Answers to RC-I comments

We thank the reviewer for their time and effort reviewing this manuscript and for providing helpful and constructive comments that have helped to improve the manuscript.

RC1: "However, I am missing some important pieces of information:

1) how to calculate the Decorrelation Index DI exactly

2) a systematic discussion of what influences DI and why

3) how were the threshold values for DI in each spectral interval derived?

The first point could easily be mended by providing a mathematical formula for DI. The second point is treated dispersed through the paper, with some information only anecdotically. The authors state that aerosols, Rayleigh scattering, blue water give rise to changes / increases in DI, and provide a few examples, without explicitly telling why DI is influenced. Without providing an exact calculation method of DI that is difficult to understand at first. Assuming that "r" is the Pearson's correlation coefficient, the slope of the spectrum also enters the equation. However, that also implies that the threshold reacts different to anomalies for ascending and descending spectral slopes. This should be discussed at the beginning, after introducing "r". Also it should be mentioned explicitly that perfectly normal spectral features, such as atmospheric absorption lines, give rise to enhanced DI. Several other aspects on the behaviour of DI are stated but may require reading between the lines to understand; in general statements should explicitly explained and without the need to reread sentences several times. Insight in the meaning of DI would be much clearer if all influences were discussed in one place, e.g. Section 2.2, instead of providing something in the introduction and letting others develop along the line as examples are given. It remains intransparent how threshold values for DI were derived. But that is important if the reader wishes to apply the method to other instruments. Were thresholds derived based on radiative transfer calculations (including Raman scattering, aerosols, generic spectral surface albedos?) or was this a trial-and-error process until some credible results were obtained?? Was there any deliberate matching to OMI saturation flags?"

AC1: We provide the requested equations in the revised section 2.2 and follow-up with an extensive discussion of the expected DI behaviour:

"Evidently $DI = 0$ for the simple case of a perfect match with $I = const * F_0$; if $I = -const * F_0$, then $DI = 2$. If I and F_0 are completely unrelated, then $DI = 1$. Considering the 'smooth' (low-frequency) component of I and F_0, we expect them to be generally correlated in the spectral regions relatively free of major atmospheric absorptions (ozone in particular). The correlation would be inevitably diminished by the wavelength-dependent Rayleigh scattering and surface reflectivity. Once a multitude of deep spectral lines is superimposed on a smooth envelope, DI will depend mainly on a match between the shape and position of these I and F_0 spectral transitions, with the correlation depending on the S/N of the tested radiances and irradiances, and even more so on slight (in OMI's case) wavelength mis-alignments between radiances and irradiances, with the steep line flanks magnifying the differences.

An additional de-correlating factor is brought forth by the omni-present atmospheric rotational Raman scattering (e.g., Joiner et al., 1995). Under the circumstances, one may never expect $DI = 0$ save the exceedingly rare cases of a perfect solar glint. It is known that Pearson's correlation coefficient is sensitive to outliers, thus simplifying detection of spectral distortions in the high-resolution data compared to the low-resolution cases, with the latter tending to lessen the impact of additive components (the shallower lines are potentially less susceptible to stray light), as well as the wavelength mis-alignment (spectral blending of multiple features leading to partial canceling of distortions in the adjacent features). At a given spectral resolution and S/N, DI sensitivity may grow with increasing numbers and contrasts (depths) of spectral features in the chosen spectral window.

At the same time, the DI is expected to be sensitive to artifacts associated with cosmic ray hits. The interval 440-480 nm, where there are few deep spectral lines, should be especially sensitive to geophysical factors, for example, to the wavelength-dependent albedo of the earth's surface. Note that there are cases when direct solar radiation F_0 is mixed with I due to instrument problems (see below).

Under this specific circumstance DI will decrease, since correlation between $(I + \delta F_0)$ and F_0 is always higher (thus DI lower) than between I and F_0. A similar effect occurs with sunglint from the water surface, when the proportion of directly reflected sunlight in I increases significantly. Note that in the current approach we do not compensate for the relatively smooth spectral differences imposed by atmospheric (Rayleigh scattering) and surface (wavelength-dependent albedo) factors, leaving this to

the next DI version. This step would make DI more sensitive to the instrument-imposed anomalies, further disentangling those from the geophysical factors (see below)".

We return to this discussion on multiple occasions (the blooming effects in particular) throughout the revised text. We also provide the explanation about the proposed thresholds:

"The provisional (the user may redefine the values using the auxiliary data provided in the OMI DI product) DI thresholds were determined as follows. We used all available, mission-long OMI UV2 and Vis radiances. For each orbit and for every spectral window we constructed DI histograms. Then we selected numerous cases sampling the tails of the DI histograms. On a case-by-case basis, for different scenes and spectral windows, we found empirically the lowest-DI values that repeatedly separate the scenes with apparently normal (spectrally smooth, with the fine-structure, low-amplitude Raman-scattering features) and distorted reflectances. These DI thresholds approximately correspond to 99.995-99.998 percentiles in the DI distributions. We plan to provide a statistically rigorous threshold definition in the improved DI version".

We have not deliberately matched the operational and the DI flags, though comparing them on multiple occasions (see Figures 1-7).

RC2: "Specific Comments:

Title (and abstract): the use of the word "non-linear" is not appropriate if also straylight is included (assumed is spatial straylight??). It may be that straylight has a non-linear effect on calibrated radiances compared to TOA, but for the instrument the amount of straylight is linearly proportional to the amount of input light into the telescope (for a certain geometry). The same for obstruction due to MLI. Also a cosmic hit may (statistically) be linear with particle flux (for a given particle type/energy/angle). It would be more precise to use the word "anomaly". (in line 73 this is correctly used for OMI flags)"

AC2: New title: ***Detection of anomalies in the UV/Vis reflectances from the Ozone Monitoring Instrument***

RC3: "Line 35: I understand this introduction is the standard advertising for OMI, but OMI NRT/VFD dissimination is really irrelevant for this paper - please remove."

AC3: The reference to the OMI NRT/VFD dissemination has been removed.

RC4: “Line 48: it is not clear that the term "blooming" is explained by the first part of the sentence. This is then better done in line 65. Please move and integrate line 65 to here.”

AC4: The explanation of the blooming effect has been moved to where it was first mentioned.

RC5: “Line 54: this may suggest all GOME-2 (2A, 2B, 2C) sensors have an issue with clouds. As far as I recollect the issue was solved by introducing coadding. Please rephrase or leave this out, since you make abundantly clear that saturation effects are common. Maybe mention that saturation is simply a common effect due to the much larger dynamic range of TOA radiance compared to detector dynamic range.”

AC5: The reference to GOME-2 is removed.

RC6: “Line 80-86: this is a general statement on radiance versus irradiance. While it is OK to make such a statement in the introduction, it is not sufficient to regard this as explanation of the behaviour of DI (see general comment).”

AC6: A systematic discussion of the factors that affect DI is provided in Section 2.2 that has been substantially revised.

RC7: “Line 83-85: ...depends mainly on the strength(depth) of solar Fraunhofer features... The depth of the solar lines by itself doesn’t change DI. What you want to say is that the low radiance in line cores makes them more susceptible to additive effects. Please rephrase. I wonder if the sentence is not better moved to a section that describes DI more in detail (see general comment).”

AC7: We amend a part of Introduction (there we talk about the ‘traditional’ correlation) and return to the subject in the revised Sect. 2.2.

RC8: “Line 116-119: a mathematical formula for the calculation of DI should be given here.”

AC8: The equation for the DI is given in the revised Sect. 2.2.

RC9: “Line 124: is DI not always >= 0 if atmospheric absorption is present ? (formula for DI needed !) and what is the influence of a non-flat spectral albedo on DI ? (again: formula for DI needed). In general, would DI not always be >= 0 unless the reflectivity decreases with wavelength? As per my general comment, I propose to discuss that here, together with noise effects, and the resulting behavior on parameters currently discussed in lines 80-86.”

AC9: Please see the revised sect. 2.2, as well as additional comments throughout the revised text.

RC10: “Table 1: the comment "strong spectral lines" is unclear. And why does this coincide with low DI thresholds?”

AC10: We provide additional details in the revised Sect. 2.2.

RC11: “Table 1, DI thresholds: see my general comment 3). Even if DI thresholds may depend on application, it must be described how the thresholds in this paper were derived, such that users may get a handle on how to set the threshold (for their application or for other instruments). "We just take these values and it works" is not enough. IMPORTANT: The paper is not acceptable without a proper description here. Scientific results must be reproducible and traceable. (as you know of course... I don’t expect pages with analysis but say what you did so others can replicate)”

AC11: Please see the answer to RC1 who had a similar concern. We have added relevant clarifications to the text of the article to address these points.

RC12: “Line 269-280: this is one of those examples where it is left to the reader to guess why exactly DI is deviant. The basics of this (spectral slope?) should have been laid out before (see comment to line 124) and it would not harm to remind here why Rayleigh scattering has an effect ("contributes significantly to the top-of-atmosphere radiance" is a bit non-descript..)”

AC12: We amend the text in the discussion of Sect. 3.3 to address this point.

RC13: “Line 302: Why does scattered light from the thermal blanket "leads to the significant decrease of DIs". Not increase? And does "the blocking of the incoming Earth shine" result in distortion of the spectral shape ??”

AC13: We expand and clarify the discussion of the RA phenomenon and link it to the observed DI patterns.

RC14: “Line 321: why does low reflectivity (solar eclipse) increase DI but scattered light (line 301-303) lowers DI (should that also be: increases)? Is the solar eclipse effect due to noise or due to spatial straylight from around the occultation zone? Is scattered light not spatial straylight?”

AC14: We elaborate on both questions in the revised text:

“…a solar eclipse zone. Though we cannot disentangle all the contributing factors for the latter, here we mention two of them as the likely causes of the high DI values (thus enhanced distortions in the

reflectances): the low S/N of the eclipsed radiances, as well as the drastically increased portion (compared to the normally-lit scenes) of the additive (straylight) component."

"...the reflective surface leads to enhanced spatial cross-talk between adjacent RA-affected scenes (an anomalous stray light that is regulated by the wavelength- and angle-dependent reflectivity of the blanket)."

RC15: "Line 329-332: I wonder if a high-pass filtering (e.g. dividing radiance by a local polynomial fit) would not largely remove the effects of aerosol and surface reflectivity on DI, and provide better sensitivity to anomalies in Vis. Also the "search for areas of clear ocean water" could probably just as well be done using a "slope index" based on 2 (continuum) wavelengths. Please comment / address."

AC15: We plan to develop a revised index DI-2 in the future, which will use the decomposition of the spectrum into high-frequency and low-frequency components; this should decrease the impact of geophysical factors. We mentioned this in the revised text. .

RC16: "Typographical comments:

The figures provide important visual information. However, most are disproportionally large compared to the text and to the required resolution (it is really not necessary to visually locate every single pixel). Especially figures [2,] 3,5,6,[8,9,] 10,11 should be reduced in size".

AC16: The sizes of all the listed figures are reduced

RC17: "Also the font size in Table 1 is disproportionally large and should be reduced."

AC17: The font in the Table 1 is reduced

RC18: "Typos:

double dots in lines 51, 94

double white space in line 169 ?

