_2 VCDs with NASA GEOS-CF v1 derived stratospheric NO_2 VCDs, and reprocessed all retrievals. Updated validation results show great improvement in NO_2 diurnal patterns between POMINO-GEMS and ground-based MAX-DOAS measurements. We also use mobile-car MAX-DOAS measurements in the Three Rivers' Source region on the Tibetan Plateau to validate POMINO-GEMS retrievals, and good agreement is also shown in terms of NO_2 diurnal variation. Responses to these general and specific comments are provided below.

My main concern is the strong correction applied to the GEMS observations. At S5P overpass time, the GEMS NO_2 SCD are basically replaced by the TROPOMI NO_2 SCD, on a grid cell basis. To my understanding, the only true GEMS NO_2 information remaining is the diurnal variation relative to the mid-morning values. Unfortunately, the MAX-DOAS validation results are poor when it comes to diurnal variations. This is a serious limitation. This should be further discussed in the paper.

Thank you very much for your suggestion. We found that the poor correlations of NO_2 diurnal variations between POMINO-GEMS and ground-based MAX-DOAS measurements are mainly caused by poor simulation of stratospheric NO_2 from nested GEOS-Chem v9-02. Therefore, we decided to use NASA GEOS-CF v1 product to re-calculate hourly stratospheric NO_2 VCDs. Updated comparison results with ground-based MAX-DOAS measurements show much better correlations in terms of NO_2 diurnal variation, and we also proved that TROPOMI-guided correction for total NO_2 SCDs makes little difference to the POMINO-GEMS NO_2 diurnal variations. We have added more discussion about this issue in the revised manuscript.

In Line 543-577, we added:

"Figure 9 compares the diurnal variation of tropospheric NO_2 VCDs between POMINO-GEMS and MAX-DOAS at eight stations. At each site, NO_2 values are averaged in JJA 2021 at each hour for comparison, and the number of valid days for each hour is also shown. The Cape Hedo site is not included because there are few valid MAX-DOAS data points at each hour. Figure 10a-f show that at the urban and suburban sites, MAX-DOAS NO_2 (black lines) peaks in the mid-to-late morning, declines towards the minimum values at noon around 13:00 LST, and then gradually increases in the afternoon. Strong

correlation of NO_2 diurnal variation between POMINO-GEMS (red solid lines) and MAX-DOAS is found at Xuzhou (R = 0.82), Hefei (R = 0.96), Fudan University (R = 0.84), Nanhui (R = 0.79) and Xianghe (R = 0.94). At the Dianshan Lake site, POMINO-GEMS NO_2 columns increase but MAX-DOAS data decrease from 08:00 to 09:00 LST, resulting in a lower correlation coefficient (R = 0.60). At Chongming and Fukue sites, MAX-DOAS NO_2 shows a peak in the morning without evident increase in the early afternoon, but this diurnal pattern is not fully captured by POMINO-GEMS. At Fukue, POMINO-GEMS NO_2 exhibit abrupt changes at 12:00 and 13:00 LST due to few valid data.

In addition, comparison of POMINO-GEMS diurnal variation with NO_2 data from GOME-2 in the morning and OMI and TROPOMI in the early afternoon shows good agreement at Hefei, Nanhui, Dianshan Lake, Chongming and Fukue sites. The differences between POMINO-GEMS to MAX-DOAS NO_2 VCDs are comparable or smaller than those between LEO satellite and MAX-DOAS NO_2 VCDs.

As we use TROPOMI total NO_2 SCDs to correct those of GEMS, this may influence the NO_2 diurnal variation of original GEMS observations. Thus we also compare MAX-DOAS data with re-calculated POMINO-GEMS tropospheric NO_2 VCDs without correction in total SCDs (red dashed lines in Figure 9). Compared to our default POMINO-GEMS data (with correction), excluding the correction leads to lower diurnal correlation coefficients at Xuzhou, Hefei, Fudan University, Nanhui and Dianshan Lake, but higher correlation coefficients at Xianghe, Chongming and Fukue. Excluding the correction increases the NMB at three sites but decreases the NMB at five sites. We conclude that at these eight sites (in the eastern areas), no significant influences on the diurnal variation of POMINO-GEMS tropospheric NO_2 VCDs are brought in through TROPOMI-based correction for total NO_2 SCDs."

