

This manuscript presents an uncertainty assessment of TROPOMI tropospheric nitrogen dioxide over Europe using Pandora and MAX-DOAS measurements. Careful attention is given to discussing various sources of uncertainty in the retrieval process. The product is evaluated against several datasets using different a priori assumptions. Overall, the manuscript is clearly written, provides comprehensive detail, and includes insightful figures that effectively support the discussion. I recommend publication after addressing the following minor revisions:

1. Introduction: The introduction is well written and provides good context. However, in line 26, I recommend revising the statement to emphasize NO<sub>2</sub> as a key indicator for emissions and air quality, including its role in ozone and SOA formation. The current wording may unintentionally suggest that NO<sub>2</sub> observations are primarily used for broader ecosystems studies rather than their central importance in air quality applications.
2. Line 305: During the representativeness error estimation, the description of the displacement of the 4 × 4 grid is unclear. It is not evident whether the ground monitor is assumed to move within the same fixed grid box or whether multiple grid box combinations are used while keeping the ground monitor location fixed. Please clarify this description to avoid ambiguity. Additionally, the rationale for comparing the results with an OMI-sized pixel is not clear and should be further justified.
3. Line 350: Please clarify whether the spatial collocation accounts for the relative location of the ground instrument within the satellite pixel. Specifically, is any weighting applied based on the instrument's position within the pixel, or are contributions from neighboring pixels considered when the instrument is located near pixel edges or corners? This is particularly important for NO<sub>2</sub>, where strong spatial heterogeneity can significantly influence the comparisons. Additionally, regarding the 1-hour collocation window, did you assess the potential influence of temporal mismatch? Ground-based monitors record data over every few seconds, while, given the TROPOMI pixel size (~6.5 × 5.5 km) and typical wind speeds (~5 m/s), an air parcel would take approximately 15–20 minutes to cross the pixel. This temporal factor may affect the representativeness of the comparison.
4. In Figure 4, the color scale is labeled as 'NO<sub>2</sub> total column,' which appears to be incorrect. Additionally, the figure caption states that it represents 'Observation minus forecast for the geometric vertical columns (slant column divided by the geometrical AMF).' This procedure should be clearly explained at the beginning of Section 3.1 to avoid confusion. It would also be helpful to include a statement clarifying that when the CTM simulation of the stratospheric NO<sub>2</sub> column achieves close agreement with the TROPOMI slant columns over clean oceanic regions in assimilation, the difference between the TROPOMI total SCDs and the assimilated stratospheric SCDs is effectively

treated as the tropospheric SCD within the TROPOMI field of view. This clarification will help readers better understand the methodology and Section 4.1.

5. Line 380, please refer to it as stratospheric NO<sub>2</sub>. Are there any observational error estimates or studies that show the latitudinal variation of this column? It may be worth citing those efforts.
6. Lines 385–390, please provide lower and upper limits of OmF differences in percentage. At high altitudes, the stratospheric column is larger and the absolute differences could be highest, but it would be useful to show the percentage differences in high-latitude regions during winter and early spring also large compared to other unpolluted regions.
7. In section 4.1 I hardly find any additional outcomes then already knows however some seasonal studies are there but the tropopause also shows strong seasonality, no explanation regarding those factors are given. Is it possible to quantify those errors and stratospheric absolute values with any limb being observations. I also suggest examining the background Pandora Direct Sun (DS) total NO<sub>2</sub> measurements, which typically represent the stratospheric NO<sub>2</sub> contribution. These data could be used to validate the TROPOMI stratospheric NO<sub>2</sub> columns and to assess the associated uncertainties in the TROPOMI observations.
8. In Section 4.1, there appear to be few novel outcomes beyond what is already known, although some seasonal analyses are presented. The influence of tropopause seasonality, which can be strong, is not discussed. It would be helpful to quantify the resulting errors and stratospheric absolute values, potentially using observations. I also suggest examining the background Pandora Direct Sun (DS) total NO<sub>2</sub> measurements, which typically represent the stratospheric NO<sub>2</sub> contribution. These data could be used to validate the TROPOMI stratospheric NO<sub>2</sub> columns and to better assess the associated uncertainties in the TROPOMI observations.
9. Line 417: I would describe the field as more spatially homogeneous rather than seasonal, also the scale is dominated by the highly polluted northeast areas.
10. Table 4: Are the smaller representation errors in winter compared to summer due to the longer lifetime of NO<sub>2</sub> during winter and the resulting weaker spatial gradients (more homogeneous) within the TROPOMI pixel size?
11. Figure 6, in x-axis, please either say [uniteless] or remove [ ].

12. Line 500: The longer lifetime of NO<sub>2</sub> in winter is generally expected to produce more spatial homogeneity in longer-term datasets, rather than pronounced inhomogeneities within typical TROPOMI satellite pixels. However, this does not appear to explain the winter high bias observed in MAX-DOAS MMF. As shown in Figure 5, the Pandora Direct Sun (DS) measurements are often viewing away from regions with high NO<sub>2</sub> columns at the TROPOMI overpass time. I recommend adding a column in Table 1 listing the Pandora Sky Scan (SS) azimuth and MAX-DOAS viewing azimuth for each instrument, as this may help better interpret the results. Additionally, including similar figures for Bremen and Thessaloniki in the Supplement, showing MAX-DOAS viewing azimuth together with the seasonal mean DS viewing direction based on solar azimuth, would be helpful. Also the horizontal footprint of the MAX-DOAS and DS is different.
13. Line 561, are there any nearby water pixel influences in TROPOMI for Athens and Thessaloniki, where fitted uncertainties are more than twice the expected values.
14. Figure 16: I recommend adding a figure showing the absolute error as a function of the NO<sub>2</sub> column (from ground-based instruments). This would help demonstrate whether the higher-column underestimation in TROPOMI operational can be explained by the use of the high resolution CAMS a priori profiles in the retrieval.