Dis -> DIs (?) line 301, 303

missing space line 303 (Fig. 11.Figure 11)

(otherwise kudos for a well-edited syntax !)"

AC18: All typos have been corrected.

Answers to RC-II comments

We thank the reviewer for their time and effort reviewing this manuscript and for providing helpful and constructive comments that have helped to improve the manuscript.

RC1: "It is not clear to me how in practice the DI coefficient was calculated and how the threshold values for different wavenumber ranges given in Table 1 were established."

AC1: The equation for the Decorrelation Index (DI) is written in 2.2. The DI is a mathematically strictly calculated parameter, and the threshold is only a rough estimate. We have added relevant clarifications to the text of the article:

"The provisional (the user may redefine the values using the auxiliary data provided in the OMI DI product) DI thresholds were determined as follows. We used all available, mission-long OMI UV2 and Vis radiances. For each orbit and for every spectral window we constructed DI histograms. Then we selected numerous cases sampling the tails of the DI histograms. On a case-by-case basis, for different scenes and spectral windows, we found empirically the lowest-DI values that repeatedly separate the scenes with apparently normal (spectrally smooth, with the fine-structure, low-amplitude Raman-scattering features) and distorted reflectances. These DI thresholds approximately correspond to 99.995-99.998 percentiles in the DI distributions. We plan to provide a statistically rigorous threshold definition in the improved DI version".

RC2: "In figure 1 authors compare DI with the number of SPW flags for a very restricted range in the spectral space (414-424nm). It is not clear why such range was chosen – it is different in Figs. 2 and 3."

AC2: DI for 14 spectral intervals of ~10 nm are calculated in the Vis range of 350-498 nanometers. The behavior of the DI in each interval has its own characteristics. In Fig. 1, we showed the spatial dependence of the DI for one interval, which has a significant sensitivity to changes in the spectrum, in Fig. 2 - spectral dependence of the DI for several intervals, in Fig. 3 - spectral dependence of the DI for the entire Vis area.

RC3: "According to table 1 on page 5, DI thresholds for damaged spectra depend on the spectral region and vary considerably (by two orders of magnitude). On the other hand, in figures 2,3,5,6 only the actual value of DI is plotted. It is therefore difficult to say how much DI exceeds the threshold. I suppose it would be better to divide the actual value of DI by the threshold value for the particular spectral range to better illustrate the degree of deviations."

AC3: Figures 2,3,5,6 show real DI, which mathematically strictly show the correlation between the terrestrial and solar spectrum. The thresholds that are set in Table 2 are very rough estimates. If we divide the exact result by an approximate factor, which varies greatly from interval to interval, we make it difficult to quantitatively interpret the picture.

RC4: "In the introduction the authors address two different effects which may deteriorate measurement data: saturation and blooming. After reading description on page 2 it is not clear to me how to differentiate in practice between the effect of the two. In both cases, as the authors write, flow of excessive electrons to neighboring pixels occurs."

AC4: An improved version of the explanation for both effects is now as follows:

"Saturation occurs when bright light causes the number of electrons in a sensor pixel to exceed either the maximum charge capacity of an individual charge-coupled device (CCD) photodiode, or the maximum charge transfer capacity of the sensor. A blooming effect occurs when electrons from a highly illuminated pixel of the CCD matrix jump to a neighboring pixel, causing distortion of its signal."

Technical corrections:

RC5: "the shortcut OMI is first used in line 20 but introduced later in line 22"

AC5: The introducing the shortcut OMI has been moved to where it was first mentioned.

RC6: "Shortcut CCD is first used in line 46 but introduced later in line 66"

AC6: The introducing the shortcut CCD has been moved to where it was first mentioned.

RC7: "line 94 "orbit orbit...13:45.." the word is written twice and there is a double dot at the end of the sentence"

AC7: All typos have been corrected.

Formatted

Detection of anomalies ~~non-linear effects~~ in thesatellite ~~spectra from : Application to~~ the Ozone Monitoring Instrument

Nick Gorkavyi[1], Zachary Fasnacht[1], David Haffner[1], Sergey Marchenko[1], Joanna Joiner[2], Alexander Vasilkov[1]

[1] Science Systems and Applications, Lanham, MD, USA
[2] National Aeronautics and Space Administration (NASA), Goddard Space Flight Center (GSFC), Greenbelt, MD, USA

Correspondence: Nick Gorkavyi (nick.gorkavyi@ssaihq.com)

Abstract. Various instrumental or geophysical ~~Non-linear~~artifacts~~effects~~, such as ~~from~~ saturation, stray light, or obstruction of light (either coming from the instrument or related to solar eclipses), negatively impact satellite measured ultraviolet and visible Earthshine radiance spectra and downstream retrievals of atmospheric and surface properties derived from these spectra. In addition, excessive noise such as from cosmic ray impacts, prevalent within the South Atlantic Anomaly, can also degrade satellite radiance measurements. Saturation specifically pertains to observations of very bright surfaces such as sun glint over open water ~~surfaces~~ or thick clouds. When saturation occurs, additional photoelectric charge generated at the saturated pixel may overflow to pixels ~~Related residual electronic cross-talk or blooming effects may occur in spatial pixels~~ adjacent to a saturated area and be reflected as a distorted image in the final sensor output. ~~Obstruction of light can occur within the zones of solar eclipses as well as from material located outside of the satellite instrument. The latter may also produce unintended scattered light into a satellite instrument.~~ When these effects cannot be corrected to an acceptable level for science quality retrievals, ~~it is desirable to~~ flagging of the affected pixels is indicated. Here, we introduce a straightforward ~~new~~ detection method that is based on the correlation, *r*, between the observed Earthshine radiance and solar irradiance spectra over a 10 nm-spectral range; our Decorrelation Index (DI for brevity) is simply defined as DI=1-*r*. DI increases with anomalous additive~~non-linear~~ effects or excessive noise in either radiances, ~~(~~the most likely cause in data from the Ozone Monitoring Instrument ~~(or~~ OMI~~, data)~~ or irradiances. DI is relatively straight-forward to use and interpret and can be computed for different wavelength intervals. We developed a set of DIs for two spectral channels of the ~~Ozone Monitoring Instrument (~~OMI~~)~~, a hyperspectral pushbroom imaging spectrometer. For each OMI spatial measurement, we define 14 wavelength-dependent DIs within the OMI visible channel (350-498 nm) and 6 DIs in its ultraviolet 2 (UV2) channel (310-370 nm). As defined, DIs reflect a continuous range of deviations of observed spectra from the reference irradiance spectrum that are complementary to the binary Saturation Possibility Warning (SPW) flags currently provided for each individual spectral/spatial pixel~~s~~ in the OMI radiance data set. Smaller values of DI are also caused by a number of geophysical factors; this allows one to obtain interesting physical results on the global distribution of spectral variations.

1 Introduction

The Ozone Monitoring Instrument (OMI) is a Dutch/Finnish ultraviolet (UV) and visible (Vis) wavelength spectrometer that is on board NASA's Aura satellite launched on July 15, 2004. It has provided one to two day global coverage for several important atmospheric trace gases including ozone (O_3), ~~sulfur~~ sulphur dioxide (SO_2), nitrogen dioxide (NO_2), and formaldehyde (HCHO) as well as information about clouds and aerosols (~~Schenkeveld et al., 2017;~~ Levelt et al., 2018). OMI has contributed to studies of atmospheric pollution, climate-related agents, and stratospheric chemistry (Levelt et al., 2018,), lead to the first observation of glyoxal ($C_2H_2O_2$) from space (Chan Miller et al., 2014), and provided precise long-term records ~~observations~~ of solar spectral irradiances (Marchenko and Deland, 2014). OMI's ~~near real-time (NRT)~~ data have ~~are disseminated within 3 hours of sensing (Krotkov et al., 2015), while the very fast delivery (VFD) data are made available within 20 minutes of overpass,~~ contribut~~ed~~ing to medium-range weather and air quality forecasts, as well as to detection and tracking of volcanic plumes (Hassinen et al., 2008; Krotkov et al., 2015; Levelt et al., 2018). OMI measurements also

provide~~s~~ estimates of tropospheric ozone columns (e.g., Sellitto et al., 2011; Ziemke et al., 2017). Several similar sensors ~~which are similar to OMI~~ are currently in orbit, including the Tropospheric Monitoring Instrument (TROPOMI) on board the Copernicus Sentinel-5 precursor (S5P) satellite, Ozone Mapping and Profiler Suite/Nadir Mapper (OMPS/NM) on Suomi NPP and NOAA-20, and Global Ozone Monitoring Experiment 2 (GOME-2) instruments on European Organisation for the Exploitation of Meteorological Satellites~~–~~ (EUMETSAT) MetOp platforms.

Non-linear effects can impact the measured signal when not properly corrected. This~~, such as saturation and blooming,~~ can degrade Earthshine radiance measurements from passive solar backscatter UV/Vis satellite spectra and thus impact retrievals of atmospheric constituents and surface properties. There are several potential sources for these effects. Saturation occurs when bright light causes the number of electrons in a sensor pixel to exceed either the maximum charge capacity of an individual charge-coupled device (CCD) photodiode, or the maximum charge transfer capacity of the sensor. Blooming and other artifacts related to charge transfer on the CCD may also affect the quality of the measured spectrum when electrons from a saturated pixel overflow to a neighboring pixel, causing distortion of its signal ~~Once the excessive electrons from the saturated pixel flow into neighboring pixels, saturation leads to CCD blooming,~~ and frequently rendering affected~~such~~ data useless. Charge transfer and readout errors can also result in a distorted spectra, as can be the case with an error correction for detector smear. Hereafter we refer to the spatial domain of the two-dimensional CCD as rows (30 or 60 simultaneously acquired scenes) ~~as rows~~, and the spectral domain as columns. Per OMI design, ~~during~~ the CCD readout is in ~~the excessive charge spreads between~~ spatial pixels more easily, and therefore the blooming or charge readout-related effects are~~is~~ expected to predominantly occur~~observed~~ between different spatial rows.~~.~~

Retrievals of atmospheric gases or aerosols can be compromised when observing very bright surfaces such as sun glint in low wind speed conditions (Cox and Munk, 1964; Kay et al., 2009; Butz et al., 2013; Feng et al., 2016), as well as over scenes predominantly covered by optically thick clouds. ~~For example, the GOME2 (launched on 2007) UV band 1 detectors experienced saturation due to reflectance from clouds; this effect was predicted in the GOME-2 error assessment study (Siddans et al., 2002).~~ Saturation caused by Sun glint routinely occurs in the visible imagery of the MODerate Resolution Imaging Spectroradiometer (MODIS) flying on NASA's Aqua and Terra satellites. MODIS data show a gradual increase of saturated data towards the red and NIR bands, reaching around 1500 pixels, or ~ 0.03% of pixels, in a granule at 869 nm (Singh and Shanmugam, 2014). The Orbiting Carbon Observatory-2 (OCO-2) and similar greenhouse gas monitoring instruments occasionally point directly at the sun-glint. The OCO-2 in-orbit checkout activities revealed an unexpectedly high signal from Lake Maracaibo, Venezuela on August 7, 2014. This signal saturated all 3 channels and was attributed to an oil slick on a wave-free lake. After this event, known as the Lake Maracaibo Saturation Incident, an automated saturation warning algorithm was incorporated into the OCO-2 processing to identify such events (Crisp et al., 2017). Solar glint from ocean and clouds, as well as "saturation tails" or blooming effects are~~—also~~ seen in many images from the Earth Polychromatic Imaging Camera on the Deep Space Climate Observatory (EPIC/DSCOVR) (Varnai et al., 2019). ~~A~~

~~blooming effect occurs when electrons from a highly illuminated pixel of the charge-coupled device (CCD) matrix jump to a neighboring pixel, causing distortion of its signal.~~ TROPOMI also experiences detector saturation and blooming problems, typically caused by bright tropical clouds seen in bands 4 (400-499 nm) and 6 (725-786 nm). Bands 7 (2300-2343 nm) and 8 (2342-2389 nm) mostly react negatively to sun glint. Currently, blooming areas are not detected by the TROPOMI L0-1b processor. A flagging algorithm is under development (Rozemeijer and Kleipool, 2019; Ludewig et al., 2019).

