Dear Editor,

We thank the reviewer for the evaluation of our paper and useful comments that helped to improve the manuscript. Below are our responses to each comment. Reviewer's comments are in blue, the responses are in black; the text added to the manuscript is in red.

On behalf of the authors,

Wenhan Qin

Responses to comments from Referee #2:

(1) The paper provides some explanations for the differences that are found between the MODIS-based GLER and the OMI-derived LER. These explanations are neglect of aerosol in the simulations, the possibility of cloud contamination in the OMI observations, and calibration differences between the MODIS and OMI instruments.

The reviewer is correct that in the simulations, we did not consider the aerosol or cloud as it is stated in the last paragraph of section 2.4 (page 9): aerosols are not included in the VLIDORT simulation of TOA radiances (from which GLER is computed). So, the GLER-LER comparison screened OMI data to minimize impacts of absorbing aerosols and clouds.

We realized even after applying proper criteria of aerosol index (AI) <0.5 and effective cloud fraction (ECF)<0.04 for data screening, the GLER- LER difference is still subject to the impact from background aerosols (mostly non-absorbing) and residual clouds, as we clarified in section 3.1. However, current generation of the satellite cloud and trace gas retrieval algorithms do not treat aerosols explicitly and this analysis is beyond the scope of this paper.

We agree that the impact of the Level 1b calibration differences between OMI and MODIS needs to be addressed, and we added more details in the updated manuscript (see detailed response below).

What is actually known about the calibration of MODIS compared to that of OMI? Haven't there been any studies comparing the two instruments? What are the calibration differences needed to explain the systematic differences between the MODIS-based GLER and the OMI-derived LER?

This is an important issue, and we are grateful to both reviewers for drawing attention to it. We have added information on the uncertainties of the MODIS instruments to section 2.2.

The calibration uncertainty for MODIS band 3 is within 2% (Xiong et al., 2005). The MODIS Aqua solar reflective bands including band 3 were corrected for a time-dependent drift in Collection 5 (Wu et al., 2013) but errors in MODIS Terra of up to 5% across the scan developed approximately 5 years after launch and this error was not sufficiently corrected in Collection 5 (Sun et al., 2014; Lyapustin et al., 2014).

We also added information on what is known of the relative calibration of the two instruments, and the effect of calibration error is on GLER and LER differences in new paragraphs in the section 4 (Discussion).

In addition to background non-absorbing aerosol and/or residual cloud contamination, it is important to consider that the GLER-LER bias may be due in part to differences in the MODIS and OMI radiance calibration. Sensitivity analysis of Eq. 2 used to compute LER and GLER shows that a 1% error in TOA radiances will produce errors in LER of up to 0.003 in surface reflectivity. A bias of 0.01 between GLER and LER requires a difference in MODIS and OMI TOA radiance of at least 3% for brighter land scenes (LER >=0.2), and differences of 10% for darker land scenes (LER <= 0.05). MODIS TOA radiances would thus have to be 3-6% low relative to OMI to explain the bias seen in GLER-LER for bright scenes, and 10-20% low for dark scenes.

Jaross and Warner (2008) compared TOA reflectances from OMI and MODIS with radiative transfer model simulations over Antarctica, accounting for the BRDF of the snow surface. By indirect comparison, OMI Collection 3 and MODIS Collection 5 agreed to within 1% at the start of the OMI mission. They estimated the uncertainty of their technique is 2%. This level of disagreement is smaller than needed to explain all of the 0.01-0.02 bias of GLER over dark scenes. We therefore conclude that only some of the bias can be attributed to calibration differences. Additional information about the relative calibration of OMI and MODIS is provided in Appendix D.

Relative sensor drift is also a concern in comparing the GLER product using the MODIS calibration with LER from OMI. Aqua MODIS appears to be well corrected in Collection 5 but the MCD43 product also uses data from the Terra instrument, which has degraded appreciably over the lifetime of the mission. However, we find no evidence of time dependent change in Collection 5 MODIS BRDF data. We suspect the time-dependent and scan angle-dependent error in the Collection 5 MODIS Terra calibration data have somehow been avoided. Since OMI drift has not been fully corrected, and the MODIS drift has been removed (or avoided in the case of Terra, apparently) the slight decrease of OMI LER relative to GLER between 2006 and 2015 in figure 8 may be due to the 1-1.5% calibration drift in OMI radiances.

We have also added Appendix D, Relative calibration of OMI and MODIS, to provide additional information about the relative calibration of the level 1b data from the two instruments used in this study.

For the GLER no OMI measurement data are used. Only the OMI ground pixel extent and geolocation are used and MODIS BRDF data are then used to calculate the GLER for that OMI pixel. So, the MODIS-derived GLER, which is to be used for OMI retrievals, inherited the calibration differences that exist between MODIS and OMI.

That is correct. The MODIS-based GLER inherits the MODIS calibration uncertainty, and when is used for OMI retrievals, it introduces the calibration differences between MODIS and OMI Level 1b data.

