Response to reviewer #3

We thank the reviewer for the evaluation of our paper and useful comments that helped improve the manuscript. We appreciate the reviewer's time and effort in reviewing the manuscript. Below are responses to each comment. All reviewer's comments are in the standard font while the responses are in the italic font.

On behalf of the authors, Alexander Vasilkov

Main comments

The method that is presented for the aerosol correction seems highly similar to the method presented by Lin et al (2014). Therefore, the claim that is made in the conclusions that this is new approach is not correct. It is a minor step forward compared to previously published work.

We agree that the method is similar to that presented by Lin et al. (2014) and we state this in Introduction Lines 71-72. However, there are two significant differences. First, it was necessary for Lin et al. (2014) to perform a lot of ad-hoc scaling of their GCM simulation results to match local aerosol observations in order to get realistic aerosol distributions. On the other hand, we are using a global assimilated aerosol product. One of the strengths of using the assimilated aerosol product is that this is generated by the assimilation system on a global scale in a seamless, consistent manner. One other important thing to note is that the GEOS aerosol assimilation product is constrained by MODIS and AERONET AOD observations at 550 nm. This is what differentiates our paper from Lin et al. (2014), and it is this consideration that allows for a global rather than regional methodology. Second, the method by Lin et al. (2014) is applicable to land surfaces only. We have developed a new treatment of surface BRDF for the ocean (Vasilkov et al., 2017). This approach for water surfaces is based on the GLER concept and has been validated in Fasnacht et al. (2019) and allows for a global processing of satellite instrument data.

We have rewritten Lines 73-75 and added the following text in Introduction:
"However, there are some significant differences. For instance, Lin et al. (2014) applied ad-hoc scaling of their global circulation model (GCM) simulation results to match local aerosol observations in order to get realistic aerosol distributions. On the other hand, we use an assimilated aerosol product (Buchard et al., 2017). One of the strengths of using the assimilated aerosol product is that it is processed on a global scale in a seamless, consistent manner. This allows for a global rather than a regional methodology as was the case in Lin et al. (2014) and Liu et al. (2020). The assimilated aerosol product provides a complete set of aerosol optical properties which include the vertically resolved aerosol layer optical depth, single scattering albedo, and phase scattering matrix computed for a given time and space location. Furthermore, the method by Lin et al. (2014) and Liu et al. (2020) is applicable to land surfaces only. We have developed a new treatment of surface BRDF for the ocean (Vasilkov et al., 2017). This approach for water surfaces has been validated in Fasnacht et al. (2019) and allows for a global and consistent processing of satellite NO_2 data."

Although the authors claim that the method can be applied globally, there are computational problems to be resolved (line 312-319). The current description of the method is therefore incomplete for its global purpose. I believe it would be better to postpone publicatios until these problems are solved and a complete description can be given.

The one case that is presented is far too limited. Given that the authors claim to present a globally applicable method, global results for representative time periods need to be presented (e.g a few months). It is impossible to base any conclusions on the one case study that is presented.

The main objective of this study is to lay out and demonstrate the end to end approach of an explicit aerosol correction for a case study in a polluted region for an approach that is ultimately intended for global application. However, we do not initially intend to demonstrate the aerosol correction applicability on the global scale, as it is beyond the scope of this initial feasibility study. We intended to analyze global NO_2 retrievals in the second part of this study. Based on reviewer's suggestion, we processed OMI cloud and NO_2 data globally for the same day of April 5, 2005 as in the manuscript. It appears that the aerosol effect on spatial distribution of NO_2 retrievals is even more complex than expected from the previous model study and existing literature. It is well known that the main aerosol effect on NO_2 retrievals depends on relative vertical profiles of NO_2 and aerosol as well as aerosol optical properties. For clear skies, the aerosol can both increase and decrease sensitivity of satellite instrument measurements to tropospheric NO_2. Of course, the magnitude of this effect depends on aerosol optical depth (AOD) and single scattering albedo and to lesser extent on the phase scattering matrix. For partly cloudy scenes, the presence of aerosol affects both the cloud radiance fraction (CRF) and cloud pressure, a.k.a. cloud optical centroid pressure (OCP). CRF mostly decreases but OCP can both decrease and increase and its retrieval is impacted by the derived CRF. Processing OMI data on an orbital basis reveals additional and complex features of the aerosol effects ultimately on NO_2. We think that the full explanations of the aerosol effects on cloud and NO_2 retrievals could be a topic of a second part of this study.