A set of 16 operational flags, called the Saturation_Possibility_Warning (SPW) flags are currently included in the OMI level 1b data set. SPWs are designed to flag OMI pixels with 16 various radiation anomalies (e.g., saturation, stray light, nonlinearity). These flags are defined for each OMI wavelength: 751 wavelengths of the Vis spectrum and 557 wavelengths of the UV2 spectrum (GDPS, 2006). All of the 16 SPW flags are binary; a pixel with any degree of abnormality (e.g., saturation) at a given wavelength is marked as possibly bad.

Here, we describe a new approach to identify potentially erroneous OMI data based on the correlation *r* between the observed back-scattered Earthshine spectrum and a reference solar spectrum computed over limited spectral regions. Earthshine spectra differ from the solar spectra due to Rayleigh, rotational-Raman, aerosol and surface scattering as well as absorption of radiation by ozone and other atmospheric components. Most of these factors, with the exception of strong ozone absorption in the UV, amount to secondary effects on the correlation coefficient between the solar and Earthshine spectra within a limited spectral window. Under normal conditions (lack of detectable instrument-imposed spectral distortions) and for a reasonably narrow (5-10 nm, for practical purposes, with a moderate-resolution spectral instrument) spectral window, ~~t~~The degree of correlation ~~under normal conditions will~~ depend~~s~~ mainly (but not exclusively) on the number and strength (depth) of solar Fraunhofer features, once we take into consideration additional factors (differences in spectral resolution, finite signal-to-noise of measurements, mis-alignment of the wavelength grids, among others) that tend to degrade~~change~~ the correlation. In the windows with well-defined solar absorption spectral features, the correlation coefficient may gradually approach unity for the scenes acquired with S/N>>100 – a~~the~~ condition met in a majority of OMI UV2-Vis reflectance spectra. ~~and is typically close to unity. In this case, t~~Assuming the radiances and irradiances have the same spectral resolution, comparable S/N, and are closely co-aligned in the wavelength domain, the correlation coefficient between the earthshine and the solar 'etalon' (assumed to be distortion-free) ~~coefficient~~ should be ~~is~~ highly sensitive to any distortions ~~additional systematic additive deviation in the Earthshine spectra~~ in the former, leading to ~~and~~ rapidly decreasing~~es~~ correlation in saturated scenes (solar glint or bright clouds) or under other~~other~~ anomalous conditions, such as cosmic ray hits on the detector.

We apply our approach to OMI data and analyze individual cases and global distributions of flagged data. While these effects have been known for some time and dealt with, to some extent ~~by instrument teams by various methods~~, the prevalence of the different effects globally for a particular instrument has rarely been documented. This work provides a

detailed analysis of ~~non-linear~~ spectrum-distorting effects in the ~~specifically as they affect~~ OMI case, as well as a general and straight-forward approach that may be applied to similar instruments (TROPOMI, OMPS, GOME-2, etc.) to identify and filter out suspect or erroneous data.

2 Data and Methods

2.1 The Ozone Monitoring Instrument ~~(OMI) description~~

The Aura satellite that hosts OMI is in a polar Sun-synchronous orbit ~~orbit~~ with a local equator crossing time of 13:45.~~.~~ OMI is a nadir-looking, push-broom UV/Vis grating spectrometer (Levelt et al., 2018). The light entering the telescope is depolarized using a scrambler and then split into two channels: the UV (wavelength range 264–383 nm) and the Vis (wavelength range 349–504 nm~~:) (~~Dobber et al., 2006; Schenkeveld et al., 2017). The UV channel is further divided into the two sub-channels, UV1 (264-311 nm, 0.63 nm resolution and 0.21 nm sampling) and UV2 (307-383 nm range, 0.42 nm resolution with 0.14 nm sampling). Measurements are collected on two-dimensional charge-coupled device (CCD) sensors used for the UV and Vis channels. Spectral information is dispersed along one dimension of each CCD and spatial is imaged on the other. Each channel has a devoted frame-transfer CCD detector with 6e5 electrons/pixel full-well capacity. To avoid blooming and ellipsoid effects, the pixel filling should be kept below 3e5 electrons (Dobber et al., 2006). OMI also measures the solar irradiance once per day through the solar port. Here, we use the UV2 sub-channel and Vis channel only; in the UV1 channel, strong, variable ozone absorption renders our approach impractical.

In the global mode, each orbit spans the pole-to-pole sunlit portion, typically comprising 1644 along-orbit exposures, referred to as iTimes hereafter. The 114° viewing angle of the telescope corresponds to a 2600 km wide swath on the Earth's surface and consists of 60 simultaneously acquired rows or ground pixels across the track. In this mode, the OMI pixel size is 13 × 24 km^2 at nadir. The in-flight performance of OMI is discussed in Schenkeveld et al. (2017). The radiometric degradation of the OMI radiances since launch ranges from ~2 % in the UV channel to ~0.5 % in the Vis channel, which is much lower than any similar sensor (Levelt et al., 2018). O~~The o~~ne major anomaly has occurred with ~~disadvantage of~~ OMI, ~~is~~ the so-called row anomaly (Schenkeveld et al., 2017); it~~, which~~ is presumably caused by a partial detachment of insulation material exterior to the instrument and produces a number of anomalous~~non-linear~~ effects on sun-normalized radiances. The row anomaly is discussed in detail in Section 3.4.

2.2 The Decorrelation Index (DI)

We introduce a new parameter, the decorrelation index (DI), and defined as $1 - r$, where r is the Pearson correlation coefficient:

$$DI = 1 - r = 1 - \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (1)$$

Formatted: Indent: Left: 1", First line: 0.5"

with $\bar{x}$ (same for $\bar{y}$)

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (2)$$

In (1) and (2) x_i and y_i are the individual sample points for radiance I and irradiance F_0, respectively. DI is derived for radiances and irradiances at each spectral region: for OMI, 14 regions of ~10 nm (n= 51 wavelengths for each spectral region) in ~~on~~ the Vis channel and 6 regions of ~10 nm (n = 69 wavelengths for each region) in ~~on the~~ UV2 ~~channel~~. For the standard solar spectrum or reference irradiance, we take an average of all solar spectra obtained by OMI in 2005. ~~We consider the atmospheric spectra of the UV2 and Vis channels separately.~~ Each earthshine ~~atmospheric~~ spectrum is re-gridded via linear interpolation to match the wavelengths of the averaged irradiance spectrum. An exact match between the radiance and irradiance spectral features gives DI = 0, whereas when the features in the radiance and irradiance spectra deviate, the DI approaches 1 to 2, where values greater than 1 indicate that irradiance and radiance spectra exhibit anti-correlation. Hence, cases of DI > 0 may indicate distortions of atmospheric spectra.

Evidently $DI = 0$ for the simple case of a perfect match with $I = const * F_0$; if $I = -const * F_0$, then $DI = 2$. If I and F_0 are completely unrelated, then $DI = 1$. Considering the 'smooth' (low-frequency) component of I and F_0, we expect them to be generally correlated in the spectral regions relatively free of major atmospheric absorptions (ozone in particular). The correlation would be inevitably diminished by the wavelength-dependent Rayleigh scattering and surface reflectivity.

Once a multitude of deep spectral lines is superimposed on a smooth envelope, DI will depend mainly on a match between the shape and position of these I and F_0 spectral transitions, with the correlation depending on the S/N of the tested radiances and irradiances, and even more so on slight (in OMI's case) wavelength mis-alignments between radiances and irradiances, with the steep line flanks magnifying the differences.

An additional de-correlating factor is brought forth by the omni-present atmospheric rotational Raman scattering (e.g., Joiner et al., 1995). Under the circumstances, one may never expect $DI = 0$ save the exceedingly rare cases of a perfect solar glint. It is known that Pearson's correlation coefficient is sensitive to outliers, thus simplifying detection of spectral distortions in the high-resolution data compared to the low-resolution cases, with the latter tending to lessen the impact of additive components (the shallower lines are potentially less susceptible to stray light), as well as the wavelength mis-alignment (spectral blending of multiple features leading to partial canceling of distortions in the adjacent features). At a given spectral resolution and S/N, DI sensitivity may grow with increasing numbers and contrasts (depths) of spectral features in the chosen spectral window. At the same time, the DI is expected to be sensitive to artifacts associated with cosmic ray hits. The interval 440-480 nm, where there are few deep spectral lines, should be especially sensitive to geophysical factors, for

example, to the wavelength-dependent albedo of the earth's surface. Note that there are cases when direct solar radiation F_0 is mixed with I due to instrument problems (see below). Under this specific circumstance DI will decrease, since correlation between $(I + \delta F_0)$ and F_0 is always higher (thus DI lower) than between I and F_0. A similar effect occurs with sunglint from the water surface, when the proportion of directly reflected sunlight in I increases significantly. Note that in the current approach we do not compensate for the relatively smooth spectral differences imposed by atmospheric (Rayleigh scattering) and surface (wavelength-dependent albedo) factors, leaving this to the next DI version. This step would make DI more sensitive to the instrument-imposed anomalies, further disentangling those from the geophysical factors (see below).

In this initial version of the OMI DI, we use the spectral range 309.9-370.0 nm for UV2 and 349.9-498.4 nm for Vis. Overlapping of these ranges is useful for assessing the calibration between the UV2 and VIS channels. For solar zenith angles (SZA) > 90^o, the radiance level drops, noise begins to dominate, and the DI grows rapidly. Therefore, we avoid ~~for the cases of~~ SZA > 90 ~~^{o},~~ cases ~~we do not compute the DI~~. The DI is sensitive to the degree of distortion of the reflectance spectrum, regardless of the cause of the distortion (saturation, crosstalk, noise etc), so that it detects distortions other than saturation. For example, the DI may detect electronic cross-talk (or blooming) effects in pixels adjacent to the saturated area. In a number of cases, the DI ~~decorrelation index~~ proves to be either more or less sensitive than the current SPW (Saturation_Possibility_Warning) flags reported in the OMI PixelQualityFlags filed of the Level 1b data, as shown in the next section.

