Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-384-AC1, 2020

Atmospheric Measurement Techniques Discussions

Interactive comment on "Improved SIFTER v2 algorithm for long-term GOME-2A satellite retrievals of fluorescence with a correction for instrument degradation" *by* Erik van Schaik et al.

Erik van Schaik et al.

boersma@knmi.nl

Received and published: 15 June 2020

We thank the reviewer for this assessment and for his/her overall insightful and constructive review that has helped us to improve the manuscript. The reviewer's comment is in normal black font, our response is in blue font. In the accompanying revised manuscript the applied corrections are visible in the track changes.

The manuscript entitled with "Improved SIFTER v2 algorithm for long-term GOME-2A satellite retrievals of fluorescence with a correction for instrument degradation" by van Schaik et al. tackles some interesting technical aspects related to the GOME-2 SIF retrievals. In particular, the assessment of degradation issues is crucial for the

Printer-friendly version

Discussion paper

interpretation of the resulting data set.

Improvements were apparently achieved by narrowing the fitting window, resulting in a very similar retrieval set-up with respect to the NASA algorithm. Therefore it is not surprising that the results converge, but reassuring. While the manuscript is generally well written, there are some important corrections necessary, and I recommend to add some additional analyses. My major concern relates to the end-to-end test, where the retrieval seems to perform worse than for the real data, even though the simulations are representing the ideal case (noise free). To identify any fitting issues (over or underfitting), adding a Chi square analysis (expected vs observed) would be necessary, at least for the real data.

The simulations in the end-to-end test included noise, and represented as much as possible scenarios that are also encountered in reality, so with noise and high water vapour concentrations that exceed the range represented in the principal component spectra. We have now added a Chi square analysis in the discussions and also in the retrieval product.

Specific comments P2L9-10 van der Tol (2014) claims that SIF is driven by CO_2 concentration? This is certainly wrong, the CO_2 concentration is relevant for the dark reaction of photosynthesis.

Point taken. We removed this part of the sentence. The photochemical yield is influenced by the rate of CO_2 fixation by plants, and this rate is not directly related to the ambient CO_2 concentration.

P2L22-24 overly complicated sentence. Please rephrase

We rephrased the sentence.

P4L20 'understood' → is due to

Agreed. We changed to 'is due to'.

P6L33-P7L3 What justifies this particular interpretation? What means “and so on”? A physical meaning may be attached to the PCs, but this is not always the case. Here, all PCs show features related to oxygen and water vapour. Maybe remove this sentence?

We compared principal component spectra collected over the non-vegetated Sahara, and noticed that the first 4 PCs were very similar between GOME-2 and DISAMAR simulations. PC1 represents the mean, and PC2 mostly represents the variability caused by water vapour. We agree that a physical meaning is difficult to assign to PCs 3 and further, so we removed the “and so on”, as indeed the PCs show features related to variability caused by H_2O, O_2, and also other issues (noise, unresolved spectral residuals). We rephrased accordingly.

P7L3 This sentence has a trivial meaning, masking the important details of how many components are actually required to model the transmission with sufficient accuracy and if all of the PCs are indeed needed. This is relevant, because it is known that the number of PCs has effects on the retrieval accuracy and precision, as reported by Guanter et al. (2013), Joiner et al. (2013), and Kohler et al. (2015).

Our tests with real GOME-2 data also showed that the choice for the number of PCs is important, with 10 being the optimal choice (most robust patterns, highest precision) for our retrieval configuration with a constant number of PCs for each retrieval. We now include: *“Higher order PCs represent variability caused by water vapour, oxygen, and from other sources such as noise, unresolved surface and instrumental effects.”*

P7Eq2 The forward model is written in a way that assumes SIF as known. If, as claimed, there is one fitting parameter for SIF, there should be a normalized spectral shape in combination with one coefficient to fit.

Thanks for spotting this. There should indeed be a fitting coefficient term in Eq. (2) to reflect that the fit model allows the SIF-contribution to vary for different scenes. We updated Eq. (2) by introducing the fit parameter c in Eq. 2.

