Reply to "RC2: 'Comment on essd-2024-616', Anonymous Referee #2, 14 Apr 2025", Citation: https://doi.org/10.5194/essd-2024-616-RC2

Glissenaar et al.

This manuscript presents a global monthly (Level 3) data set of tropospheric nitrogen dioxide, specifically as a climate data record for the ESA Climate Change Initiative project. This product is accompanied by uncertainties in the Level 3 data set. Careful attention is taken to discuss how random uncertainties cancel when averaging over both space and time. The product is evaluated against several datasets.

Overall, the manuscript is written clearly with fairly comprehensive detail and insightful figures that support the text. I thought the derivation and discussion of uncertainties as the observations are averaged over space and then time was extremely valuable and comprehensive. While important enough to accompany this dataset anyway, the topic is also very useful to the community more widely. Quantifying how the average relative uncertainty in valid L2 pixels compares with the uncertainty in the L3 monthly product is very enlightening. I almost wished there were more examples in Table 3 and discussed in Lines 389-393, but I understand that a user could simply consult the data itself.

I thought the product evaluation was robust and comprehensive. Sources of discrepancies were discussed.

We'd like to thank the reviewer for their thoughtful observations and recognition of the relevance of this study.

The first anonymous reviewer comment points to the work by Rijsdijk et al. (2024), which may not get enough attention in the motivation of this manuscript (throughout the introduction, for example), and I agree with this assessment. If I understand correctly, the main contribution of this manuscript from an algorithmic perspective (compared to Rijsdijk et al., 2024) is the addition of temporal averaging in the Level 3 product here. Perhaps the authors could spend a bit more time really clarifying exactly which other developments (if applicable) are introduced here.

We agree with this assessment. We have introduced a paragraph about Rijsdijk et al.'s work in the introduction. We have also included another reference to Rijsdijk et al. (2025) in the overview of methods in Line 130. We have also included a note in

Line 224 (the start of the temporal averaging section) that this is new in the present study.

I also had a question about the potential cloud selection bias that could be introduced in the averages, *in addition to* the spatial representativeness uncertainty. I did not see that addressed here in this text. I see two points that are worth considering:

(1) Using this Level 3 dataset as a "clear sky average": for example, comparing with *coincidently sampled cloud-free model data*, the importance of a potential bias in your monthly average from selecting relatively cloud-free conditions seems to be avoided. The two quantities (cloud free satellite average and cloud-free model average) are directly comparable and the uncertainty estimate reported in the Level 3 product stands as appropriate for this application.

Yes, this is correct. We have now included a note in the dataset description (section 4.1) clarifying this.

(2) Other users may expect to consider the uncertainty in this Level 3 product as an uncertainty in the *true monthly mean*. In this case, it seems as though the reported uncertainty doesn't actually account for a potential bias due to selection of cloud-free conditions. Cloudiness could adjust photolysis rates and chemical lifetimes, or it could be correlated with atmospheric transport conditions that may bring cleaner or more polluted conditions to a particular location. The potential "bias" in a monthly average due to selecting cloud-free conditions could conceivably be positive *or* negative depending on location. Reviewer #1 alludes to this in their first comment on Section 3.3.1: "If I understood correctly, superobservations, which are probably clear-sky observations, have a higher weight; do these have a tendency to lower/higher NO2 concentrations and create a bias?". I suspect the answer could be either, and depends on local chemistry and atmospheric transport conditions relative to the location of pollutant hotspots (e.g., Geddes et al., 2012: https://doi.org/10.1016/j.rse.2012.05.008).

We have added a discussion on the difference between the clear-sky and true monthly average in the dataset description (section 4.1), including what the assessed uncertainty represents (not the uncertainty to the true average, but indeed to the clear-sky average) and how overcast conditions could affect the true monthly mean. We studied this issue before in Boersma et al. (2016), who compared the true NO2 monthly mean to the clear-sky NO2 monthly mean and found that the former is biased high by up to 25% over polluted regions.

While not exactly within the scope of this study (which focuses on propagating errors appropriately across spatial and temporal averaging), I think this point deserves some brief discussion/clarification, and especially advice to users when considering the data as a monthly average (and not when just comparing it to other clear-sky sampled data sets).

We agree, and adding the discussion and reference to the Boersma et al. (2016)-study as outlined above will help to that effect.

Finally, I wondered about the cadence of updates for this particular product. The data in this manuscript is for May 2018 to December 2021. Can users expect this Level 3 algorithm and data product to be regularly updated, and if so, at what frequency? Will the record soon be available for more recent TROPOMI years? What are the plans for long-term processing?

The L2 v2.3.1 (consistent with OMI L2) which is used in this dataset is only available until December 2021, and will not be extended. Therefore we cannot extend this specific L3 dataset.

We are planning to process TROPOMI 2.4/2.5 (consistent algorithm version) for May 2018 – Nov 2023 into a L3 dataset using the same methods as discussed in this paper.

TROPOMI L2 will be reprocessed (2.9) in early 2026. This could give an opportunity to have a L3 dataset over a long consistent period.