The DI provides a range of values that describes the deviation of observed spectra from the reference irradiance spectrum, while the SPW flag is a binary value. The DI therefore allows flexibility in setting detection thresholds for damaged spectra for different applications. The DI value for a given spectral interval depends strongly on the number of Fraunhofer lines as well as presence of strong ozone absorption features within the wavelength range. Therefore, the DI values corresponding to likely damaged spectra vary somewhat for each spectral region. For example, the 14 DI divisions of the Vis spectrum generally fall into two distinct groups; for the first group, the value of DI above 0.01-0.03 is a sign of a significant distortion of the spectrum, while for the second group a typical distortion threshold value is larger (~0.1-0.4). The provisional (the user may redefine the values using the auxiliary data provided in the OMI DI product) DI thresholds were determined as follows. We used all available, mission-long OMI UV2 and Vis radiances. For each orbit and for every spectral window we constructed DI histograms. Then we selected numerous cases sampling the tails of the DI histograms. On a case-by-case basis, for different scenes and spectral windows, we found empirically the lowest-DI values that repeatedly separate the scenes with apparently normal (spectrally smooth, with the fine-structure, low-amplitude Raman-scattering features) and distorted reflectances. These DI thresholds approximately correspond to 99.995-99.998 percentiles in the DI distributions. We plan to provide a statistically rigorous threshold definition in the improved DI version.

Formatted: Font: Times New Roman, 10 pt

Table 1 summarizes the DI wavelength bands and suggested threshold~~critical~~ values corresponding to damaged spectra. These critical values should be treated as indicative. A user may define different thresholds, depending on their application. We chose row 20 to determine these critical values.

Table 1. Chosen OMI DI spectral intervals and indicative DI thresholds for damaged spectra.

Interval (UV2)	**Wavelengths (nm) (for row 20)**	**Value DI as signature of distortion of spectra**	**Comments**
1	309.94-320.61	- *	Strong ozone effects
2	320.76-331.08	>0.20-0.25	Ozone effects
3	331.23-341.24	>0.35-0.45	Weak ozone effects
4	341.39-351.11	>0.02-0.03	Strong spectral lines
5	351.25-360.70	>0.02	Strong spectral lines
6	360.84-370.02	>0.01	Strong spectral lines
(Vis)			
1	349.93-360.33	>0.03	Strong spectral lines
2	360.54-370.93	>0.01	Strong spectral lines
3	371.14-381.52	>0.02	Strong spectral lines
4	381.73-392.11	>0.01	Strong spectral lines
5	392.32-402.70	>0.01	Strong spectral lines
6	402.91-413.29	>0.06-0.08	
7	413.50-423.89	>0.1-0.15	
8	424.10-434.50	>0.02-0.03	Strong spectral lines
9	434.71-445.12	>0.05-0.1	
10	445.32-455.74	>0.25	
11	455.95-466.39	>0.4	
12	466.60-477.05	>0.4	
13	477.26-487.72	>0.03	Strong spectral lines
14	487.93-498.41	>0.2	

*The threshold depends on the row number.

The dependence of the threshold DI values on cross-track position is relatively minor, except for the first UV2 interval. For this interval, other cross-track position may carry different values, primarily due to ozone absorption (increasing towards the swath edges). The ~~critical~~ DI threshold~~value~~s depend on spectral resolution and S/N of the reflectances, hence the indicative values~~thresholds~~ from Table 1 may vary for different instruments.

Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt, Not Bold
Formatted: Font: 10 pt
Formatted: Font: 10 pt, Not Bold
Formatted: Font: 10 pt
Formatted: Font: 10 pt, Not Bold
Formatted: Font: 10 pt

3 Results

To study the DI, we first concentrate on scenes that are most likely to contain saturation ~~and blooming~~ effects: sun glint areas with relatively calm water surfaces and contiguous bands of deep convective clouds. Next, we examine the global DI distribution ~~of the DI~~, which reveals other effects that damage observed spectra. We then investigate the impact of the row anomaly on the DI.

3.1 Saturation over clouds

A typical problematic cluster of bright clouds in the Pacific Ocean is shown in Fig.1a, where two zones are highlighted, a small northern zone (denoted A) and a large southern zone (marked as B). Figure 1c shows the number of wavelengths for a given pixel marked with the SPW flag as saturated. Figure 1b -1d shows the corresponding DI values for the Vis interval 414-424 nm. The DI indicates that the spectra in zone A are weakly affected, and in zone B they are badly damaged. Figure 1c shows the number of wavelengths for a given pixel marked with the SPW flag as saturated.

Figures 2 and 3 illustrate the properties of the DI that characterize the quality of a given part of the spectrum using a single parameter. Figure 2 shows an example of a spectrum with slight distortions that are captured by the SPW flags, but nevertheless has low values of the DI. Small deviations of the DI from 0 can result from geophysical effects, for example, an increased amount of ozone, and minor damage to the spectrum, ~~–~~as shown in Fig. 2. Those users who have strict requirements for the quality of the spectra should use the SPW flag in this case, which detects minor damage to the spectrum. Figure 3 shows the Vis spectrum for a pixel in zone B (indicated by an arrow in Fig. 1) corresponding to iTimes = 807, Row = 20~~)~~. The radiance spectrum is saturated in the 400-465 nm range. In contrast with Fig. 2, damage in this spectrum is manifested in both the SPW flag and the DIs. The DIs reflects the degree of spectral damage, which in this case reaches a maximum near 450 nm. Based on the problem under study, a user can determine whether the spectrum is useful despite minor damage such as in zone A. In such cases, the SPW and DI may provide complementary information.

Reflectance on Figures 2 and 3 is defined as $\pi \cdot I/[F_0 \cdot \cos(\theta)]$, where I is the top-of-the-atmosphere (TOA) radiance, F_0 is the extraterrestrial solar flux, θ is the solar zenith angle (SZA). Usually, t~~T~~he wavelength dependence of ~~a~~ TOA reflectance ~~in a usual situation~~ is ~~has the form of a~~ fairly smooth ~~curve~~, albeit with relatively low-amplitude, ~~small amount of~~ high ~~–~~frequency structures due to rotational Raman scattering, also known as the Ring effect. ~~(e.g., Joiner et al., 1995).~~ Both zones in Fig.1 have high values of reflectance; for zone A, reflectance is between 0.95 and 1.0 (Fig. 2), while for zone B, reflectance is between 1.0 and 1.1 (Fig. 3). In some viewing directions the reflectance can exceed unity due to anisotropic angular distribution of the TOA radiance.

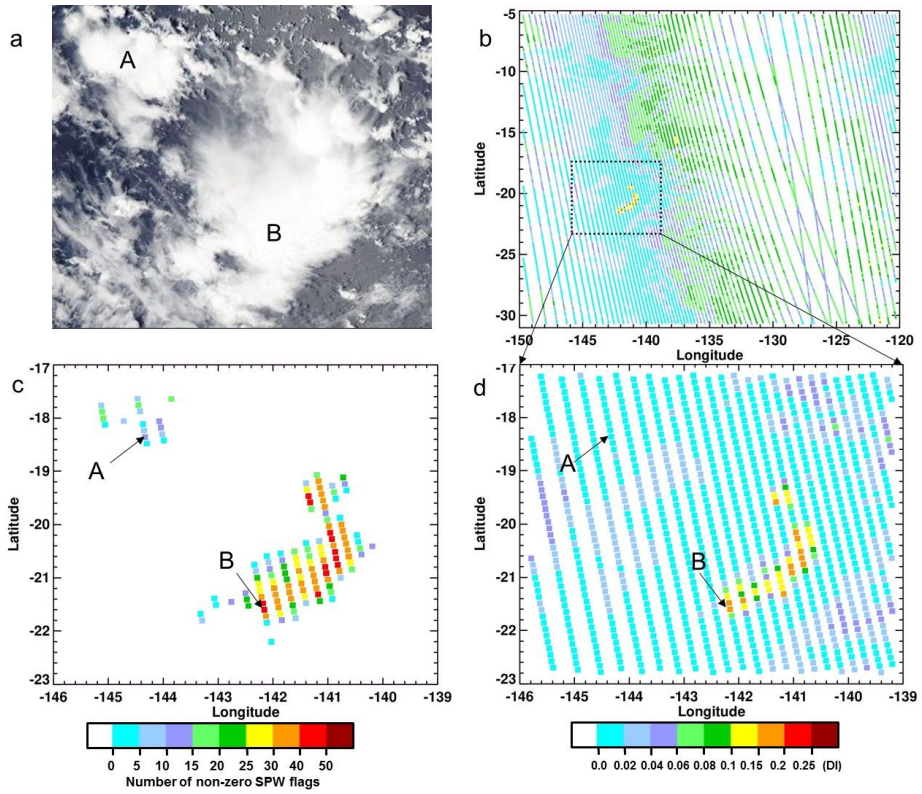

Figure 1: a. Two cloud zones in the south Pacific on January 14, 2006 for orbit 7990: small northern zone labeled "A" and large southern labeled "B" (a) Aqua MODIS image; (b), (d) DI maps for the Vis spectral region 414-424 nm; (c) The number of wavelengths for a given pixel marked with the SPW (Saturation Possibility Warning) flag as saturated (the maximum number is 51 in this Vis spectral region).

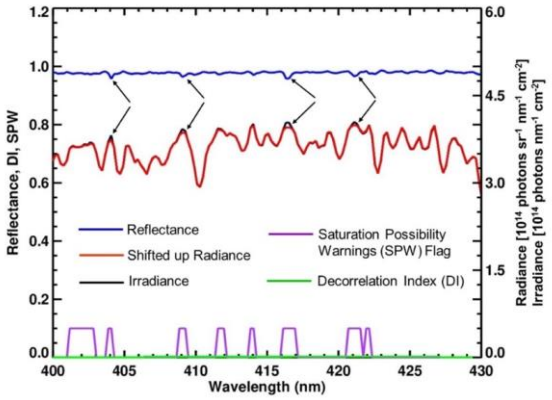

Figure 2: Data for iTimes 839, row 15, orbit 7990, January 14, 2006 in zone A (this pixel is marked by arrows in Fig. 1(c,d). ~~The red and black lines are the radiance and irradiance, respectively (the magnitude of the irradiances is shifted to line up with the radiance. Radiance is reported in photons sr^{-1} nm^{-1} cm^{-2}, irradiance – in photons nm^{-1} cm^{-2}. Both radiance and irradiance are divided by 10^{14}.~~ The blue line at the top of the picture is reflectance $\pi I/[F_0 \cos(\theta)]$, where θ is solar zenith angle. Reflectance in this zone has slight variations caused by minor saturation in the atmospheric spectrum as indicated by the arrows. The purple line shows the binary SPW (Saturation Possibility Warning) flags multiplied by 0.1. The green line is the DI<0.01 for bands 403-413, 413-424, 424-434 nm. The intensity of the radiance spectrum is shifted upwards slightly for clearer comparison with the irradiance spectrum.

