

# Referee #1

We thank Referee 1 for this review of our manuscript. Below, we address the comments with the comments of Referee 1 in bold and our reply in normal font.

**This manuscript presents a revised long-term GOME-2 solar-induced chlorophyll fluorescence (SIF) data record with substantial improvements in instrumental degradation correction, inter-sensor consistency, and temporal stability. The topic is highly relevant to the ESSD community, and the authors make a significant effort to diagnose and mitigate known issues in the GOME-2 SIF record. Overall, the manuscript is well written and technically sound, and the dataset is potentially valuable for long-term ecosystem productivity and carbon-cycle studies. However, several aspects of the data processing and correction methodology would benefit from additional clarification, validation, and uncertainty characterisation.**

**In section 3.1, the instrumental degradation correction is derived by fitting long-term trends in globally averaged reflectance, under the assumption that the global mean reflectance should exhibit no secular trend and that any observed long-term change is attributable to sensor degradation. While this assumption is reasonable at first order, it would be helpful for users if the authors could further justify or validate it using the following ways:**

- 1. Have the authors evaluated whether long-term changes in cloud fraction, aerosol loading, or land surface properties could contribute to non-instrumental trends in global mean reflectance?**

Reply: The impact of other factors, including trends in cloud fraction, aerosols, and noted global greening on the global mean reflectance are intensively studied, but their combined global impact are uncertain, and contrasting results are found (Li et al., 2022). While, long-term global greening is expected to decrease planetary albedo, an observed increase in cloud cover over ocean regions acts in the opposite direction (Mao, 2019). Moreover, the reported overall geophysical trends on a global scale are much smaller, with MODIS data estimating a global decrease in global albedo of 0.0004 between 2002 and 2016 (Li et al., 2018), as compared to the impact of throughput loss due to instrument degradation of up to +10% in terms of reflectance (EUMETSAT, 2022). Hence, the impact of long-term changes in geophysical effects would be smaller than the accuracy of the degradation correction method used (Tilstra et al., 2012). Finally, the dominant role of instrumental effects is further evidenced by the pronounced scan-angle dependent trends in reflectance, which are of similar magnitude to the temporal trends and are not expected from geophysical drivers.

We include a brief discussion on the potential contribution of long-term geophysical variability to the global reflectance trend in the revised manuscript. Specifically, we add the following lines after line 134 of the manuscript:

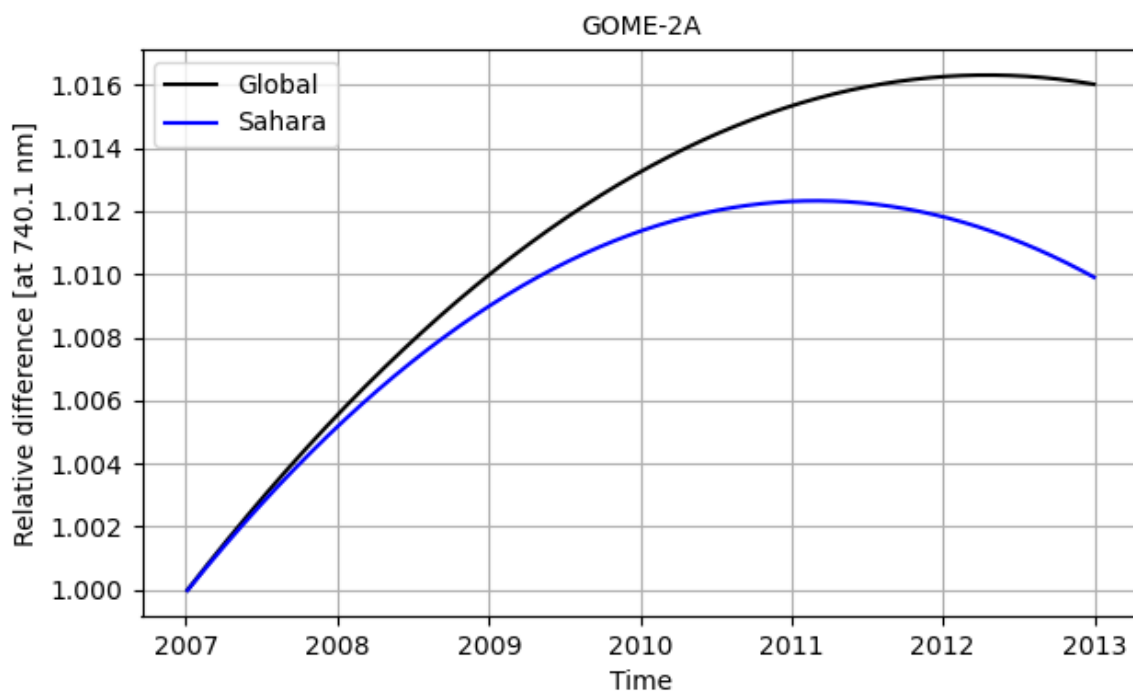
“While geophysical changes, such as variations in cloud fraction, aerosols, and global greening, can affect long-term trends in global mean reflectance, their impact is expected to be minor compared to the observed trends and attributed impact of throughput loss following instrument degradation (Li et al., 2018; EUMETSAT, 2022). Moreover, the pronounced scan-angle dependence of the observed long-term trends strongly indicates an instrumental origin.”