On the same idea, the authors present a comparison between POMINO-GEMS and POMINO-TROPOMI. The comparison results are obviously very good, but the study is biased. I strongly recommend to use independent satellite NO_2 products; such as OMI and GOME-2 products. The addition of OMI and GOME-2 would allow to compare with the GEMS observed diurnal variation.

Thank you for your constructive suggestion. We have added comparisons between POMINO-GEMS and OMNO2 v4 and GOME-2 GDP 4.8 tropospheric NO_2 products, and made additional comparisons in the discussion for NO_2 diurnal variations.

In Line 496-506, we added:
"Figure 7d-f and g-i show the comparison results of POMINO-GEMS tropospheric NO_2 VCDs with OMNO2 v4 on a 0.25° × 0.25 °grid and GOME-2 GDP 4.8 on a 0.5 °× 0.5 °grid averaged over JJA 2021, respectively. POMINO-GEMS NO_2 VCDs exhibit good spatial consistency with the two independent products (R = 0.87 and 0.83), although with slightly lower values than OMNO2 v4 (by 16.8%) and GOME-2 GDP 4.8 (by 1.5%). These VCD differences are expected, considering the differences in the retrieval algorithm. For example, the POMINO-GEMS algorithm implements explicit aerosol corrections in the radiative transfer calculation, while OMNO2 v4 and GOME-2 GDP 4.8 treat aerosols as "effective clouds". POMINO-GEMS accounts for the anisotropy of surface reflectance by adopting MODIS BRDF coefficients, whereas OMNO2 v4 and GOME-2 GDP 4.8 use geometry-dependent and regular LER, respectively. The horizontal resolution of a priori NO_2 profiles in POMINO-GEMS is 25 km (and interpolated to 2.5 km), 1° × 1.25 °in OMNO2 v4 and 1.875 °× 1.875 °in GOME-2 GDP 4.8."

In Line 556-559, we added:

"In addition, comparison of POMINO-GEMS diurnal variation with NO_2 data from GOME-2 in the morning and OMI and TROPOMI in the early afternoon shows good agreement at Hefei, Nanhui, Dianshan Lake, Chongming and Fukue sites. The differences between POMINO-GEMS to MAX-DOAS NO_2 VCDs are comparable or smaller than those between LEO satellite and MAX-DOAS NO_2 VCDs"

In Line 663-666, we added:
"Meanwhile, surface NO_2 concentrations derived from LEO satellite observations also agree well with those of POMINO-GEMS, except that POMINO-GEMS derived surface NO_2 concentrations are higher than those of GOME-2 GDP 4.8 by about 40% – 60%."

Since this is the first study about GEMS NO_2 measurements, the paper should provide a section where the GEMS operational VCDs are compared to the presented product and provide some conclusions on the regions and periods where the GEMS NO_2 tropospheric VCD are performing good or bad.

Thank you for your suggestion. We agree that comparison between POMINO-GEMS and GEMS operational NO_2 product is necessary. Unfortunately, we found that tropospheric NO_2 VCDs in GEMS v1 operational product in summer are unavailable (no valid data), so we couldn't perform the comparison. As soon as the reprocessing of GEMS v2.0 operational product is finished, we will compare the updated GEMS operational tropospheric NO_2 VCDs with POMINO-GEMS retrievals.

In the diurnal variation plot (figure 9 and figure 11), the uncorrected GEMS NO_2 VCD should also be plotted. (uncorrected GEMS NO_2 VCD = uncorrected GEMS NO_2 SCD – NO_2 stratospheric columns)/POMINO GEMS AMFs.

Thank you for your suggestion. We have added the diurnal variation of uncorrected POMINO-GEMS tropospheric NO_2 VCDs in Figure 9 and Figure 10. The comparison results of corrected and uncorrected POMINO-GEMS derived surface NO_2 concentrations against MEE data are very similar, so we listed the statistics in Table S4 of the Supplement Information (SI).