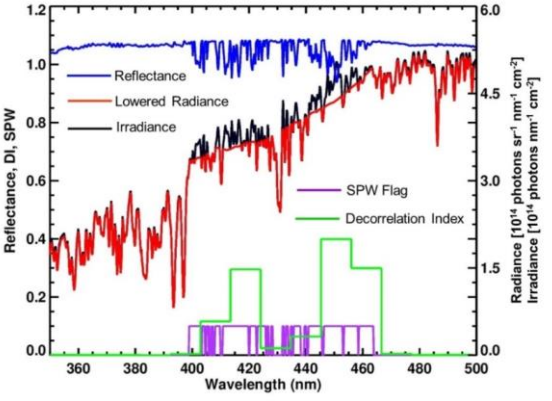

Figure 3: Similar to Fig. 2 but for iTimes - 807, Row - 20, orbit 7990, January 14, 2006 in cloud region B (see arrows in Fig.1 (c,d). The radiance was lowered by a few percent for better comparison.

3.2 Saturation over lakes and oceans

The South American lake Salar de Uyuni is used for calibration of many satellite sensors (Lamparelli et al., 2003; Fricker et al., 2005). Salar de Uyuni is dry for most months of the year, but during the rainy season, it is filled with shallow water with strong direct reflectance from the sun. This may cause saturation of OMI's detectors. The lake, covered with shallow water, generated strong solar glint, for example, for ~~the~~ ~~on~~ orbit 7987 (January 14, 2006). Figure 4a shows this shallow lake on January 14, 2006 as observed by the Aqua MODIS sensor. The SPW flags (Fig. 4c) and DIs (Fig. 4b,d) for this case show that the lake generates two bright spots: southern and northern. The solar glint from the northern spot is so bright that the signal extends to nearby pixels (iTimes 823-825, Rows ~~#~~ 11-14)~~. The resulting spectral distortions, called (blooming), are caused by artificially increased radiance~~ and are detected by the DI (see also Cao et al., 2019; Shen et al., 2019 for examples of blooming in other sensors). The SPW flags is unset for s~~S~~ignificant portions of the affected OMI pixels ~~that underwent blooming effects have zero values for the SPW flag; the SPW flag typically marks only the saturated pixels~~ (see Fig. 4).

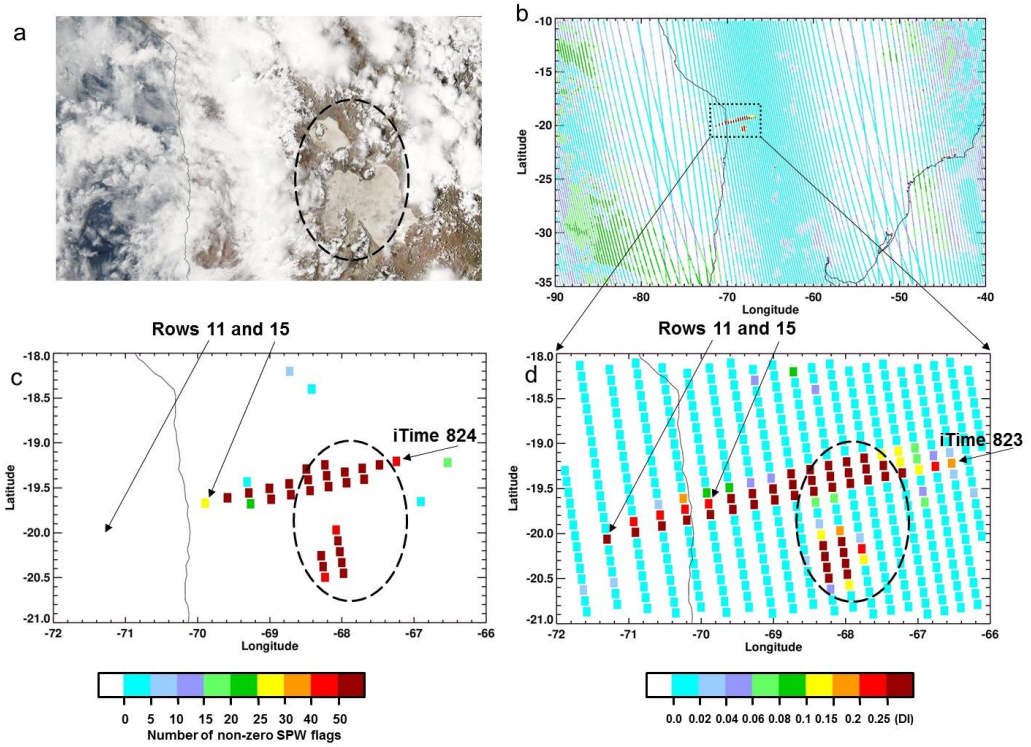

Figure 4: Similar to Fig. 1 but for an area near Salar de Uyuni, January 14, 2006, orbit 7987.

Figure 5 shows the Vis spectrum for a pixel where on the edge of the highly saturated region. This pixel may have ~~a strong distortion due to a complex~~ superimposed effects~~position~~ of moderate saturation of the pixel itself as well as charge overflow from due to significant saturation in the neighboring pixels~~ and blooming, effect produced amplification in strongly distortinged the radiance spectrum~~ (iTimes - 824, Row - 15). While ~~Most of~~ the reflectance values of many of these ~~for the blooming affected ese~~ pixels (rows ~11-15 in Figure 4) are in the expected ~~normal~~ range (0.3-0.6) there are numerous cases where~~. However, blooming related perturbations, in the form of peaks in the reflectance due to blooming, often boost reflectances~~ the final radiance signal is well beyond normal~~these~~ range~~s~~, thus leading to high DIs. ~~exceed unity. The DI highlights the affected parts of the spectrum.~~

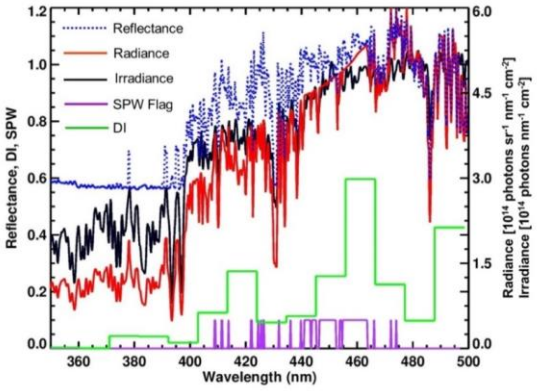

Figure 5: Similar to Fig. 2 but for pixel iTimes 824, row 15, orbit 7987 indicated by arrows in Fig. 4(c,d) showing solar glint from Lake Salar de Uyuni.

Figure 6 shows the spectrum of ~~for~~ a pixel where~~ith here~~ the prevailing ~~spectral distortion due to the~~ anomalous~~blooming~~ effect~~s~~ do not appear to be direct saturation~~causes peaks in the radiance and reflectance~~ (iTimes: 824, Row: 11). High DI values are seen for a number of corrupted parts of the spectrum where the SPW flags are zero. While the saturated case is straightforward to detect and interpret (e.g., the practically featureless radiances in Figures 3 and 4), these other ~~'blooming'~~ spectral distortions may have a complex spectral envelope due to the differences in the wavelength sampling of the sequential OMI rows. The completely saturated spectral domains may trigger effects in the neighboring rows, indiscriminately affecting the involved wavelengths. However, in~~f~~ the case of a less severely saturated scene, there might be

additional effects to consider.: Pper instrument design, the OMI wavelength grids are form aing 'spectral smile' in the row-wise direction. Inspecting the wavelength registration for a given CCD column (the spectral domain) while moving from the swath's edges towards nadir, one may notice gradual wavelength shifts between the adjacent rows. The wavelengths are increasing while moving from the edges to the center of the swath, thus forming a 'smile'. This may result in occasional augmented distortions around narrow, well-defined features in the spectral image in non-saturated pixels, while the signal for other wavelengths in the spectrum may remaining intact. Such occasional distortions could be mimicked and greatly outnumbered by a different effect that also stems from the 'spectral smile'. In some cases, brightly lit (but not saturated) scenes border low-reflectance areas: e.g., the studied Salar de Uyuni case, the cloud-front edges, or the edges of extended fresh snow/ice fields. In these bordering low-reflectance scenes the greatly augmented spatial stray light could mimic aemulate the blooming effect caused by the spectral smile, thus leading to higher DI around strong, deep spectral transitions that may exceed the imposed threshold. In the OMI data sampling the high-contrast scenes, the spatial stray light effects induce wavelength shifts that affect trace-gas retrievals (Richter et al., 2020). Some of the above-threshold DIs in the global maps (mid-to-high latitudes, open-water scenes - see below) could be triggered by the high-contrast scenario.

Users who make special demands on the quality of the spectra can use spectra whose quality is confirmed by both parameters.

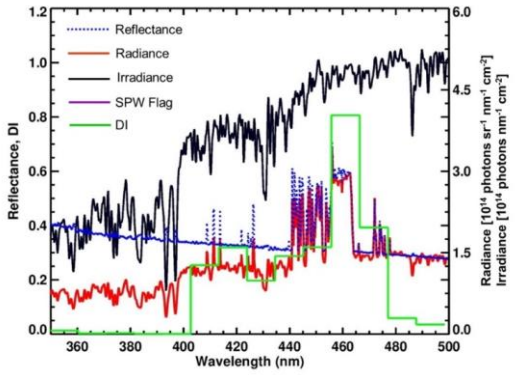

Figure 6: Similar to Fig. 5 but for pixel iTimes 823, row 11, orbit 7987 (see arrows in Fig. 4d). SPW flags are zero for this case and are not shown.

An example of solar glint in the Caribbean Sea is shown in Fig. 7 for July 26, 2013. Effects of the glint for this case are detected in both the DI and SPW flags. Some of the pixels not marked by the SPW flags show high DI values that may be

~~are most likely~~ related to blooming or other effects associated with the impaired performance of neighboring pixels on the detector.