P7L15 The spectral shape of SIF in Daumard et al. (2012) is based on leaf level measurements, not "spectral measurements over various vegetated fields". ". . .considerable uncertainty on the shape. . ." is inaccurate. If fact, Magney et al. (2019) showed that the spectral shape of SIF in the far-red is remarkably stable across species and environmental conditions. However, in wavelengths $< \sim 730$ nm the shape is controlled by re-(absorption) effects, which is the origin of the uncertainty.

Daumard et al. [2012] measured fluorescence fluxes at two different wavelengths: 685 nm and 760 nm, and obtained fluorescence spectra based on these measurements. Their fluorescence spectra show variability especially below 730 nm as driven by chlorophyll concentration, a conclusion in line with Magney et al. [2019]. We updated the text accordingly.

P7L17 The sentence about the spectral shape is misleading as the two tested spectral shapes in Parazoo et al. (2019) are very similar. The main message here should be that the fitting window ideally covers the spectral region where SIF is stable, otherwise re-absorption effects may interfere with the retrieval. However, this limits the number of spectral points, which increases the retrieval noise. Conversely, a wider retrieval window makes the retrieval less noisy, but affects the retrieved SIF magnitude (introduces a bias). The overlap with solar Fraunhofer lines does not change the retrieved magnitude, it is a necessary prerequisite. Please rephrase according to my comment.

Agreed. We now adapted the text to read: *"A careful trade-off is required in the selection of the fitting window: it should preferably overlap with the stable part of the fluorescence reference spectrum (less uncertainty), contain several Fraunhofer lines, but preferably avoid strong absorption features from oxygen and water vapour, as we will see."*

P7L21 This sentence appears out of context.

We removed the sentence.

P7L28 I don't understand this sentence. What means "...so orthogonality between the albedo plus two-way transmission and SIF plus one-way transmission terms is important."?

We removed this sentence.

P7L30 I don't think that an "Internship Report" is a legit, citeable reference. Unrealistic albedo or transmittance values occur under "some viewing conditions" and lead to negative SIF estimates? This raises several questions and appears to be out of context, please remove this sentence entirely.

Since the research described in that internship report (an open source, but grey literature) does not help the clarity of the paper here, we have now removed these sentences.

P8L20 The opening of the sentence sounds odd. Replace "our understanding" by "our assumptions to model..."? Provide a reference for the DISAMAR radiative transfer model, the abbreviation has not been introduced.

Thanks for the suggestion: "assumptions to model" indeed captures better what we intend to say. A reference for DISAMAR, De Haan [2011], has now been added, and the acronym has been written out.

P9Fig. 2 Please check your y-axis, spelling and unit. What does absorption > 1 mean? Explained variances could be added.

We corrected the spelling of y-axis in Figure 2. Absorption refers to the two-way slant optical thickness caused by the absorption of light by water vapour and oxygen. We adapted the caption of Figure 2 accordingly.

P10L9 This sentence is odd. First, "spectra" is misleading, because PCs are compared. Second, it should be self-explanatory that a radiative transfer model is able to "capture the various relevant atmospheric processes".

The sentence refers to the Figure 2 that compares slant atmospheric absorption (optical thickness) spectra, so we don't think that that is misleading. As concerns the second point, we adapted the sentence and removed the part that DISAMAR captures the various relevant processes.

P10L11 The informative value of comparing PCs from observations and simulations is not obvious to me. The sign could be flipped, higher order PCs explain only a fraction of the variance and might occur at different positions (e.g. PC3 from simulations could be PC5 in observations). Would you clarify what this comparison should tell the reader and why it deserves a Figure?

The comparison was carried out to obtain confidence that we understand the main drivers of variability in the reference spectra, and to confirm that we can use DISAMAR as a tool to test retrieval settings. This was not obvious at first, but we agree that the (lower panel of) Figure 2 and associated text are better moved to the Supplement, which is what we have now done.