For the GLER calculation, not only the OMI ground pixel extent and geolocation are used, but the OMI geometries (solar, viewing and relative azimuth angles) are needed because GLER depends on all these angles.

It might be good to mention (more explicitly) in the paper that using the GLER for OMI-based retrievals can result in the introduction of calibration inconsistencies in these retrievals.

Thank you for pointing this out. We added the following into the 3rd paragraph of the section 4 (Discussion).

Sensitivity analysis of Eq. 2 used to compute LER and GLER shows that a 1% error in TOA radiances will produce errors in LER of up to 0.003 dependent in the surface reflectivity. A bias of 0.01 between GLER and LER requires a difference in

MODIS and OMI TOA radiance of at least 3% for brighter land scenes (LER >= 0.2), and differences of 10% for darker land scenes (LER <= 0.05). MODIS TOA radiances would thus have to be 3-6% low relative to OMI to explain the bias seen in GLER-LER for bright scenes, and 10-20% low for dark scenes.

(2) What are, to your knowledge, the expected differences between MODIS-derived GLER and the MODIS BRDF? In other words, what are the errors the user would introduce if he/she would use the MODIS-derived GLER as if it would be a BRDF?

GLER is calculated by exactly matching OMI-measured and calculated TOA radiances over a non-Lambertian surface as described in Eq.2. However, the average photon path lengths for the measured and calculated signals can still be different and the larger the viewing zenith angle, the larger the difference. Our simulations have shown the difference between GLER and BRF does not exceed 0.01 at SZA of 45° and viewing zenith angle < 70° within the solar-OMI principal plane. Vasilkov et al. (2017) reports the simplification of using the GLER approach relative to the full BRDF treatment can lead to small biases in the calculation of AMFs of up to 6% (see Fig. 10a) and similar biases in the retrieved NO2 vertical columns.

Is the 0.8% mentioned in section 2.1, page 5, lines 18-19 a representative percentage for this? Is this 0.8% valid for 466 nm?

The referenced value of the error 0.8% due to Lambertian assumption does look small compared to the current literature. Although this result is for band 1 (645 nm), a similar result is expected for band 3 (466nm) which we use in this study, because for land surfaces other than deserts the surface reflectance in these two bands is similar.

Please note that we removed this statement in our revision because other independent studies have shown different numbers for the Lambertian assumption. More importantly, the uncertainty in our GLER product is tied more to uncertainty of MCD43 data than MOD09, which the above number pertains to.

Does the term "green band" in line 20 refer to the 555-nm MODIS band (band 4)?

Yes, this is correct: "green band" for MODIS is band 4 (550 nm) in their terminology.

(3) The paper introduces GLER calculated at 466 nm. For this, MODIS BRDF information is used from MODIS band 3. However, MODIS band 3 is centered around 470 nm, and not around 466 nm. On page 7, lines 6-7, an explanation for the 4-nm difference is given which suggests that the 470-nm MODIS BRDF is just used "as is" at 466 nm

If this is the case, wouldn't it be better to say that the GLER is representative for 470 nm(and not for 466 nm, even though the GLER retrieval is performed at this wavelength)? What are the technical reasons for doing the GLER retrieval at 466 nm and not at 470nm? Is it because of absorption by trace gases at 470 nm? If so, which trace gases?

MODIS band 3 has a 20 nm bandwidth (459-479 nm) centered near 470 nm. The land surface reflectivity does not have high-frequency spectral structure within this 20 nm range so wavelengths in this range, such as 466 nm, will be representative of the band in terms of land surface reflectance.

However, the same is not true for the atmospheric effects. The reason we picked 466 nm for the GLER calculation is because (1) 466 nm is further away from the O2-O2 absorption centered around 477 nm, and from an ozone absorption feature near 470 nm, and (2) 466 nm is the wavelength used in our cloud algorithm to retrieve effective cloud fraction (ECF) (Vasilkov et al., 2017). Observations at 466 nm are relatively free of atmospheric inelastic, rotational-Raman scattering (RRS) as well, which is important because OMI has narrow band pass and thus is sensitive to the RRS errors.

We modified the last paragraph of section 2.2 as follows.

Since kernel coefficients depend on wavelength, for the present study we selected MODIS band 3, the shortest wavelength in the MCD43GF product, with a center wavelength of 470 nm (ranging from 459 to 479 nm) to represent 466 nm, which is the wavelength used in our cloud algorithm to retrieve effective cloud fraction (ECF) (Vasilkov et al., 2017). Observations at this wavelength are relatively free of atmospheric rotational-Raman scattering (RRS) and trace gas absorption.

(4) section 1, page 3, line 9:

Here the paper mentions a few papers on LER databases. Could you also mention the databases/instruments behind these references?

Please also add the (more recent) reference to Tilstra et al. (2017), about LER retrieval from the GOME-2 instrument. https://doi.org/10.1002/2016JD025940

We appreciate the reviewer pointing out these LER datasets from GOME/2 and SCIAMACHY. We added references to those databases as well as the recent publication as suggested.

