RC2: 'Comment on egusphere-2024-617', Anonymous Referee #2

Legend: Referee comments in **blue**, author comments in **black**

The manuscript by Lange et al. presents a very thorough evaluation of the performance of the operational GEMS tropospheric NO2 product and the scientific NO2 product from the University of Bremen over Seoul. The relevant aspects of the retrieval are evaluated in an exhaustive manner: absolute magnitude, seasonality, weekend effect, and diurnal cycle. Also some plausible interpretation of the measurements is provided in terms of emissions, transport, and atmospheric chemistry, which strengthens the study. Overall, I fully support publication of this work, and only make a few remarks and suggestion for corrections below. I agree with the other reviewer that the amount of material presented is quite overwhelming and some condensing would benefit the readability of the paper.

We would like to thank the reviewer for their helpful comments. We hope that we have adequately answered all questions and that our explanations are satisfactory.

General comment about larger changes we have made during the review process:

Upgrade to the new pandora version for pandora Seoul and changes in quality filtering for all pandora data, results in slight changes in several figures involving pandora data and improved data availability for some seasons and sites. For all pandora except the pandora Seoul, the most recent data version (rnvh3p1-8) was available when writing the manuscript. For pandora Seoul, we used the available data product at that time, which was rnvh1p1-7. In the meantime the new version is available for pandora Seoul. Since the column retrieval was improved and changes in the columns are expected, we decided to update to the new version. Additionally, we adapted the quality filter from filtering low quality and unusable data to filtering only unusable data and introduced instead an additional wrms (Normalized rms of fitting residuals weighted with independent uncertainty) filter. This results in a somewhat higher data availability for some seasons. Since ground-based data are only used when quality filtered satellite observations are available, this acts as a further indirect filter. Overall, values in the comparisons have changed only slightly, and the conclusions drawn from the figures remain the same.

To improve readability we moved some plots to the appendix and condensed the discussion about diurnal variability in section 5 (Diurnal variability of GEMS and ground-based tropospheric NO2 VCDs) into the seasonality subsection.

We added a comparison of subversions of the GEMS IUP-UB product using different stratospheric VCD products in the discussion in section 5.4.

General remarks

* It may be useful to discuss the differences between stratospheric NO2 in the operational, IUP-UB and TROPOMI products in more detail. Especially since stratospheric NO2 is argued to be one of the reasons for the overestimation in the operational GEMS product. Is there a clear reason why the method from Bucsela et al. (2013) would result in too low stratospheric columns?

Thank you for your comment. To investigate the influence of the different stratospheric VCD products, we created subversions of the GEMS IUP-UB product using different stratospheric column products. Figure 1 shows scatter plots of coincident satellite and ground-based

tropospheric NO2 VCD observations for (a) the original GEMS IUP-UB product using the STREAM-based stratospheric VCDs, (b) the GEMS IUP-UB using the TM5 stratospheric VCDs, (c) the GEMS IUP-UB using the GEMS L2 stratospheric VCDs, and (d) the original GEMS L2 product. Replacing the STREAM-based stratospheric VCD with the TM5 data increases the bias from 3% (-22% - 38%) to -20% (-41% - 5%) and changes the offset from +1.6e15 to -3.9e14 molec cm-2. This illustrates that the TM5 model stratospheric VCDs are too large, resulting in too low and even negative tropospheric NO2 VCDs. Using the GEMS L2 stratospheric VCDs for the GEMS IUP-UB product increases the bias and the offset, illustrating that the GEMS L2 stratospheric VCD product is too low. This results in an overestimation of the GEMS IUP-UB tropospheric NO2 VCD compared to the station data. The correlation stays constant for both subversions as there is little correlation between the stratospheric NO2 columns and the tropospheric NO2 variations at the stations. The higher scatter seen in the operational GEMS L2 product is caused by the surface reflectivity as shown in Sect. 5.4.