P10L20 As there is no word about noise in the simulations, I have to assume that they are noise free, in which case the poor retrieval performance shown in Table 2 is surprisingly disappointing. Without noise, the RMSE could be zero.

The caption of Table 1 mentions that "signal-to-noise ratio of the simulated reflectance spectra was 1,000". We now also included that information in the caption of Table 2.

P12-13 Even though it is interesting to explore alternative ways to identify poor retrievals. The authors should consider to use the reduced Chi square instead. The reduced Chi square would reveal any deficiencies in the retrieval set-up.

We thank the reviewer for the suggestion. Besides the metrics included in Table 2, we checked the reduced Chi squared metric to ensure that the retrieval is not suffering from overfitting or underfitting. We found that the non-faulty retrievals all have values close to 1, and faulty retrievals generally have a χ^2_{red} >3. We now included this information in

the main text, and report χ^2_{red} in the SIFTER v2 data product.

P13L9-21 The paragraph reveals that there is a serious issue when modelling water vapour absorption. This is an idealized scenario in which 65

As indicated in Table S2, the scenario is not exactly 'ideal'. Our purpose was to test the robustness of the fitting windows for situations with a much moister atmosphere (30-65 g m^{-2}) than in the reference PC set (4-40 g m^{-2}). The faulty cases occur for those scenes where water vapour > 40 g m^{-2}. For non-faulty retrievals, the 734-758 nm window provides the most unbiased results. In an additional test, we performed the retrieval with a PC set based on 30-65 g m^{-2}water vapour columns, and the number of faulty retrievals reduced strongly. We now explain this in the manuscript.

P14L21 Is there a reference to justify that the Saharan desert has a "dynamic range" of water vapour in the first place?

We found the dynamic range in the ECMWF ERA-Interim meteorological fields [Dee et al., 2011], which have been evaluated for water vapour with satellite measurements in Grossi et al. [2015]. These references have been included in the manuscript.

P15L5 "but not dramatically so"? This seems to be a bold claim.

The ECMWF data suggest that by column, water vapour over the Sahara is between 20-80% lower than over tropical forests. We removed the "dramatically so".

P17L15 Could you elaborate on the exact mechanism by which the slit function introduces a bias or is this hypothesis purely based on the latitudinal offset?

The GOME-2A slit function is known to change significantly over time because of temperature changes [Munro et al., 2016]. Changes in the shape of the slit function are known to have caused highly structured spectral responses and thereby problems with the fitting of HCHO, O_3, and NO_2 in the UV-Vis part of the GOME-2A spectra (e.g. De Smedt et al. [2012], Miles et al. [2015], Azam and Richter [2015], Beirle et al. 2017]). These changes occur along an orbit (thus with latitude) and (we now include

in the manuscript text) *"appear as an increase in the width of the slit function, with implications for the depth of the Fraunhofer structures: the wider the slit, the less deep the Fraunhofer lines. Shallower Fraunhofer lines may then be interpreted by the fitting algorithm to have been caused by fluorescence, which explains the positive fluorescence bias in the southern hemisphere (Figure 4). Conversely, sharper and deeper Fraunhofer lines (relative to the width of the lines over the Sahara), may well be interpreted as caused by negative fluorescence, explaining the negative bias for latitudes north of the Sahara."*

The problems have been mitigated in the trace gas algorithms by extending the fitting approach with dynamical fit parameters describing the width and shape of the slit function, with good results [Beirle et al., 2017]. Such an approach could also be attempted for GOME-2 SIF retrievals in the far-red part of the spectrum. These are limited by the reference spectra taken over the middle of the orbit, when the slit function has an intermediate width that is likely not representative for smaller widths north of 30°N, and the larger widths in the southern hemisphere.

P18L5 Could you add a Figure to illustrate the bias correction? I am particularly interested to see if a linear fit with radiance levels is justified.

