We would like to thank the reviewer 2 for their helpful comments that improved this manuscript. Below in *italics* please, find our replies to the reviewer's comments. We have revised our manuscript by adding two new figures (and related discussion) comparing the two versions over source regions and highly polluted cities, and addressing all other comments raised by the reviewer.

==========

The authors present a new version of NASA's standard OMI NO2 algorithm, which includes several improvements: better surface reflectance treatment, new cloud product based on the updated surface reflectance, and several other improvements. The effects on AMFs are shown for global maps for specific days. The effects on VCDs are shown for specific days and long-term average. The new product has been evaluated with ground-based and airborne observations. The manuscript is generally well written and the new product version fits well within the scope of the AMT. I recommend publication of the manuscript after considering the comments below:

(1) The analysis of the AMFs focus on some daily global maps. The largest difference between V3.1 and V4.0 should be actually noticeable at the regional and local scale as also discussed in the manuscript. I suggest to add some example figures that show the improvement at that scale, for example, along polluted coastlines, in the presence of snow, or in mountainous terrain. These figures would also be important to demonstrate the new version is actually superior to the previous version.

*We agree that additional figures could help show the extent of changes between V3.1 and V4.0. Our global and regional maps already show changes over coastlines, snow, and mountains. Following the suggestions from both reviewers, we decided to include additional examples from source regions as well as five highly populated cities that can represent different geographic terrain (e.g., polluted coastal city). We have expanded Section 2.5 discussing the impact of the changes as follows:*

*"Figure 11 shows some examples of how changes in the algorithm from V3.1 to V4.0 affect monthly domain average tropospheric $NO_2$ columns over areas affected by various $NO_x$ sources. In contrast to minor changes over the pristine Pacific Ocean, month-to-month changes over source regions vary considerably. The differences in tropospheric $NO_2$ columns between V4.0 and V3.1 range from -11 to 15% over Beijing, China and from 0 to 29% over the Ruhr area in Germany, suggesting variations in relative difference among cities and industrial areas. The changes over a major biomass burning area of Democratic Republic of Congo, Angola, and Zambia range 13-56% during the biomass burning season of May through August, but are <5% in other months. Differences between the two versions are small over areas influenced by lightning $NO_x$ emissions. In Figure 12, we examine monthly variation of tropospheric $NO_2$ columns from the two versions over five highly populated and polluted cities that vary in terrain types ranging from coastal (e.g., Shanghai, Tokyo) to mountainous (e.g., Mexico City). $NO_2$ columns in V4.0 are generally higher than V3.1 by 0-30%, but the difference can occasionally reach up to 50% in some months. Changes of that order of magnitude in highly polluted areas have implications for estimation of $NO_x$ emissions and trends using these data."*

(2) The evaluation of the new product is quite short and could be extended with some additional analysis. In particular, it is currently difficult to judge if the new version significantly improves the product, because the authors do not evaluate both V3.1 and V4.0 for all data. It would also be

helpful to have table with correlation coefficient, bias and other parameters to give an overview over these numbers currently spread throughout the manuscript.
*This is indeed a good suggestion. Some of the validation results shown here for V4.0 are extension of our previous study with V3.1 discussed in detail in Choi et al (2020), which we have cited in several places of this manuscript. Therefore, we chose not to include them for V3.1. We have added a summary table (Table 2) for the validation results as suggested by the reviewer.*

L55ff: The row anomaly is only mentioned in Section 2.4 but I would consider already briefly mention it at the beginning of Section 2 because the impact on data availability is unfortunately quite severe.
*We have added the following statement in Section 2 "OMI's full daily coverage has been affected by data loss due to an anomaly presumably caused by material on the spacecraft outside the instrument that results in reduced coverage to about half of its original swath as discussed in Section 2.4."*

L99: MODIS surface reflectance has also been used in the HKOMI product (Kuhlmann et al. 2015, https://doi.org/10.5194/acp-15-5627-2015).
*Thanks. The new reference is added on Page 4, line 115.*