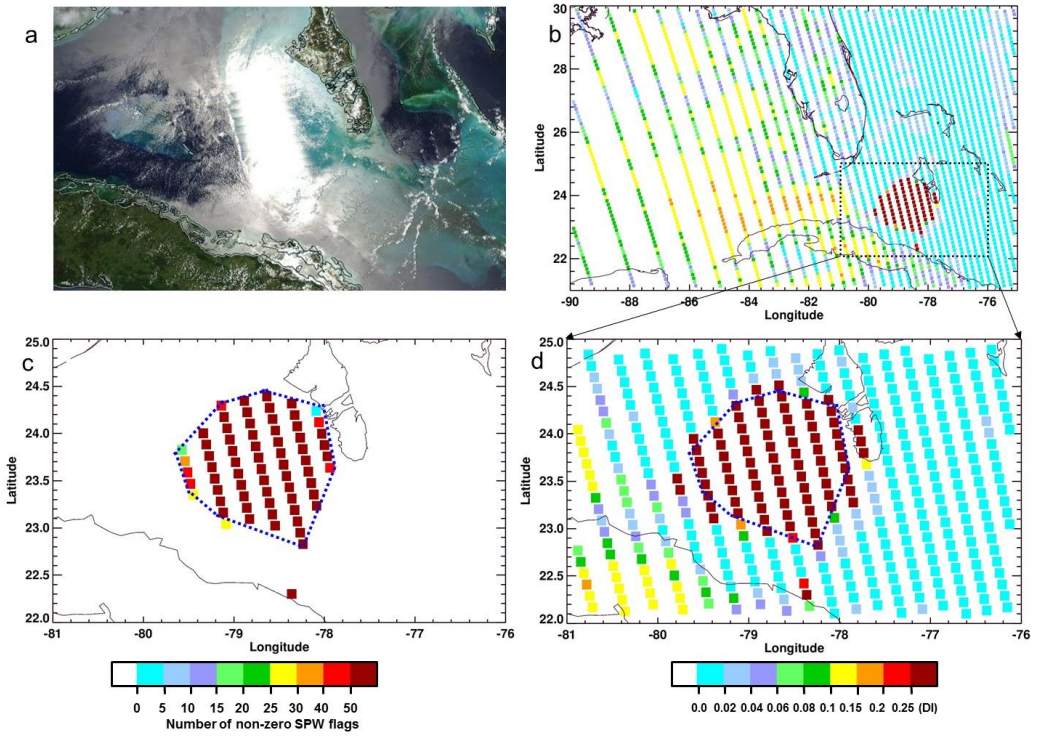

Figure 7: Similar to Fig. 1 but for an area showing solar glint near the Bahamas (orbit 48034, July 26, 2013). Glint positions in (c) and (d) do not exactly match those in the RGB image (a). **The area of pixels marked by the SPW flags are approximately delineated by the blue dotted line.**

3.3 Orbital and Global Distribution

Figures 8-10 show global distributions of the number of affected Vis spectra with DI > thresholds for spectra for March 2006 (Figs. 8 and 9) and for the entire 2006 (Fig. 10). The global DI behavior ~~distributions of DI show that it~~ depends on many

geophysical and instrumental processes. For example, Fig. 8a shows the spatial distribution of the number of spectra with ~~where~~ DI > 0.03 for the 424.1-434.5 nm range. Figure 8b similarly shows distributions for ~~where~~ DI > 0.25 in ~~for~~ the 445.3-455.7 nm spectral window. Despite the different threshold DI values that characterize the distorted spectra, these two DIs show similar ~~approximately the same~~ distributions of corrupted spectra associated with enhanced cosmic ray hits on the detectors within the South Atlantic Anomaly (SAA) region, glint, and a solar eclipse zone. Though we cannot disentangle the contributing factors for the latter, here we mention two of them as the likely causes of the high DI values (thus enhanced distortions in the reflectances): the low S/N of the eclipsed radiances, as well as the drastically increased portion (compared to the normally-lit scenes) of the additive (straylight) component.

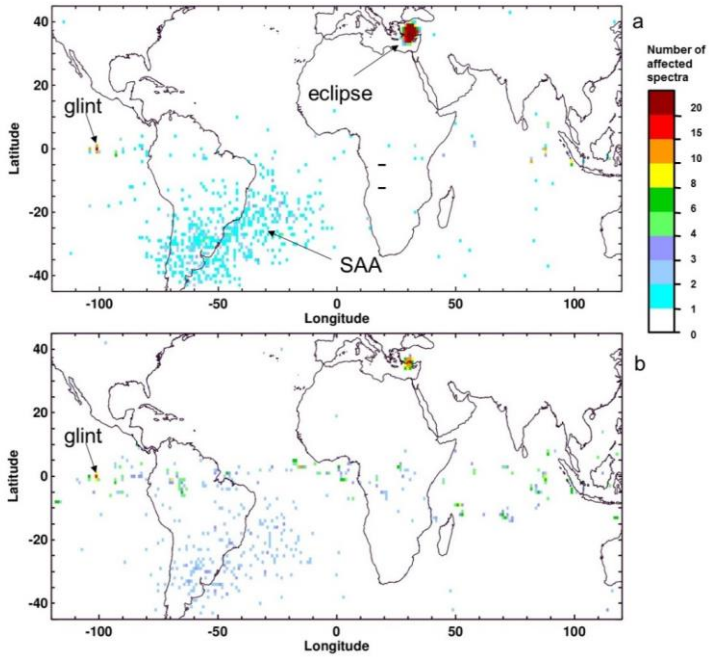

Figure 8: Gridded (1^{o}x1^{o}) distributions of the number of affected spectra for March 2006; (a) Vis band 424.1-434.5 nm, DI > 0.03; (b) Vis 445.3-455.7 nm, DI > 0.25. The South Atlantic Anomaly (SAA) and a region affected by solar eclipse are clearly visible; the remaining pixels with high DI values are mostly associated with sun glints and bright clouds.

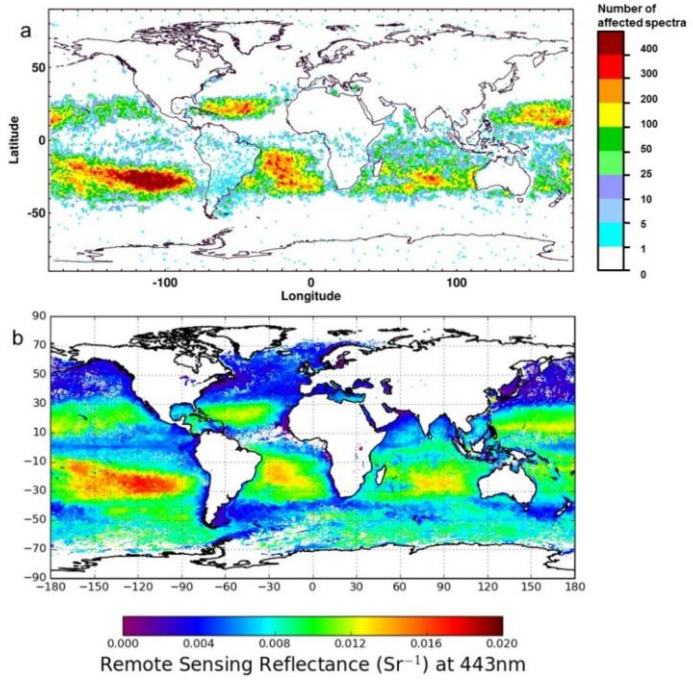

Figure 9: (a) Gridded (1°x1°) distribution of a number of spectra with DI>0.1 for Vis 445.3-455.7 nm in March 2006; (b) ocean remote sensing reflectance for March 2006 at 443 nm from Aqua MODIS.

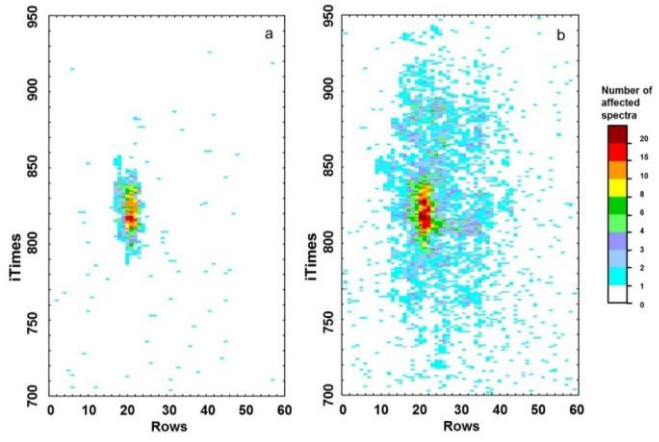

Figure 10: Distribution of a number of affected spectra for 2006, Vis 445.3-455.7 nm; (a) DI > 0.6; (b) DI>0.25. Usually each pixel collects ~ 5000 spectra per year.

The interpretation of low DI for normal spectra (for example, spectra with DI > 0.1 for Vis 445.3-455.7 nm) is quite complicated as low DI values depend on many factors. Figure 9 shows the spatial distribution of the number of spectra with DI > 0.1 in the 445.3-455.7 nm region compared with ocean reflectance at 443 nm. There is obvious spatial correlation between the spectra the DI identifies and ocean reflectance: larger numbers of such spectra correspond to ocean areas with higher reflectance. This is particularly pronounced in the southern Pacific Gyre whose waters exhibit extremely low bio-productivity and thus are very bright in the blue region (Tedetti et al., 2007). The strong spectral dependence of water-leaving reflectance in the blue region in these extremely clear waters results in lower correlation with the solar spectrum. This may be attributed in part to vibrational Raman scattering that is prevalent in clear ocean waters (Vasilkov et al., 2002; Westberry et al., 2013). Additionally, the Pacific Gyre area is characterized by low cloudiness and low aerosol loadings. Therefore, in this area the relatively high proportion of the shown data comes from the surface, thus being more susceptible to the Rayleigh and Raman scattering effects. These change the TOA radiances in different ways: the low-frequency spectral 'envelope' is affectamended by Rayleigh, while the fine-scale structures are introduced by the vibrational and rotational Raman scattering. Both effects lead to 'distorted' reflectances, thus higher DIs. ~~in the atmosphere contributes significantly to the top-of- atmosphere radiance. The atmospheric rotational Raman scattering effect additionally lowers the correlation of the TOA radiance with the solar spectrum.~~

Figure 10 a,b shows the orbital distributions of the 445.3-455.7 nm DI for DI > 0.25 and DI > 0.6 , respectively, plotted for OMI detector rows (generally oriented east-west across the satellite track) versus iTimes (north-south orbital direction) for 2006. A block of 250 along-orbit exposures (iTimes) approximately covers 30° ~~degrees~~ in latitude. The middle of this band falls on the equator on March 22. During the year, this band shifts by 22.4° to both the north and south. The zone around row 21 and iTimes 820 is an area of solar glint from the ocean surface (case DI > 0.6, Fig.10b) that does not change with season. The distribution of bright clouds with DI > 0.25 also shows a strong propensity for the geometrical conditions of solar glint (Fig.10a). This is consistent with EPIC/DSCOVR data showing solar glint from clouds that contain oriented ice plates (Varnai et al., 2019). In the OMI case, the ~~These~~ strongly saturated (or damaged) spectra with DI > 0.60 number about 2500 (~0.0005%) or ~7 spectra/day. Slightly affected spectra (0.25 < DI < 0.6) occur at a rate of ~0.002% or ~33 spectra/day.