Accounting for all those considerations we have reworded Lines 284-286 as follows:

"The application of our approach of the explicit aerosol correction to the selected area shows that the NO_2 increase due to the correction is in the right direction of reducing the documented low biases in the NO_2 retrievals with respect to ground- and aircraft-based observations."

and added to the conclusions the following text:

"It should be noted that the above estimates of the explicit aerosol correction effects on cloud and NO_2 retrievals are valid for the selected area. More detailed investigation of the aerosol effects on the global scale will be carried out in the future work".

As described in the literature and section 3.1, the aerosol effect depends on both the aerosol vertical profile and the NO2 vertical profile. Whereas it is clear that the work uses Merra-2 profiles for the aerosols, it is not clear where the NO2 profiles are coming from. This should be

described clearly, and in case there are not coming from Merra -2, it should be made clear why not.

NO_2 profiles and other model-derived information used in the computations are taken from the Global Modeling Initiative (GMI) model. The GMI simulation is driven by the meteorological fields from the MERRA-2. This is clarified in the revised manuscript as follows:

"NO_2 profiles and other model-derived information (e.g., temperature profiles, tropopause pressure) used in the computations are taken from the Global Modeling Initiative (GMI) model. The GMI simulation is driven by the meteorological fields from the MERRA-2. We use the GMI model because the simulations have been run consistently from the start of the OMI mission and this allows us to reprocess results from the entire OMI mission with the proposed aerosol correction."

The fact that this method brings even more model in formation into the satellite retrievals is a concern. How can a user judge how much of the final retrieval product is model and how much is based on the measurements? This should be addressed in detail. Related to this, current NO2 retrievals include averaging kernel information which allow users to replace the assumed NO2 profile with their own profiles. It is preferable if a similar approach is implemented for the assumed aerosol profiles. This should also be addressed in the manuscript.

We do not agree with the reviewer's characterization of our approach. MERRA-2 includes assimilation of aerosol optical depth from various ground and space-based remote sensing platforms. Our motivation for using aerosol information from MERRA-2 instead of a chemical transport model (e.g., GEOS-Chem) was to include observationally constrained data which takes advantage of the relative strengths of a model and observations.

The NASA NO_2 standard product does not include averaging kernel, but rather it includes scattering weight (SW) profiles and tropospheric and stratospheric AMFs that allow users to re-calculate the averaging kernels and VCDs using their own a-priori NO_2 profiles. This is possible because SWs are independent of a-priori NO_2 profiles. The new SWs accounting for aerosol profiles and surface BRDF will be provided in the new NO_2 product. The users can use the new SWs to correct for their custome NO_2 profiles. However, using a different aerosol profile and surface BRDF would require re-calculating SWs. It is an inherent issue with any explicit aerosol correctio approach. We believe that our approach represents a reasonable compromise, which is feasible for global processing.

The sophistication of the aerosol and the cloud model seems to be out of balance. Whereas for the aerosol model a state-of-the-art aerosol model is applied, clouds are represented by a simple Lambertian clouds. The choice of this cloud model should be substantiated. Note that the retrieval results be as good as the weakest link in the chain.

For trace-gas retrievals it is important to estimate photon path lengths in the atmosphere that determine trace-gas absorption and thus affect the measured TOA radiances. Photon path lengths in a cloudy atmosphere are determined by the following most important cloud