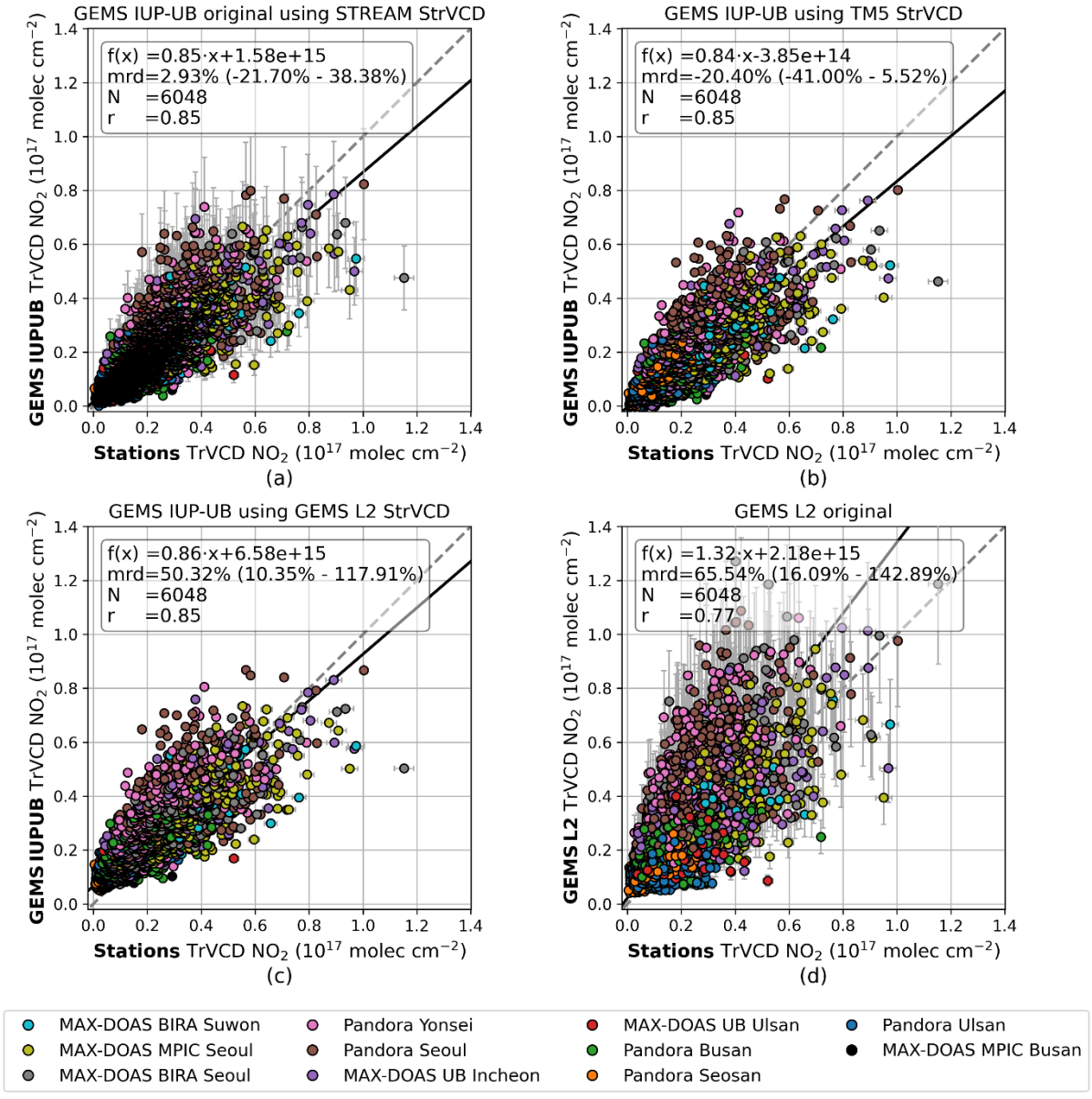

Figure 1. Scatter plots of satellite vs. co-located ground-based NO_2 tropospheric VCDs; for (a) the original GEMS IUP-UB product using the STREAM-based stratospheric VCDs, (b) the GEMS IUP-UB using the TM5 stratospheric VCDs, (c) the GEMS IUP-UB using the GEMS L2 stratospheric VCDs, and (d) the original GEMS L2 product.

Unfortunately, it is not yet clear why the Bucsela et al. (2013) based GEMS L2 stratospheric VCD is too low, one possible reason might be the chosen threshold value to find tropospheric contamination.

Have the stratospheric NO2 columns been validated?