We also added the following to the Introduction:

For example, Herman and Celarier (1997) from the Total Ozone Mapping Spectrometer (TOMS) at 340 and 380 nm, Koelemeijer et al. (2001) from the Global Ozone Monitoring Experiment (GOME) in 11 wavelengths between 335-772 nm, Kleipool et al. (2008) from OMI in 23 wavelengths at 328-499 nm, and more recently Tilstra et al. (2017) from GOME-2 in 21 wavelengths between 335-772 nm as well as from the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) in 29 wavelengths from 335-1670 nm. These climatologies are constructed by computing statistical values representative of multiple years of observations made with different sun-satellite viewing geometries. In order to minimize cloud contamination, they may be based on a lower percentile (e.g., Herman and Celarier, 1997) and/or the mode of the LER histogram depending on surface type (e.g., Koelemeijer et al., 2001; Kleipool et al., 2008; Tilstra et al., 2017).

(5) section 2.3/Fig. 3:

When explaining how the point-in-polygon method is used, please mention briefly that the "real" OMI pixel is not rectangular, and that this fact alone already can lead to (small) differences between the (MODIS-derived) GLER and the OMI-derived LER.

We agree with this point made by the reviewer and we moved this to the Appendix A1 with more information describing our collocation methodology, which now says the following:

The collocation methodology is shown schematically in Fig. 12. The OMI pixel is first defined from the four ground pixel corner points provided in the OMPIXCOR data product as a 4-sided polygon. A sample space is then constructed along constant latitudinal boundaries, with the corner points tangent to the boundaries of the sample space as shown. All pixels from the MODIS BRDF/Albedo product and ancillary data sets inside the sample space are tested using the point-in-polygon method (Haines, 1994). For this application, we used the corner points for the VIS channel corresponding to 75% of the energy in the along-track field of view. This definition assumes the pixels across the track share boundaries with their two adjacent neighbors (except for the pixels at the far edge of the swath), while the pixels along the track of the satellite overlap (reference to OMPIXCOR Readme). De Graff et al. (2016) showed the actual shape of the OMI pixel is not exactly a rectangular polygon but rather is best represented by a super Gaussian. They also showed that the optimal overlap function between OMI and MODIS depends on the scene and the time difference between the satellites. We do not consider these factors as critical to this application because the GLER is based on MCD43GF, an 8-day gridded MODIS BRDF product from Terra and Aqua. Small errors in the pixel shape should only have a minimal impact on our results.

Perhaps you could refer to the paper by De Graaf et al. (2016) about the size and shape of the OMI pixels. https://doi.org/10.5194/amt-9-3607-2016

We thank the reviewer to point this out. Indeed, the actual shape of the OMI pixel is not exactly a rectangular polygon but rather is best represented by a super Gaussian. We added the reference De Graff et al. (2016) as suggested.

In this paper, OMI and MODIS (band 3) reflectances are compared to each other using different sizes and shapes for the OMI point-spread function.

This is correct, and so we have to collocate MODIS pixels at 30 arc-second resolution within the OMI pixels as described in Appendix A1.

(6) section 3.2, page 13, line 16:

The paper mentions here that surface BRDF does not change on a day-to-day basis. But this can happen in certain cases, as explained in section 3.3, page 15, lines 12-14. Perhaps you could change the sentence to "While surface BRDF in general does not ..." (or something similar)?

We agree with this comment and have made minor changes to this sentence in section 3.2.

Surface BRDF or albedo change is small on a day-to-day basis, with the exception of extreme events such as fires and floods that are not captured with the 16-day MODIS dataset. There is, however, noticeable variability in the BRDF and albedo between seasons due to land cover changes throughout the year.

(7) In the paper the GLER and the OMI-derived LER are also compared to the Kleipool climatology. Which field is taken from the Kleipool climatology? Is it the field "MonthlyMinimumSurfaceReflectance" or is it the field "MonthlySurfaceReflectance"? If it is the "MonthlySurfaceReflectance" field, then that would probably explain part of the higher values of the Kleipool climatology compared to the OMI-derived LER for at least some of the land cover types in Figure 9. In fact, I think that in these analyses it would be better to use the traditional "MonthlyMinimumSurfaceReflectance" field.

In any case, it would be good to mention in the paper which of the two fields was used in the analyses.

The field "MonthlySurfaceReflectance" is used because this is what was used in our NO2 and O2-O2 cloud algorithms before switching to GLER. But we acknowledge this is a common point of confusion, so for comparison purposes, we include data from both the “Monthly Surface Reflectance” and “Monthly Minimum Surface Reflectance” fields in figures in our revised manuscript.

(8) small typo in a reference, page 27, line 20:

Haines, E.,: –> Haines, E.:

Thank you for pointing this out this typo. This has been corrected in the revised version.