In Line 569-577, we added:
"As we use TROPOMI total NO_2 SCDs to correct those of GEMS, this may influence the NO_2 diurnal variation of original GEMS observations. Thus we also compare MAX-DOAS data with re-calculated POMINO-GEMS tropospheric NO_2 VCDs without correction in total SCDs (red dashed lines in Figure 9). Compared to our default POMINO-GEMS data (with correction), excluding the correction leads to lower diurnal correlation coefficients at Xuzhou, Hefei, Fudan University, Nanhui and Dianshan Lake, but higher correlation coefficients at Xianghe, Chongming and Fukue. Excluding the correction increases the NMB at three sites but decreases the NMB at five sites. We conclude that at these eight sites (in the eastern areas), no significant influences on the diurnal variation of POMINO-GEMS tropospheric NO_2 VCDs are brought in through TROPOMI-based correction for total NO_2 SCDs."

In Line 586-588, we added:
"In contrast, POMINO-GEMS without total SCD correction exhibits much poorer correlation with mobile-car MAX-DOAS data, due to the erroneous increase in the afternoon."

In line 639-640, we added:
"Note that the consistency between POMINO-GEMS and MEE data does not depend on the total SCD correction (Table S4)."

Specific comments:

Abstract

Line 33: I suggest to remove the very first sentence, that sounds a bit obvious and is already in the introduction: Nitrogen dioxide (NO_2) is a major air pollutant.

Done.

Line 35: LEO NO_2 retrievals are not limited only by insufficient temporal sampling, but also by retrieval uncertainties and spatial resolution. The two limitations exist also for GEMS.

Revised.

Line 37: at an unprecedented hourly resolution during the daytime.

Revised.

Line 41: "We then derive tropospheric NO_2 air mass factors (AMFs) with explicit corrections for the anisotropy of surface reflectance and aerosol optical effects, through pixel-by-pixel radiative transfer calculations." The authors do not present the impact of those two corrections in the rest of the paper. It should be either be presented in the manuscript (see my AMF comments later) or removed from the abstract.

Thank you for your suggestion. In this study, we didn't perform sensitivity tests to discuss the impacts of surface reflectance and aerosol optical effects, but we compare these ancillary parameters when comparing POMINO-GEMS with other satellite products. Therefore, we decided to keep the sentence in the abstract.

In line 500-506, we added:
"These VCD differences are expected, considering the differences in the retrieval algorithm. For example, the POMINO-GEMS algorithm implements explicit aerosol corrections in the radiative transfer calculation, while OMNO2 v4 and GOME-2 GDP 4.8 treat aerosols as "effective clouds". POMINO-GEMS accounts for the anisotropy of surface reflectance by adopting MODIS BRDF coefficients, whereas OMNO2 v4 and GOME-2 GDP 4.8 uses geometry-dependent and regular LER, respectively. The horizontal resolution of a priori NO_2 profiles in POMINO-GEMS is 25 km (and interpolated to 2.5 km), 1° × 1.25 °in OMNO2 v4 and 1.875° × 1.875 °in GOME-2 GDP 4.8."

Line 44: The term "reveals" is overused, since the NO_2 hotspot signals are well known from LEO observations.

Revised.

Line 45: As intended by the presented method, POMINO-GEMS NO_2 VCDs agree well with POMINO-TROPOMI v1.2.2 product. Please indicate in the abstract that the remaining differences are coming from AMF differences.

Revised.

Introduction

Line 66: the provided references are for NO_2 datasets rather than LEO mission themselves. Please add more appropriate references for GOME, OMI, GOME-2, TROPOMI.

Done.

Line 90: Validation results have shown the overall capability of the official GEMS NO_2 algorithm. I'm not sure this is true. You should provide reference to support this affirmation.

Done.

Method and data

Line 128: Please explain briefly what is meant by "continuum reflectances".

Done.

Line 140: Please explain briefly what is meant by "area-weighted oversampling technique".

Done.

Please provide basic information on the slant columns retrieval settings for GEMS and TROPOMI operational products: wavelength interval, cross-sections, reference spectrum.