The stratospheric NO2 columns haven't been validated in this study. Some preliminary evaluation of the GEMS L2 v2 stratospheric product was done within the PEGASOS ESA project. The TROPOMI stratospheric NO2 column was validated by Verhoelst et al. (2021) using zenith-sky DOAS measurements during twilight, showing a slight negative median difference for the stratospheric column data of -2% in summer and -15% in winter. Unfortunately, it is beyond the scope of this paper to go further into details regarding the validation of stratospheric GEMS data.

* How does stratospheric NO2 change throughout the day in the two GEMS products studied?

Thank you for this question. We investigated this in more detail and created plots showing in dark blue the stratospheric NO2 VCD of the GEMS L2 product based on the method from Bucsela et al. (2013), in light blue the stratospheric NO2 VCD based on the STREAM algorithm, and in green the TM5 model stratospheric columns, used in the TROPOMI product which is shown in black. As expected, the TM5 and TROPOMI stratospheric NO2 VCDs agree well. The GEMS IUP-UB STREAM-based stratospheric column shows a very similar diurnal evolution as the TM5 data but is slightly lower. The GEMS L2 product shows a similar but reduced variability over the day and is lower by a factor of around 2.5 when compared to the TM5 and GEMS IUP-UB stratospheric columns. We added plots and discussion in Sect. 5.4 of the manuscript.

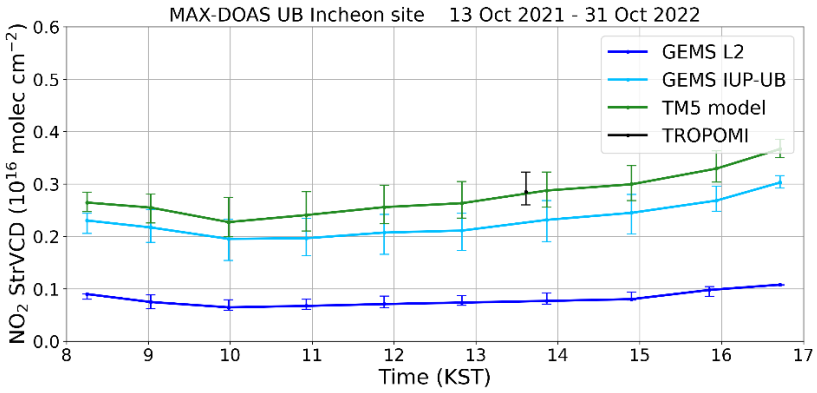

Figure 2. Diurnal variability of median stratospheric NO_2 VCDs for the GEMS L2 product based on the method from Bucsela et al. (2013) in dark blue, the GEMS IUP-UB STREAM-based product in light blue, and in green the TM5 model stratospheric VCDS, used in the TROPOMI product which is shown in black.

* A clear message what the authors think is the main reason for better validation results around noon than in the morning or late afternoon would be appropriate.

Thank you for your comment. We mention the larger biases in the morning and late afternoon now in section 5.1, and added the following:

"These differences for observations at larger SZA can be explained by a lower sensitivity of GEMS and more uncertain AMFs for these scenes, which is amplified for larger aerosol loads and low boundary layer heights in combination with a lack of knowledge of the tropospheric

aerosol in the AMF calculation for the GEMS IUP-UB product. This is further discussed in Sect. 5.4."

and point to the discussion in section 5.4, where this question is also discussed.

"Another already mentioned aspect, which possibly contributes to the differences, especially at larger SZA, is the lack of knowledge of tropospheric aerosol in the calculation of the AMF for the GEMS IUP-UB product. However, the L2 product considers aerosol parameters from GEMS observations in the AMF determination and should correct for their influence. The expected improvement is not reflected in the comparisons.
Due to less sensitivity at higher SZA (and VAA), AMFs are expected to be more uncertain for these scenes. This uncertainty is further enhanced for larger aerosol loads and with low boundary layer heights in the morning and evening."

We hope that our text is now clearer.

Minor issues

L179-181: does the GEMS IUP-UB product have a similar quality assurance flagging system as TROPOMI?

Yes, the GEMS IUP-UB and TROPOMI quality flagging system are similar but the GEMS IUP-UB has not yet a full error propagation. We have added this in the text.

L291: typo 'sight' --> slight

Changed.

L445-446: this sentence was a bit difficult to follow. Please consider rephrasing.

This sentence was rephrased. We hope it is better to follow now.

L512: stratospheric NO2 columns are usually on the order of 10^15 molec. cm-2

Yes, thank you for the comment, we changed this to 10^15.