

We thank the reviewer for their comments. Responses to comments are shown below in red, while quotations from the revised manuscript are indicated in blue.

**Editor:**

I also suggest including satellite product names or versions. For example, for line 624 (ATC version), there are several different versions of NO<sub>2</sub> data (in evolution) from KNMI and the readers may want to specify. Thank you very much.

We have added the following sentence (Line 622):

All TROPOMI retrievals used here, unless otherwise indicated, are the reprocessed version 1 products.

**Reviewer 1:**

I read the replies and revised paper. I have no major issues. Maybe better to refer to Spivakovsky (2020) instead of (1990).

We have replaced the Spivakovsky, 1990 reference with a Spivakovsky, 2000 reference, the most recent, relevant Spivakovsky et al paper we know of.

Spivakovsky, C. M., Logan, J. A., Montzka, S. A., Balkanski, Y. J., Foreman-Fowler, M., Jones, D. B. A., Horowitz, L. W., Fusco, A. C., Brenninkmeijer, C. A. M., Prather, M. J., Wofsy, S. C., and McElroy, M. B.: Three-dimensional climatological distribution of tropospheric OH: Update and evaluation, *Journal of Geophysical Research: Atmospheres*, 105, 8931-8980, 10.1029/1999jd901006, 2000.

**Reviewer 2:**

I thank the authors for their effort. The revised results, especially the improved agreement with the Atom observations, are encouraging. Nevertheless, I still have a few concerns, as described below. If the authors address them, this paper can be published.

"The 1.0° × 1.0° and monthly resolutions, in combination with the long data record, are sufficient to understand regional trends in TCOH"

Unfortunately, this is still unclear from the present study. The potential added value of using finer-scale input data needs to be discussed more carefully. In addition, since this study does not fully utilize individual satellite observations, the implications for future observations obtained may be limited or biased. This also needs to be discussed more carefully.

We have reworded the above quoted sentence as follows (Line 248):

Nevertheless, the 1.0° × 1.0° and monthly resolutions, in combination with the long data record, provide new constraints on regional trends in TCOH and some aspects of TCOH temporal and spatial variability.

We also have added the following sentence (Line 242):

The study domain partially mitigates limitations of the 1.0° × 1.0° resolution, as spatial heterogeneity of the relevant species is generally much lower over the remote tropical oceans than over land.

As we described in the previous response and in the current paper, HCHO and NO<sub>2</sub> retrievals over the remote oceans are not of sufficient quality for higher spatial or temporal resolution than that presented here. We acknowledge that over land, where S/N of these retrievals is higher, we would likely be able to produce a higher resolution TCOH product. We discuss that here (Line 724):

A higher resolution TCOH product over land would also likely be feasible, because of the increased signal to noise of the NO<sub>2</sub> and HCHO retrievals.

"We have trained a model that omits HCHO as an input variable, finding little difference from the satellite constrained OH product that includes all inputs. "

" Both these examples suggest that NO<sub>2</sub> values are more important to the final TCOH than HCHO, whereas the feature importance has HCHO as the most important variable. "

These statements appear to be contradictory. Clarification is needed.

We have found in this and other XGBoost-related work that the feature importance is not necessarily the best metric to understand the relative effect changes in inputs can have on the model output. Whereas the feature importance indicates the relative importance of the inputs for accuracy, that does not necessarily imply that the model's response to changes in that variable will have the largest impact on the magnitude of the output. That is the distinction that we were trying to indicate in our previous response. Further, we often find that, when omitting the input with the greatest feature importance from an XGBoost model, in this case HCHO, the accuracy/bias/etc. is frequently unchanged as the model is often able to use information from the other inputs to compensate for the omission. We do not make any changes to the text in response to this comment because these quotations are from the previous response to reviewers. In the paper, our only mention of feature importance is (Line 584):

"This uncertainty analysis is in general agreement with the model feature importance (Supplementary Fig. 11), a measure of the relative importance of GBRT model inputs, where HCHO and NO<sub>2</sub> consistently have the largest values of the satellite inputs."

We feel that inclusion of this discussion in the manuscript would create an unnecessary complication.