The fit of the SIF zero level offset against radiance level is mostly a smoothing operation. The new figure S2 (right panel) shows the linear fit of SIF against radiances of all pixels with cloud fractions < 0.4 that have been observed at 45°N between 130-150 °W in July 2007. There is a weak relationship that suggests a more prominent non-zero bias for high radiance levels, i.e. when Fraunhofer lines are relatively well-defined. The mean bias at 45°N based on all pixels is -0.095 mW m^{-2} sr^{-1} nm^{-1} and the median value is -0.088 mW m^{-2} sr^{-1} nm^{-1}. For comparison, we also plotted the binned mean SIF values with intervals of 0.1×10^{13} photons cm^{-2} s^{-1} sr^{-1} nm^{-1} (light green squares). The linear regression through all data (black circles) follows the binned values quite well, suggesting that a linear fit is reasonable. We now include this figure (Figure 1 below and as Figure S2 in the revised Supplemental Material) and discussion

in the Supplemental Material.

P18L10 Does this sentence mean that you apply a daily and monthly bias correction? Please clarify.

We now include in the manuscript that *"In the operational retrieval, the bias correction is determined based on the most recent days as soon as a sufficient number of pixels has been collected in a latitude bin (at least 10 per bin). In practice, the bias correction is often based on the last day, but for some latitude bands, the correction can be based on pixels dating back at most 14 days."*

P18L22-26 Again, it is not appropriate to cite an "Internship Report" which is not publicly available.

We checked the guidelines of Copernicus Publications and note that "Informal or so-called "grey" literature may only be referred to if there is no alternative from the formal literature. Works cited in a manuscript should be accepted for publication or published already." In this case, we think it is relevant to refer to the Report by van Schaik [2016] which is publicly available via the url provided in the reference list.

P20L1 Could you explain how the degradation relates to the different L1 processors? Please discuss what kind of effect could potentially result from the change in L1 processors.

Figure 5(b) and the consistent trends in reflectance over Libya4 (Figure 1) suggest that the degradation is mostly progressing in time rather than with level-1 processor version (Table S1). For the changes in processor version, some impact of the change from 6.0 to 6.1 (in June 2015) might be expected – the other 2 changes (from 5.3 to 6.0 and from 6.1 to 6.2) did not affect calibration. For the processing of v6.1 data, the in-flight derived BSDF for solar radiometric calibration was introduced, which may have resulted in changes of the radiometric accuracy to unknown extent. However, we do not find any clear evidence for this in Figures 1, 6, or 10.

P21L11 Again, “Internship Report”.

See above.

P21L14 Is there a reason to assume that the uncertainty is driven by the SIF signal level? It should be driven by the radiance level. Please add (or replace) the uncertainty vs radiance level in Figure 6.

Figure 6 shows that the uncertainty is not driven by the SIF signal itself, which is reassuring. This was already discussed in the text. Thanks for the suggestion to also show the uncertainty vs. reflectance level. We now do that in the new Figure 6(b).

P22L6 Replace “yields” by values. How come that the filter rejects suddenly only 5-10

Done. The filter itself is the same. However, the retrieval scenes and reference PC set are different between GOME-2A and the DISAMAR tests. For example, the GOME-2A set contains a much larger set of surface albedo values, viewing geometries, SIF values, etc. A similar rejection rate should therefore not be expected.

P27Fig.9 The Amazon time series still shows a significant SIF decrease after 2013. Even though the reliability for the later period is already extensively discussed, are there other potential sources for artificial trends?

We have recently compared trends in SIF over the Amazon from our GOME-2A SIFTER v2 product to trends in OCO-2 SIF and NIRv from MODIS. The SIFTER v2 product is not showing a decreasing trend relative to these other data products over the Amazon (preliminary results). Other possible errors could be the representativeness of the PC set for later years, albedo trends, and insufficient correction for water vapour absorption.

P27L6 I see this as an overstatement. No retrieval algorithm has been developed, but different parameters have been optimized.

We removed this sentence.

References

Azam, F. and Richter, A.: GOME2 on MetOp: Follow-on analysis of GOME2 in orbit degradation, Final report, EUM/CO/09/4600000696/RM,2015, available at: http://www.doas-bremen.de/reports/gome2_degradation_follow_up_final_report.pdf (last access: 7 September 2016), 2015.