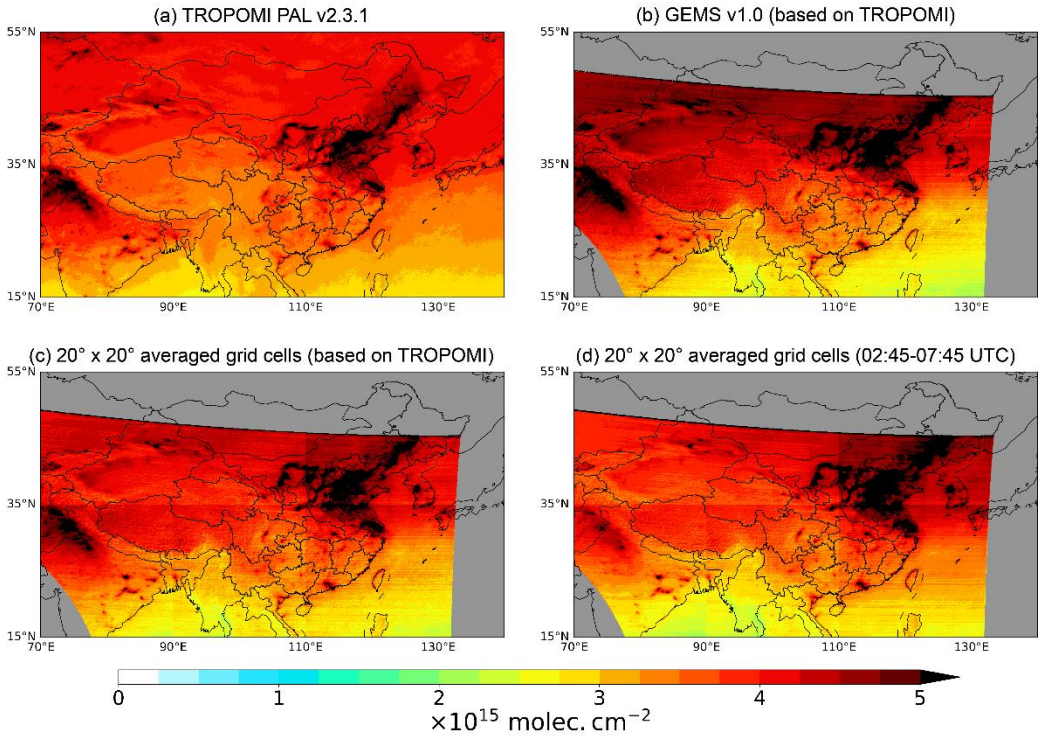
Done.

Total NO_2 SCDs

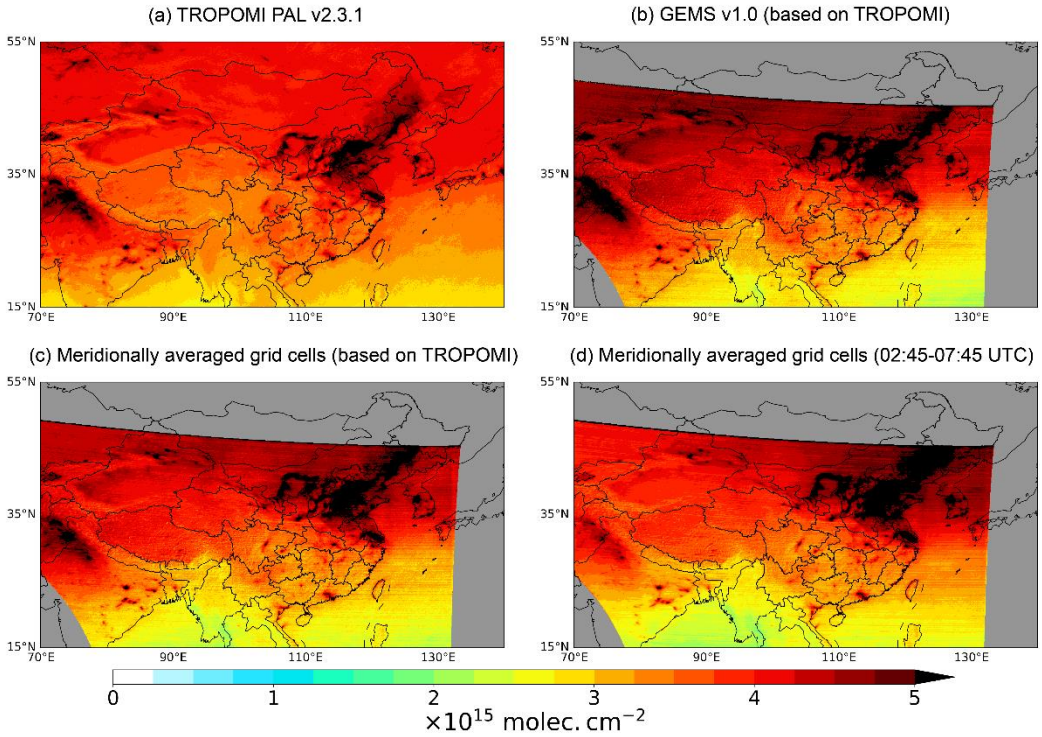
The correction based on the TROPOMI SCDs is somehow radical, since it is calculated for every grid cell. Have you tested more softer corrections, for example based on much larger grid cells, or based on meridionally averaged grids?