L435: Please mention somewhere that "coastal areas" refers to the high NO2 values labelled with "ocean" in Figure 6.
*This is clarified as suggested. We now state "cities and highly polluted coastal areas".*

L441ff: The term "lower troposphere" is somewhat confusing here, because it should be the additional layers between the new and old cloud pressure and not the full lower troposphere.
*We agree; the statement now reads as: "Higher values of OCP in OMCDO2N will include additional portions of scattering weights between the OMCDO2N- and OMCLDO2-based OCPs, especially in the lower troposphere, thereby reducing the tropospheric AMF."*

L444f: ". . .in the calculation of tropospheric AMF." -> ". . .in the calculation of tropospheric AMF increasing the AMF."
*The statement is modified as suggested "On the other hand, the higher CRF values lead to an increased contribution of the cloudy AMF in the calculation of tropospheric AMF, thereby increasing its value."*

L459ff: What is the reason for the increase of AMFs over ocean in Fig. 5c?
*The apparent change in tropospheric AMFs over ocean in Figure 5c due partly to changes in terrain pressure and partly to color bar issue. The changes are small at about 1%. The relevant statement is modified on Page 17, Line 477.*

L528: Please add a sentence that explains what kind of improvement is expected when improved NO2 profiles become available.
*The statement is modified to include possible approaches for improving a priori $NO_2$ profiles as follows: "Further improvement to the retrievals is possible by enhancing the quality of a priori $NO_2$ profiles through improvements in model resolution, emissions, and chemistry, which remain unchanged in the current version."*

L532f: The effect of the a priori is not really removed "altogether" when NO2 profiles are used for model comparison but remains as part of the model error.
*Indeed, the use of model-derived inputs affects retrievals. The context here though is different, and is related to an issue that is often raised while comparing simulated $NO_2$ with retrievals. Eskes and Boersma (2003) discuss an approach of using averaging kernel to remove the effect of a priori $NO_2$ profiles used in retrievals while comparing model-derived $NO_2$ columns with retrievals. An alternative approach of using scattering weights for the same purpose is discussed in Lamsal et al (2014). We have added these references to clarify the context.*

L676: What do you mean by the "alternating nature of the variation"? Please provide more details.
*We realize that the statement was not clear. The statement is now revised as follows: "This alternating nature of the variation in results in polluted versus clean areas suggests that OMI's large footprint size and narrow spiral radius (~4 km) of the aircraft are likely the primary cause for the observed differences. This was demonstrated in Choi et al. (2020) by using high-resolution Community Multi-scale Air Quality Model (CMAQ) simulations."*

L684: Please specify how the "agreement" was computed here and which parameter has improved by 20-35
*The statement is modified in the revised version for clarity as follows: "The use of observed profiles in the OMI retrievals leads to a slight change in correlation, but 20-35% reduction in mean difference between OMI and aircraft observations, highlighting the role of a priori profiles in $NO_2$ retrievals as suggested by previous studies (Russell et al., 2011; Lamsal et al., 2014; Goldberg et al., 2017; Laughner et al., 2019; Choi et al., 2020)."*

L725ff: The sentence is a bit confusing. Does the 0.3 refers to the GLER or to the difference between GLER and LER?
*This value of 0.3 refers to GLER. The statement is correct, but is modified for clarity as follows: "The data from GLER (a unitless value with 0.0-1.0 range) are generally lower, by <0.05, than the climatological LER data over land and ocean outside of sunglint areas; GLER is much higher over the sunglint areas that reaches more than 0.3 due to the geometry-dependent Fresnel reflection."*

L726 L732: The formulation "lower by "the optimized"?
*It should have been "lower than V3.1". It is corrected.*

L742: "may" -> "can"
*Done.*

Figure 1 lists many abbreviations used in the paper. It would be better to list them in a table instead of the caption of a figure.
*Those abbreviations are already defined in the text. Re-defining them in the figure caption helps readers understand various terminologies and acronyms without switching back and forth between other texts or table and this figure.*