Furthermore, we soften lines 128-129 of the manuscript and change it to “Although no major long-term trends are expected, substantial long-term trends are noted (Fig. 2).”

- 2. Could the degradation polynomial be derived or cross-validated using more radiometrically stable reference targets (e.g. selected desert regions or invariant ocean areas), and if so, how consistent are the resulting correction factors with those derived from global averages?**

Reply: We appreciate this comment, and agree that a cross-validation over a more radiometrically stable reference target would provide additional confidence in the derived degradation correction. We, therefore, repeated the reflectance correction trend analysis over the Sahara desert region (16–30° N, 8 W–29° E), see Fig. 1, which is the same region over which the PC's are constructed. The analysis was performed by averaging over all scan-angles to maintain sufficient sampling.

The resulting trend over the Sahara shows similar overall behaviour compared to the Global trend in reflectance corrections, with an increase from 2007, flattening around 2011-2012, followed by a decrease in reflectance. However, the magnitude of the trend is reduced compared to the global average. This difference might result from overall higher residuals in the fit due to restricting the trend analysis to a smaller region, which increases sensitivity to variability of conditions and sampling (e.g., cloud conditions and viewing geometry).



**Figure 1.** Reflectance degradation correction of GOME-2A as a function of time at  $\lambda=740.1$  nm over the Global region (60° S, 60° N) (in black) and the Sahara region (16–30° N, 8° W–29° E) (in blue). The degradation is shown as relative difference from the reference date 5 January 2007. Both degradation fits used a polynomial order of  $p=2$  and Fourier order of  $q=6$ , these settings resulted in the best fit.

To understand the impact of this difference in degradation corrections on SIF, we compare this difference against the sensitivity test performed in the manuscript. In the manuscript, we include a sensitivity test for the degradation correction as part of the systematic error analysis in S2.5. In this sensitivity test, the polynomial order of the correction was altered from  $p=5$  to  $p=2$ , and its impact on GOME-2B SIF was evaluated over 615 pixels in the Congo Basin on 14 January 2017. Table 1 shows the results of this sensitivity test (as shown in Table S2.8 in the manuscript), including the relative difference in reflectance at 740.1 nm for both degradation correction settings.

This corresponds to a relative reflectance difference of 0.005, and is therefore comparable to the maximum difference between the global- and Sahara-based degradation corrections shown in Fig. 1, which is approximately 0.006 at 31 December 2013. This suggest that the impact on SIF is similar to the  $0.06 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$  for the tested case. Therefore, the difference between

the reference target remains well within the SIF uncertainty range identified in the sensitivity analysis (Table 1). This indicates that the degradation polynomial derived over the Sahara is consistent with that obtained from the global approach.

Given, the comparable behaviour and magnitude of the degradation corrections, the use of global averaging is preferred, as it enables a robust and stable estimation of trends, due to the substantially larger number of observations available per day and per scan angle.

**Table 1. Summary of sensitivity tests on the degradation correction, comparing polynomial order  $p=5$  (default) and  $p=2$  for GOME-2B. Results are shown for 615 pixels in the Congo Basin (13 S–6 N, 14–31 W) on 14 January 2017. Table from the manuscript (Table S2.8). The relative difference [at 740.1 nm] reflects the averaged over all wavelengths.**

Degr. Correction settings	Relative difference [at 740.1 nm]	SIF value [mW m <sup>-2</sup> sr <sup>-1</sup> nm <sup>-1</sup> ]	SIF uncertainties [mW m <sup>-2</sup> sr <sup>-1</sup> nm <sup>-1</sup> ]
$q=6, p=5$	0.996	$1.24 \pm 0.68$	$0.69 \pm 0.08$
$q=6, p=2$	1.001	$1.30 \pm 0.67$	$0.68 \pm 0.08$

Max ( $\Delta$ SIF value) = 0.06 mW m<sup>-2</sup> sr<sup>-1</sup> nm<sup>-1</sup>

- 3. In Section 3.2, the authors retrieve SIF using the 735–758 nm spectral window. However, Guanter et al. (2021) indicated that this window is more sensitive to atmospheric effects, particularly water vapour and cloud contamination, compared to narrower windows such as 743–758 nm. This raises the possibility that part of the observed inconsistency between GOME-2A and GOME-2B SIF over regions such as the Amazon and Eastern China, which were characterised by frequent cloud cover and high atmospheric humidity. This inconsistency may be driven by differences in atmospheric conditions rather than instrumental effects alone. Consequently, the robustness of the product in these regions may be partially limited by atmospheric influences inherent to the chosen fitting window. The authors are encouraged to discuss this potential limitation explicitly and, if possible, assess the sensitivity of the inter-sensor consistency to the choice of spectral window.**

Reply: The referee is right in raising concern regarding the sensitivity of SIF retrievals to atmospheric effects over regions characterised by high atmospheric humidity and frequent cloud cover. This sensitivity is a well-known challenge in SIF retrieval.