Thank you for your comment. We have tested three different softer corrections, and the results are shown below, respectively.

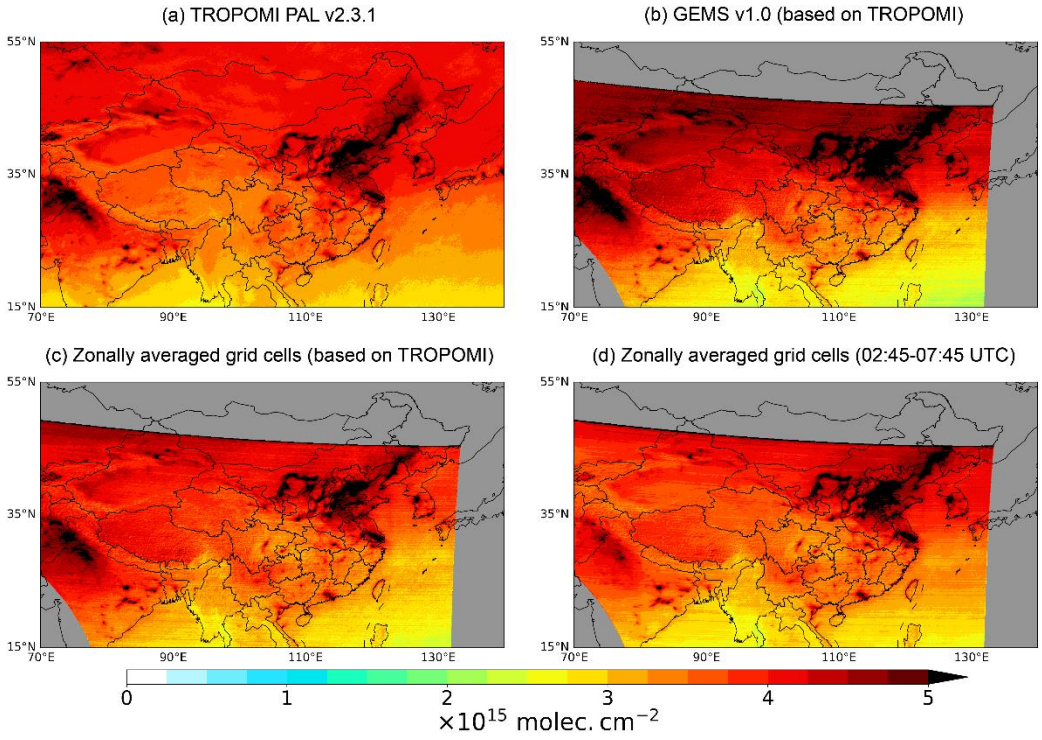
Correction based on 20°×20 °averaged grid cells:



Correction based on meridionally averaged grid cells:



Correction based on zonal averaged grid cells:



These correction methods can reduce the high bias over northern and northwestern GEMS FOV to various extents, but are not capable to remove stripes. Therefore, we think the correction method applied in our algorithm is effective enough to address those systematic issues in official GEMS product.