Beirle, S., Lampel, J., Lerot, C., Sihler, H., and Wagner, T.: Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives, Atmos. Meas. Tech., 10, 581–598, https://doi.org/10.5194/amt-10-581-2017, 2017.

De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J.-F.,Lerot, C., Theys, N., Valks, P., Hao, N., and van der A, R.: Im-proved retrieval of global tropospheric formaldehyde columnsfrom GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, Atmos. Meas. Tech., 5, 2933–2949,doi:10.5194/amt-5-2933-2012, 2012.

Miles, G. M., Siddans, R., Kerridge, B. J., Latter, B. G., andRichards, N. A. D.: Tropospheric ozone and ozone profiles re-trieved from GOME-2 and their validation, Atmos. Meas. Tech.,8, 385–398, doi:10.5194/amt-8-385-2015, 2015.

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lind-strot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A.,Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop series of satellites: instrument de-sign, calibration, and level 1 data processing – an overview, At-mos. Meas. Tech., 9, 1279–1301, doi:10.5194/amt-9-1279-2016, 2016.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-384, 2020.

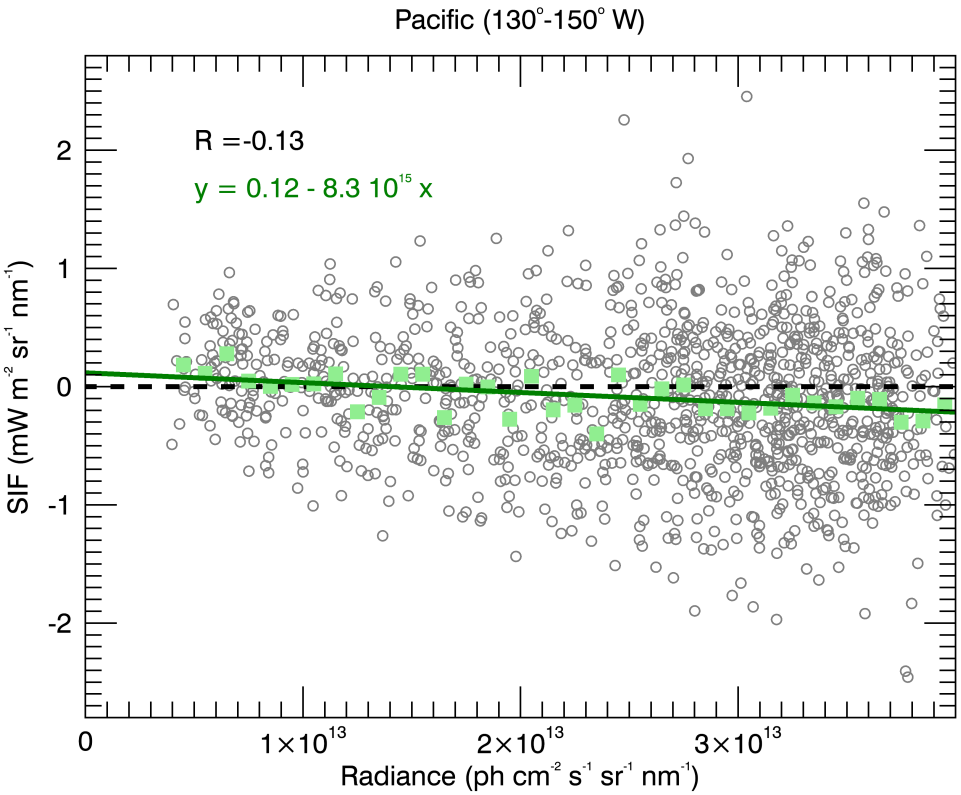

Fig. 1. SIF over the Pacific reference sector against radiances of all pixels with cloud fractions < 0.4 that have been observed at 45°N in July 2007.

Printer-friendly version

Discussion paper