Formatted: Superscript

3.4 The Row Anomaly

The row anomaly (RA) renders a significant portion of the OMI rows as unusable. The anomaly was clearly detected in two rows in June 2007. In May 2008 the row anomaly spread to several other rows on the sensor. The row anomaly has continued to develop since then, with particularly swift changes around January 2009 and early fall of 2011. Currently about

33% of the UV2 rows are affected in the southern hemisphere parts of the OMI orbit. This increases to ~57 % in the northern hemisphere. These estimates are comparable in the Vis channels (Schenkeveld et al., 2017). Figure 11 ~~similarly~~ shows DI distributions in the ~~space~~ row versus~~s.~~ ~~by~~ iTimes format (traditionally used for RA tracking) ~~, but~~ for the overlapping region of the UV-2 and Vis detectors. ~~This shows that the row anomaly is well detected by the DI in both detectors though with more impact on the Vis detector.~~ The row anomaly is a complex phenomenon that may result in artificially low or high values of reflectances, ~~radiances~~ additionally affecting their wavelength dependence~~,~~. The RA stems from ~~depending on the~~an ~~-~~interplay of multiple ~~-two major~~ factors that may affect the DI values. The RA is likely linked to a gradual detachment of the thermal blanket partially blocking some FOVs (rows). Since this blanket is highly reflective, its warped surface causes occasional (solar-angle dependent, predominantly affecting northern portion of the OMI orbit) reflection of the direct sunlight into some RA-affected rows. In addition, the reflective surface leads to enhanced spatial cross-talk between adjacent RA-affected scenes (an anomalous stray light that is regulated by the wavelength- and angle-dependent reflectivity of the blanket). The time-, space- and wavelength-dependent combination of 3 factors may lead to increasing or decreasing DIs. Even more complications stem~~come~~ from the fact that the RA may increase inhomogeneity of the spectral-slit illumination, thus causing substantial (and unaccounted for) wavelength shifts and ensuing spectral distortions in the reflectances. ~~: the blocking of the incoming Earth shine and the solar light scattering.~~

Deciphering the complex RA-related patterns in Figure 11, we~~,~~ first of all~~,~~ relate them to the pre-RA epoch that shows increase in the above-threshold cases in the equatorial regions, with a pronounced minimum centered on the sunglint domain (rows 10-30 and iTimes~ 650-1000). At the same time, the numbers of the~~–~~ above-threshold cases diminish towards the higher latitudes and the OMI's swath edges. In the TOA reflectances coming from the sunglint areas, the higher proportion of the directly reflected sun light leads to a higher radiance-irradiance correlation, thus lower DIs. The latitudinal and swath-angle dependent trends can be linked to the gradually diminished influence of the surface that modulates the TOA reflectances due to the wavelength-dependent surface albedos. In the planned upgrade of the DI algorithm, we intend to address this component, thus decreasing the impact of geophysical factors.

Turning our attention to the RA-affected areas in Figure 11, we notice that DIs closely delineate the RA-affected areas and show pronounced north-south asymmetry, with intricate patterns of the relatively higher/lower DIs compared to the RA-free plots. The north-south asymmetry is caused by the well-documented northward growth (Schenkeveld et al. 2017) of the blanket-reflected direct-sunlight component in the RA-affected radiances. This inevitably lessens the corresponding DIs. The solar influence appears to be strongly cross-swath modulated. This is a new aspect that requires a detailed follow-up study that is beyond the scope of the paper. At the same time, the remainder of the RA-affected areas show significant increase in the above-threshold DIs. This likely comes from the RA-imposed and unaccounted for wavelength shifts.

~~The former generally increases DIis. At the same time, the additive component related to the reflected scattered (most likely, by the instrument's thermal blanket) solar light is known to predominantly affect the radiances in the northern part of OMI~~

orbit. Under some conditions (This leads to the significant decrease of DIis in the RA-affected areas seen in Fig. 11. Figure 11 shows that the number of spectra with DI > 0.01 generally increases from polar to equatorial regions. This may be explained by an increase in the contribution from the oceans that undergoes vibrational Raman scattering (see Fig. 9). But in the area of rows 10-30 and iTimes 650-1000, there is a region with lower DI values. This spot is clearly associated with the growth of bright sun glare from the sea and cloud that reduces the amount of Raman scattered light from the atmosphere and ocean and thus reduces the DI. We expect that the value of the DI in these regions depends on the average slopes and wave heights in a given month as well as on wind speed.

The lower DI counts in the sunglint areas in Figure 11 This behavior seemingly contrasts with the increased DI values at the center of these regions in the Vis (Figure. 10). One should note that Figures 10 and 11 sample different spectral domains, with the 445.3-455.7 nm range (Figure 10) known to be highly susceptible to saturation, contrary to the exceedingly rare incidence of saturation in the 349.9-360.3 nm band (Figure 11).

This is because an increase in the fraction of solar radiation decreases the DI, but when the solar radiation exceeds the limits of saturation, the spectra will begin to be damaged and a local zone with a high DI (see Fig. 10) will appear in the center of the zone with reduced DI.

Formatted

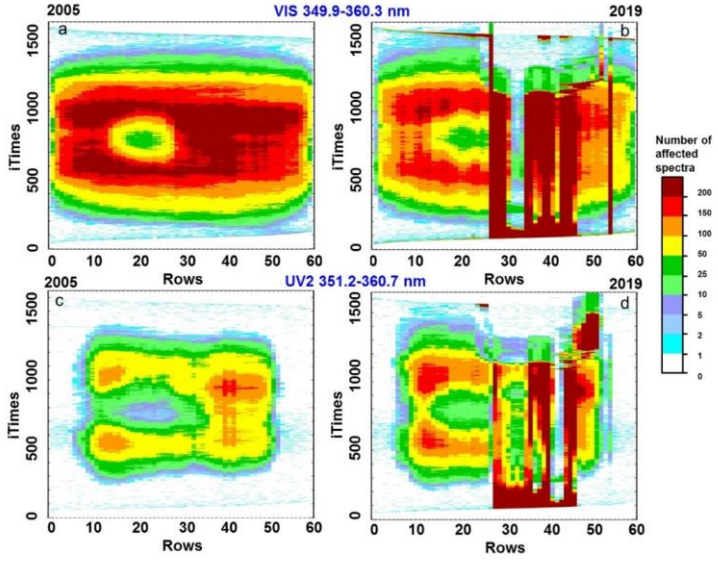

Figure 11: Distribution of the number of affected spectra (DI>0.01) for March 2005 (left) and 2019 (right) for the Vis 349.9-360.3 nm (top) and UV2 351.2-360.7 nm (bottom). The row anomaly is responsible for the stripes of high values shown in 2019.

4 Discussion and Conclusions

The OMI ~~We are convinced of the completeness of the~~ PixelQualityFlags (PQF) were designed to characterize each wavelength of the OMI spectrum (SPW flag is just one of the 16 bits in the PQF). The DI, developed on the basis of the correlations between observed and solar spectra, can serve as a simple but effective and complementary method for detecting and discarding anomalous UV and Vis satellite spectra, for example associated with detector saturation, blooming, charge transfer or readout, excessive noise, cases of very low reflectance~~ivity~~ such ~~(~~as ~~in~~ solar eclipse~~)~~, or the OMI row anomaly. The DI summarizes all changes in the spectrum in one parameter and eliminates the need ~~-~~to examine all the available flags for a given pixel. An important motivatione for introducing such an index is the convenience of handling it. For example, to infer enhanced information of the quality of spectra in the Vis region, we introduce 14 scalar-valued DIs for regions of the spectrum. For comparison, there are 751 binary saturation flags per spectrum in the level 1b. Similarly, we use 6 DIs for the UV2 spectrum; much less than the 577 flags assigned in the level 1b. Interpreting a large number of flags can be difficult. The DI product gives an indication of spectral quality based on overall correlation that is easier to interpret. Assessment of the DI the OMI Collection 3 L1b record has motivated improvements in detector corrections for the next version of the L1b product to be released in OMI Collection 4. The continuous nature of the DI allows data users to assign lower confidence to regions of the spectra that may not be completely saturated as detected by an electronic saturation algorithm. DI values vary for spectra that do not experience any anomalies. These variations of the DI may carry information that can be used for other purposes. For instance, the DI can be used to search for areas of clear ocean water, in which the spectra are not abnormal, but experience significant deviations from the solar spectrum due to geophysical reasons.

The DI can be a useful tool for analyzing spectra obtained from other current and future space-borne sensors that may suffer from saturation and blooming such as TROPOMI (launched in 2017) or the similar Environmental trace gases Monitoring Instrument (EMI) on the GaoFen-5 satellite (Chen, 2016) (launched in 2018). Similar sensors include the~~operate on~~ OCO-2 (launched in 2014)~~,~~ and OCO-3 (launched in 2019) (Eldering et al., 2019), ~~and~~ South Korea's g~~C~~eostationary Environment Monitoring Spectrometer (GEMS) (launched Feb. 18, 2020), ~~and will fly in the future on~~ NASA's geostationary Tropospheric Emissions: Monitoring of Pollution (TEMPO) (Zoogman et al., 2017) (planned for launch in 2022), ~~the~~ Copernicus geostationary Sentinel-4 (planned for launch in 2023)~~,~~ and low-Earth orbit Sentinel-5 (planned for launch in 2023). Many of these sensors have a smaller pixel size and/or smaller field of view (FOV) than OMI. For such instruments, this may lead to an increase in the effects of sun glint. Studies utilizing the DI with current instruments may benefit the design of future instruments by identifying how often and under what conditions spectra are impacted by non-linear effects.

Data availability

The Decorrelation Index data for OMI Collection 3 data will be available at NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). The OMI Level 1b data used for calculations of the DI are available at https://aura.gesdisc.eosdis.nasa.gov/data/Aura_OMI_Level1/. MODIS data are available at https://worldview.earthdata.nasa.gov/.

Formatted: Left, Line spacing: single

Formatted: Font: Times New Roman, 10 pt

Formatted: English (U.S.)

Competing interests

The authors declare that they have no conflicts of interest.

Authors contributions

NG developed ~~a~~ computer code~~s~~, analyzed the DI results, and wrote the manuscript. ZF supported the development and implementation of the algorithms and comparison DI results with the ocean reflectance. ~~(Fig.9).~~ DH proposed a concept of DI and wrote the manuscript. SM proposed a concept of DI and supported the development and implementation of the algorithms. JJ set the task of developing an effective method for determining solar glints, supported the development of the algorithm, and wrote the manuscript. ~~(chapters 3.1-3.4).~~

Financial support

This work was supported by the NASA Aura project (OMI core team) managed by Ken Jucks.

Acknowledgements

The authors thank the OMI and MODIS instrument teams for providing the OMI and MODIS data presented, respectively. We dedicate this work to the memory of Remco Braak, whose early work on saturation in OMI spectra helped to motivate this work.

References

Butz, A., Guerlet, S., Hasekamp, O. P., Kuze, A., and Suto, H.: Using ocean-glint scattered sunlight as a diagnostic tool for satellite remote sensing of greenhouse gases, Atmos. Meas. Tech., 6, 2509–2520, https://doi.org/10.5194/amt-6-2509-2013, 2013.

Cao, X., Hu, Y., Zhu, X., Shi, F., Zhuo, L., and Chen, J.: A simple self-adjusting model for correcting the blooming effects in DMSP-OLS nighttime light images, Rem. Sens. Environ., 224, 401-411, https://doi.org/10.1016/j.rse.2019.02.019, 2019.

Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D. J., and Chance, K.: Glyoxal retrieval from the Ozone Monitoring Instrument, Atmos. Meas. Tech., 7, 3891–3907, https://doi.org/10.5194/amt-7-3891-2014, 2014.

Chen, L.: Mission overview GaoFen-5. CEOS-ACC-12 meeting, Seoul, Korea, Oct. 13-15, 2016, http://ceos.org/document_management/Virtual_Constellations/ACC/Meetings/AC-VC-12/Day%201/5.%20Linagfu%20Chen%20-%20Gaofeng-5%201013.pdf, 2016

Cheng, L., Tao, J., Valks, P., Yu, Ch., Liu, S., Wang, Y., Xiong, X., Wang, Z., and Chen, L.: NO_2 retrieval from the Environmental trace gases Monitoring Instrument (EMI): preliminary results and intercomparison with OMI and TROPOMI, Rem. Sens., 11, 3017, https://doi.org/10.3390/rs11243017, 2019.