In Line 252-258, we added:
"Our correction method is done for each grid cell. We tested other correction methods by applying the same correction value to grid cells within a 20° × 20 ° domain, at the same latitude, or at the same longitude. These alternative methods can reduce the high bias over the northern and northwestern GEMS FOV to various extents, but cannot remove the stripes (not shown). We also note that our simple correction is a temporary solution before the aforementioned systematic problems in the official GEMS SCD retrieval are solved by improving spectral fitting. In Sections 3.3 and 3.4, we compare the diurnal variations of tropospheric NO_2 VCDs based on corrected and uncorrected GEMS SCDs."

More examples of figure 2b and d could be shown for other GEMS hours (maybe in the supplement).

Done.

Line 199: Please comment on the diurnal variation of the GEMS systematic problems. For example, is the high bias over northern and northwestern part of GEMS FOV constant during the day or does it increase?

Thank you for your suggestion. Comparisons of two products at different hours are shown in Figure S1 of the SI. The stripes remain significant at all hours, which is expected because this problem has nothing to do with the observation time. However, since GEMS observations are spatiotemporally matched with

those of TROPOMI, there is no direct comparison over the northwestern GEMS FOV from mid-morning to noon, so the diurnal variation of systematic high bias of GEMS total NO_2 SCDs cannot be clearly depicted and hence discussed yet.

AMFs

A figure presenting the POMINO GEMS amfs should be added, as well as a comparison with the POMINO TROPOMI AMFs.

Done.

Estimation of surface NO_2 concentrations

Please specify if the Rgc GEOS-Chem simulated ration is time dependent or constant. In other words, is there a diurnal variation of the model introduced with this correction? If yes, what is the observed GEMS diurnal variations if you use a constant ratio?

Thank you for your suggestion. The GEOS-Chem simulated column-to-surface ratio is time independent, so there is a diurnal variation of the model introduced with this correction. We have added the discussion of the GEMS NO_2 diurnal variations using a daily ratio.

In Line 641-647, we added:
"To quantify the influences of the diurnal variation of hourly column-to-surface ratio from GEOS-Chem simulations, we compare the MEE measurements with POMINO-GEMS derived surface NO_2 concentrations using daily column-to-surface ratio (Figure S15). As expected, POMINO-GEMS derived NO_2 concentrations show a similar diurnal variation as the tropospheric NO_2 VCDs do, with two peaks in the mid-morning and afternoon, and a minimum at noon. The temporal correlation coefficient with MEE is only about 0.23. Thus it is more reasonable to use hourly ratio for comparison with MEE measurements, as done in our study."

Line 277: Please explain briefly what is the grubbs statistical test and provide a reference.

Done.

Results and discussion

As the NO_2 total and stratospheric SCDs are almost the same by definition of the presented "fusion" technique between GEMS and TROPOMI, I suggest to skip section 3.2 and to replace it by a comparison of AMFs from POMINO GEMS and TROPOMI.

Thank you for your suggestion. Even though the fusion method leads to very similar NO_2 total SCDs and stratospheric VCDs, there are still slight differences in tropospheric NO_2 SCDs which is caused by different geometries between GEMS and TROPOMI. Therefore, we have added more detailed discussion of the reasons for the differences between POMINO-GEMS and POMINO-TROPOMI v1.2.2

tropospheric NO_2 VCDs in Section 3 of the SI.

Figure 8: the regression line values are exactly the same between plots a and b. this seems strange, please check.

Thank you for your comment. The updated regression results are shown in the revised manuscript.

Figure 9: please use a fixed scale, or at least only two different scales for high and background NO_2 levels.

Done.

Figure 11: I suggest to detail the comparison with MEE diurnal variations for different groups of sites (urban, rural, northeast, southwest China). This could provide more information on the regions where the GEMS diurnal variation is valid or not.

Thank you for your constructive suggestion. We have added detailed comparison with different groups of MEE sites in the revised manuscript.