parameters: the geometrical cloud fraction, the cloud optical depth, and the cloud vertical extent. Because of limited cloud informational content in TOA radiances, these three parameters cannot be retrieved simultaneously from the radiance measurements. That is why it is necessary to take on additional cloud assumptions. For instance, if a model of the Mie scattering cloud layer is used (Loyola et al., AMT, 2018), there is a need to assume a priori values for the cloud microphysical parameters and cloud vertical extent, to assume a homogeneous cloud layer and to add information about the cloud fraction from other measurements. We use a simpler model, the so-called mixed Lambertian-equivalent reflectivity (MLER) model that combines the independent pixel approximation and the treatment of cloud and ground as horizontally homogeneous, opaque Lambertian surfaces (Koelemeijer et al., JGR, 2001). The MLER model compensates for photon transport within a cloud by placing the equivalent Lambertian surface somewhere in the middle of the cloud instead of at the top (Stammes et al., JGR, 2008; Vasilkov et al., JGR, 2008; Sneep et al., JGR, 2008). As clouds are vertically inhomogeneous, the pressure of this surface does not necessarily correspond to the geometrical center of the cloud, but rather to the so called optical centroid pressure (OCP) (Joiner et al., AMT, 2012). The cloud OCP can be thought of and modeled as a reflectance-averaged pressure level reached by backscattered photons. The cloud OCP is the appropriate quantity for use in trace-gas retrievals from satellite instruments. Cloud-top pressures, e.g. those derived from thermal infrared measurements, are not equivalent to OCPs and do not provide good estimates of the required solar photon path lengths through clouds that are needed for trace-gas retrievals from UV–vis backscatter measurements (Vasilkov et al., JGR, 2008; Joiner et al., AMT, 2012). It has been demonstrated that the MLER model works reasonably well for trace-gas and cloud algorithms (Koelemeijer et al., JGR, 2001; Veefkind et al., IEEE T. Geosci. Remote, 2006; Stammes et al., JGR, 2008; Boersma et al., AMT, 2011; Bucsela et al., AMT, 2013; Veefkind et al., AMT, 2016).

Detailed comments

Section 2.5.1 This section should also describe cases for which the a-priori information is inconsistent with the measurements. For example, the following cases may occur: -The ECF becomes less than zero; -The ECF becomes larger than one; -The ECF is zero, but the SCD for O2-O2 is not consistent with aerosol profile. These are important details and all known geophysical situations should be covered in an algorithm paper.

We added the following text at the end of Section 2.5.1:

For a very small fraction of the ECF retrievals, ECF values can be outside the physically meaningful range of zero to one. We keep all the ECF retrievals in output orbital files thus providing the necessary diagnostic information on these physically unreasonable cases. Additionally we provide the clipped ECF retrievals, that is negative retrieved ECF values are replaced with zero and ECF values greater than one are replaced with one. Similarly, we provide these clipped CRF values as the input for the OMI NO_2 algorithm. A small fraction of the cloud OCP retrievals can also appear to be unphysical (values greater than surface pressure) (Veefkind et al., 2016, Vasilkov et al., 2018). Again, we keep all OCP retrievals in output files and additionally provide clipped cloud OCP retrievals by replacing OCP values greater than the surface pressure with the actual surface pressure.

Section 2.5.1, line 183. Details shall be provided on how the equation is solved. What numerical methods are used?

We added the following at the end of Section 2.5.1:

To solve Eq. (4) we rewrite it in the form:
$SCD_c(P_c) \equiv AMF_c(P_c)*VCD(P_c) = [SCD - AMF_g*VCD_g*(1-f_r)] / f_r$
where quantities on the right hand side of the equation are known, in particular, the quantity SCD is retrieved from the spectral fit of the OMI measurements around the O_2-O_2 absorption band at 477 nm (Vasilkov et al., 2018). Using LUT values of $AMF_c(P_c)$ and calculated $VCD(P_c)$ we then find the LUT pressure nodes P_1 and P_2 for which the following inequality is valid:
$AMF_c(P_1)*VCD(P_1) < AMF_c(P_c)*VCD(P_c) < AMF_c(P_2)*VCD(P_2)$
or equivalently,
$SCD_1 (P_1) < SCD_c (P_c) < SCD_2 (P_2)$.
Then P_c can be obtained by linear interpolation of P over SCD:
$P_c = [(SCD_c - SCD_1)*P_2 + (SCD_2 - SCD_c)*P_1] / (SCD_2 - SCD_1)$.