Cox, Ch. and Munk, W.: Measurement of the roughness of the sea surface from photographs of the Sun's glitter, J. Opt. Soc. Amer., 44, 838-850, https://doi.org/10.1364/JOSA.44.000838, 1954.

Crisp, D., Pollock, H. R., Rosenberg, R., Chapsky, L., Lee, R. A. M., Oyafuso, F. A., Frankenberg, C., O'Dell, C. W., Bruegge, C. J., Doran, G. B., Eldering, A., Fisher, B. M., Fu, D., Gunson, M. R., Mandrake, L., Osterman, G. B., Schwandner, F. M., Sun, K., Taylor, T. E., Wennberg, P. O., and Wunch, D.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, Atmos. Meas. Tech., 10, 59-81, https://doi.org/10.5194/amt-10-59-2017, 2017.

Dobber, M. R., Dirksen, R. J., Levelt, P. F., van den Oord, G. H. J., Voors, R. H. M., Kleipool, Q., Jaross, G., Kowalewski, M., Hilsenrath, E., Leppelmeier, G. W., de Vries, J., Dierssen, W., and Rozemeijer, N. C.: Ozone Monitoring Instrument calibration, IEEE Trans. Geosci. Rem. Sens., 44, 1209-1238, https://doi.org/10.1109/TGRS.2006.869987, 2006.

Eldering, A., Taylor, T.E., O'Dell, C.W., and Pavlick, R.: The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data, Atmos. Meas. Tech., 12, 2341-2370, https://doi.org/10.5194/amt-12-2341-2019, 2019.

Feng, L., Hu Ch., Barnes B. B., Mannino, A., Heidinger, A. K., Strabala, K., and Iraci, L. T.: Cloud and sun-glint statistics derived from GOES and MODIS observations over the Intra-Americas Sea for GEO-CAPE mission planning, J. Geophys. Res., 122, 1725-1745, https://doi.org/10.1002/2016JD025372, 2017.

Fricker, H. A., Borsa, A., Minster, B., Carabajal, C., Quinn, K., and Bills., B.: Assessment of ICESat performance at the Salar de Uyuni, Bolivia, Geophys. Res. Lett., 32, L21S06, https://doi.org/10.1029/2005GL023423, 2005.

GDPS Input/Output Data Specification (IODS). v. 2. Level 1B Output product and Metadata. SD-OMIE-7200-DS-467, 44-45, 25 August 2006, https://manualzz.com/doc/28490086/gdps-input-output-data-specification--iods--volume-2-leve..., 2006.

Hassinen, S., Tamminen, J., Tanskanen, A., Leppelmeier, G., Mälkki, A., Koskela, T., Karhu, J. M., Lakkala, K., Veefkind, P., Krotkov, N., and Aulamo, O.: Description and validation of the OMI very fast delivery products, J. Geophys. Res., 113, D16S35, https://doi.org/10.1029/2007JD008784, 2008.

Joiner, J., Bhartia, P. K., Cebula, R. P., Hilsenrath, E., McPeters, R. D., and Park, H.: Rotational Raman scattering (Ring effect) in satellite backscatter ultraviolet measurements, App. Opt., 34, 4513-4525, https://doi.org/10.1364/AO.34.004513,

Kay, S., Hedley, J. D., and Lavender, S.: Sun glint correction of high and low spatial resolution images of aquatic scenes: a review of methods for visible and near-Infrared wavelengths, Remote Sens., 1, 697-730, https://doi.org/10.3390/rs1040697, 2009.

Krotkov, N. A., Li, C., and Leonard, P.: OMI/Aura ~~sulfur~~ sulphur dioxide (SO_2) total column L3 1 day best pixel in 0.25 degree x 0.25 degree V3, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), https://doi.org/10.5067/Aura/OMI/DATA3008, 2015.

Lamparelli, R. A. C., Ponzoni, F. J., Zullo, J., Jr., Pellegrino G. Q., and Arnaud, Y.: Characterization of the Salar de Uyuni for in-orbit satellite calibration, IEEE Trans. Geosci. Rem. Sens., 41, 1461-1468, https://doi.org/10.1109/TGRS.2003.810713, 2003.

Levelt, P. F.: OMI instrument description and level 1B product, ATBD-OMI-01, August 1, 2002, https://eospso.nasa.gov/sites/default/files/atbd/ATBD-OMI-01.pdf, 2002.

Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan, B. N., Streets, D. G., Eskes, H., van der A, R., McLinden, C., Fioletov, V., Carn, S., de Laat, J., DeLand, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering, K., Apituley, A., González Abad, G., Arola, A., Boersma, F., Chan Miller, C., Chance, K., de Graaf, M., Hakkarainen, J., Hassinen, S., Ialongo, I., Kleipool, Q., Krotkov, N., Li, C., Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Tilstra, L. G., Torres, O., Wang, H., and Wargan, K.: The Ozone Monitoring Instrument: overview of 14 years in space, Atmos. Chem. Phys., 18, 5699-5745, https://doi.org/10.5194/acp-18-5699-2018, 2018.

Formatted: Font: Not Bold
Formatted: Font: Not Bold
Formatted: Heading 2
Formatted: Font: 10 pt, Not Bold, Font color: Black
Formatted: Font: 10 pt, Not Bold, Font color: Black
Formatted: Font: Not Bold
Formatted: Font: 10 pt, Not Bold, Font color: Black
Formatted: Font: Not Bold
Formatted: Font: 10 pt, Not Bold
Formatted: Font: Not Bold
Formatted: English (U.S.)

Ludewig, A., Loots, E., Bartstra, R., Landzaat, R., Rozemeijer, N., Vonk, F., Leloux, J., van de Sluis, E., van der Plas, E., Harel, R., van Kempen, T., Tol, P., van Hees, R., and Kleipool, Q.: Level 1b product status/ Sentinel-5 Precursor Validation Team Workshop, ESRIN, Frascati, Italy - 11/14 November 2019.

Marchenko, S. and Deland, M.: Solar spectral irradiance changes during cycle 24, Astrophys. J., 789, 117-134, https://doi.org/10.1088/0004-637X/789/2/117, 2014.

Richter, A., Hilboll, A., Sanders, A., and Burrows, J.P., Inhomogeneous scene effects in OMI and TROPOMI satellite data, OMI-TROPOMI Workshop, convened on-line, October 26-29, 2020.

Rozemeijer, N.C. and Kleipool, Q.: S5P Mission Performance Centre. Level 1b Readme, S5P-MPC-KNMI-PRF-L1B, 2019-08-05. https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Level-1b-Product-Readme-File, 2019.

Schenkeveld, V. M. E., Jaross, G., Marchenko, S., Haffner, D., Kleipool, Q.L., Rozenmeijer, N. C., Veefkind, J. P., and Levelt, P. F.: In-flight performance of the Ozone Monitoring Instrument, Atmos. Meas. Tech., 10, 1957-1986, https://doi.org/10.5194/amt-10-1957-2017, 2017.

Sellitto, P., Bojkov, B. R., Liu, X., Chance, K., and Del Frate, F.: Tropospheric ozone column retrieval at northern mid-latitudes from the Ozone Monitoring Instrument by means of a neural network algorithm, Atmos. Meas. Tech., 4, 2375–2388, https://doi.org/10.5194/amt-4-2375-2011, 2011.

Shen, Zh., Zhu, X., Cao, X., and Chen, J.: Measurement of blooming effect of DMSP-OLS nighttime light data based on NPP-VIIRS data, Ann. GIS, 25, 153-165, https://doi.org/10.1080/19475683.2019.1570336, 2019.

~~Siddans, R., Latter, B.G., Kerridge, B. J., Weber, M., de Beek, R., Burrows, J. P., Aben, I., Tanzi, C., Hartmann, W., and Wickett, M. G.: GOME-2 error assessment study. Eumetsat contract: EUM/CO/01/901/DK, 18 December 2002, https://www.iup.uni-bremen.de/UVSAT_material/reports/GOME-2_FR_Ph_I-IV_-_Final_Report.pdf, 2002.~~

Singh, R. K. and Shanmugam, P.: A robust method for removal a glint effect from satellite ocean colour imagery, Ocean Sci. Discuss., 11, 2791-2829, https://doi.org/10.5194/osd-11-2791-2014, 2014.

Tedetti, M., Sempere, R., Vasilkov, A., Charriere, B., Nerini, D., Miller, W.L., Rawamura, K., and Raimbault. P.: High penetration of ultraviolet radiation in the south east Pacific waters, Geophys. Res. Lett., 34, L12610, https://doi.org/10.1029/2007GL029823, 2007.

Formatted: Font: 10 pt

Formatted: Left

Formatted: Font: (Default) Times New Roman, 10 pt, Not Bold

Varnai, T., Kostinski, A.B., and Marshak, A.: Deep Space Observations of sun glints from marine ice clouds, IEEE Geosci. Rem. Sens. Lett., 1-5, https://doi.org/10.1109/LGRS.2019.2930866, 2019.

Vasilkov, A. P., Joiner, J., Gleason, J. F., and Bhartia, P. K.: Ocean Raman scattering in satellite backscatter ultraviolet measurements, Geophys. Res. Lett., 29, 1837-1840, https://doi.org/10.1029/2002GL014955, 2002.

Westberry, T. K., Boss, E., and Lee, Zh.: Influence of Raman scattering on ocean color inversion models, Appl. Opt., 52, 5552-5561, https://doi.org/10.1364/AO.52.005552, 2013.

Ziemke, J. R., Strode, S. A., Douglass, A. R., Joiner, J., Vasilkov, A., Oman, L. D., Liu, J., Strahan, S. E., Bhartia, P. K., and Haffner, D. P.: A cloud-ozone data product from Aura OMI and MLS satellite measurements, Atmos. Meas. Tech., 10, 4067–4078, https://doi.org/10.5194/amt-10-4067-2017, 2017.

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., Hilton, B. B., Nicks, D. K., Newchurch, M. J., Carr, J. L., Janz, S. J., Andraschko, M. R., Arola, A., Baker, B. D., Canova, B. P., Chan Miller, C., Cohen, R. C., Davis, J. E., Dussault, M. E., Edwards, D. P., Fishman, J., Ghulam, A., González Abad, G., Grutter, M., Herman, J. R., Houck, J., Jacob, D. J., Joiner, J., Kerridge, B. J., Kim, J., Krotkov, N. A., Lamsal, L., Li, C., Lindfors, A., Martin, R. V., McElroy, C. T., McLinden, C., Natraj, V., Neil, D. O., Nowlan, C. R., O'Sullivan, E. J., Palmer, P. I., Pierce, R. B., Pippin, M. R., Saiz-Lopez, A., Spurr, R. J. D., Szykman, J. J., Torres, O., Veefkind, J. P., Veihelmann, B., Wang, H., Wang, J., and Chance, K.: Tropospheric emissions: monitoring of pollution (TEMPO), J. Quant. Spectrosc. Rad. Trans., 186, 17-39, https://doi.org/10.1016/j.jqsrt.2016.05.008, 2017.