In Line 389-395, we added:
"The spatial distribution of all MEE sites in the GEMS FOV is shown in Figure S8a, and that of MEE sites over urban, suburban and rural regions are shown in Figure S8b–d, respectively. The classification of sites is based on Tencent user location data with a horizontal resolution of 0.05° × 0.05 °for every 0.5 second from 31 August to 30 September 2021 (Figure S8e), adopted from previous work (Kong et al., 2022). Here, urban MEE sites are defined as where the mean location request times is larger than 50 times per second, suburban sites refer to 5-50 times per second, and rural sites refer to less than 5 times per second. The number of sites for urban, suburban and rural sites are 808, 554 and 71, respectively."

In Line 661-669, we added:
"Figure 12b-d show the comparison of NO_2 diurnal variations for different groups of MEE sites. The diurnal variations of POMINO-GEMS derived surface NO_2 concentrations show similar characteristics over urban, suburban and rural regions, and all correlate well with those of MEE data. Meanwhile, surface NO_2 concentrations derived from LEO satellite observations also agree well with those of POMINO-GEMS, except that POMINO-GEMS derived surface NO_2 concentrations are higher than those of GOME-2 GDP 4.8 by about 40% – 60%. We conclude that validation with extensive MEE measurements presents promising performance of POMINO-GEMS retrievals, especially the great agreement of POMINO-GEMS NO_2 diurnal variation with MEE data over urban, suburban and rural regions."

Since the uncertainties on the measured diurnal variations appear to be large, I suggest to applied to the MAX-DOAS measurements a similar "column to surface column transformation" as for the satellite columns, and to compare directly MEE and MAX-DOAS diurnal variations.

Thank you for your suggestion. We have added discussion about this comparison.

In Line 648-652, we added:
"To further test the reliability of our VCD-to-surface-concentration conversion method (Eq. (9)), we apply the same method to MAX-DOAS NO_2 VCDs and compare the resulting surface NO_2 concentrations with MEE data. As shown in Figure S16, the diurnal variation of MAX-DOAS derived surface NO_2 concentrations correlates well with that of MEE measurements ($R = 0.96$), in support of our conversion method."

Error estimates

10% error on the GEMS NO_2 SCD (or we cloud say on the TROPOMI NO_2 SCDs) seems to be underestimated. Furthermore, the diurnal variations of the error on the GEMS fits is not taken into account.

Thank you for your comment. We have re-written Section 3.5 to discuss the error estimates in a more detailed way. Although we are not able to assess the diurnal variations of the error on the GEMS fit alone, we have added a quantitative discussion about the diurnal variation of spatiotemporal correlation coefficients and NMBs of POMINO-GEMS to ground-based MAX-DOAS and MEE measurements in Section 3.3 and 3.4. We will do the detailed error analysis in the future.

Conclusions

The observed added-value of GEMS should be discussed in a more balanced way in the conclusions, as well as the current limitations.

Thank you for your comment. We have added the discussion of current limitations in this study.

In Line 756-768, we added:
"However, there are still several limitations in our study. To address the systematic overestimation and stripes problems in the original GEMS data, we correct GEMS total NO_2 SCDs by using TROPOMI data as a temporary solution. For example, we implement a simple geometric correction to combine GEMS and TROPOMI total NO_2 SCDs, but their differences in scattering geometry are only partly accounted for. Thus this correction works well in most regions but may introduce uncertainties up to 30% in the northwestern GEMS FOV. Currently, the Environmental Satellite Center of South Korea is updating the NO_2 SCD data to v2.0. We will update our POMINO-GEMS algorithm accordingly, once the updated official NO_2 product becomes available to provide necessary inputs for our research product. In addition, in the conversion from NO_2 VCDs to surface concentrations, we use a constant correction factor of 2 to account for the strong NO_2 vertical gradient near the surface. This simple treatment does not account for the diurnal variation of the correction factor, and thus may introduce errors in the derived surface NO_2 concentrations."

References:

Kong, H., Lin, J., Zhang, Y., Li, C., Xu, C., Shen, L., Liu, X., Yang, K., Su, H., and Xu, W.: Unexpected high NOX emissions from lakes on Tibetan Plateau under rapid warming, 10.21203/rs.3.rs-1980236/v1,

2022.