Section 2.2, line 121 How does Merra-2 deal with the very strongly non-linear growth of aerosol particles for relative humidities > 90%. This may be a frequently occurring at the top of the boundary layer for partly cloudy conditions and have significant effects on the AMF.

We agree that non-linear growth of aerosol particles can be important. However this topic is beyond the scope of our paper. Here we note that MERRA-2 does account for particle hygroscopic growth. Aerosol hygroscopic growth depends on simulated relative humidity and is considered in computations of particle fall velocity, deposition velocity, and optical parameters (Randles et al., 2017).

Section 2.3 Provide detailed information on the setup of the RT calculations wrt the aerosol optical properties, such as the number of streams used in calculations, etc.

In Section 2.3 we have added the following:
VLIDORT computes the single scattering contribution exactly in a spherically-curved atmosphere using the full scattering matrix. For multiple scattering, VLIDORT treats the direct solar beam attenuation in the pseudo-spherical approximation. This study used the delta-M scaling option to treat sharply peaked aerosol phase functions (Nakajima and Tanaka, 1988). We used 12 discrete ordinate streams in the polar hemisphere half space for the computation.

Nakajima, T., and M. Tanaka, Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation. J. Quant. Spectrosc. Radiat. Transfer, 40, 51-69, 1988.

Section 3.2, line 249 The demonstrated effect is clearly not only because of BRDF effects. The largest effect is due to different source of the surface reflectivity data. If non BRDF effects were taken into account a similar effect could be expected.

This is true and it was discussed in Vasilkov et al. (2017). Here we have added the following:

"It should be noted that the differences include both BRDF effects and biases between the MODIS and OMI-based surface reflectance data sets. This is because the BRDF data and thus the GLERs are derived from atmospherically-corrected MODIS radiances while the climatological LERs are inherently affected by residual aerosols. Additionally, climatological LERs can be contaminated by clouds due to the substantially larger OMI pixel size as compared with MODIS footprints."

Section 3.2 line 261 Also calibration differences between OMI and MODIS and atmospheric correction in MODIS should be discussed here.

We added the following:

"Calibration differences between OMI and MODIS are discussed in Qin et al. (2019) and specific details are provided in Appendix D: "Relative calibration of OMI and MODIS" of that paper. To summarize: MODIS Collection 5 radiances (used to derive BRDF kernel coefficients and thus GLER values) are higher than OMI Collection 3 radiances by approximately 1%. A sensitivity analysis of the equation used to compute GLER shows that a 1% error in TOA radiances will produce errors in LER of up to 0.003 in surface reflectivity. This value is much lower that the reported average difference between the climatological LER and GLER of 0.03. The atmospheric correction for MODIS band 3 used in this study has a theoretical error budget of about 0.005 reflectance units (Qin et al., 2019). Again, this error is much lower than the reported average difference suggesting that neither the calibration differences nor the MODIS atmospheric correction are major contributors to the observed difference between climatological LER and GLER."

Section 3.2 line 298: "An interesting feature of the explicit aerosol correction on OCP is that the OCP can be reduced for a small fraction of the pixels." This sentence is not understood.

We have reworded this sentence as follows: "An interesting effect of the explicit aerosol correction on OCP is that OCP values for high altitude clouds are lower for a few pixels within the selected area, while in general OCP are higher for the remaining bulk of pixels."

Section 3.2 line 297: At low cloud fractions the errors in the OCP will explode. What is the assumed error in the OCPs? How does an increase of 50 hPa relate to the expected accuracy of the OCP?

That is correct. To clarify the issue we have added the following text:

"An OCP error is amplified with lower cloud fraction values. This is true all cloud pressure algorithms. In addition to OCP, we retrieve the so-called scene pressure (Vasilkov et al., 2018). In the absence of clouds and aerosols, the scene pressure should be equal to the surface pressure. A difference between the scene pressure and surface pressure can be considered as estimates of the OCP retrieval bias. This bias is about 40 hPa. Thus an increase of 50 hPa is comparable to the expected accuracy of the OCP retrievals. However, in our work we compare the OCP retrievals with and without the explicit aerosol correction. Even though these retrievals

possess bias, difference between them, e.g. increase of 50 hPa due to the implicit aerosol correction, does make sense."