The 734–758 nm spectral window indeed partly overlaps with spectral regions affected by water vapour absorption features. However, the selected spectral window represents a compromise between minimising H<sub>2</sub>O absorption features and selecting sufficient bandwidth and data points for a robust fit. A sensitivity study by Parazoo et al. (2019) showed that, due to the coarser spectral resolution of GOME-2 (~0.5 nm), reducing the retrieval window to 742–758 nm (e.g., as used by Köhler et al. (2018) for TROPOMI SIF) leads to increased retrieval biases for GOME-2 SIF. Therefore, reducing the spectral window for GOME-2 is not expected to reduce overall bias and performance for GOME-2 SIF retrievals. Notably, Guanter et al. (2021) retrieves SIF from TROPOMI (~0.37 nm), which has a higher spectral resolution than GOME-2 (~0.5 nm) in the near-infrared (NIR). In addition to the retrieval window selection, the retrieval methodology, and particularly, the representation of atmospheric transmittance, plays a key role in mitigating atmospheric sensitivity. In Anema et al. (2025), we demonstrated that the SIFTER v3 algorithm, which is used in this work, advances the SIF retrieval robustness under humid conditions by better disentangling atmospheric effects from the SIF signal, due to improvements in the PC construction. We will clarify this by including the following lines after line 158 of the manuscript: “Residual sensitivity due to inclusion of water vapour absorption features may remain under humid conditions. This sensitivity is reduced in the SIFTER v3 algorithm through an improved representation of atmospheric variability in the principal component construction (Anema et al., 2025).”

Due to the consistent retrieval methodology to both GOME-2A and GOME-2B, atmospheric

sensitivity is expected to affect both sensors in a similar manner. However, we acknowledge that the sensitivity to atmospheric effects might amplify small instrumental differences in high humidity regions. Attributing the drivers of remaining intersensor differences, such as noticed over the Amazon and Eastern China, is challenging due to the interplay of various factors and goes beyond of this study. Future work will be needed to disentangle these contributions. We include the following sentences after line 362 of the manuscript to address this point:

“Both regions are characterised by higher atmospheric humidity. The enhanced retrieval sensitivity under such conditions may contribute to the observed intersensor difference. A full attribution of these remaining differences, however, is beyond the scope of this study and requires further investigation.”.

- 4. In Section 3.3, the authors assume that the systematic bias is primarily latitude-dependent and use the Pacific Ocean as a reference region for deriving the correction. However, ocean surfaces have intrinsically low radiance/reflectance, which may lead to systematically low retrieved SIF values and potentially different error characteristics compared to vegetated land. It is therefore unclear whether a bias estimated over water can be transferred to land vegetation in a physically consistent way. Please justify why an ocean-based reference is representative for correcting land SIF, and provide the post-correction seasonal SIF time series over the selected PCS regions to demonstrate that the correction preserves realistic seasonal dynamics. Moreover, previous studies have suggested that retrieval biases can be related to the level of reflected radiance (i.e., scene brightness) rather than latitude per se. Given this, the manuscript should clarify why radiance/reflectance-dependent effects were not explicitly evaluated (e.g., by stratifying bias as a function of reflectance or radiance, in addition to latitude). Including such an analysis would help disentangle whether the observed bias is genuinely latitude-driven or instead a manifestation of brightness- or scene-type-dependent retrieval errors.**

Reply: We thank the referee for this detailed comment. This manuscript follows the bias correction method as described and applied to GOME-2A in Anema et al. (2025). We realise that the description of the applied method in the current manuscript was not sufficiently clear and might have suggested that the bias correction is purely latitude-dependent. However, the correction is characterised as a function of reflectance for each latitude band, thereby accounting for brightness-dependent effects (see Anema et al. (2025) for details). We will revise the manuscript to clarify and expand the description of this aspect of the methodology. After line 204 of the manuscript we will add: “The correction is characterised as a function of reflectance (at 744 nm) per latitude band and on a daily basis, thereby accounting for brightness-dependent effects.”

The systematic bias described in Section 3.3 largely originates from instrumental effects, in particular the evolution of the slit-function across the orbit, which causes widening or narrowing of the Fraunhofer lines. This leads to apparent (too) positive or negative SIF offsets. Due to the instrumental origin of the effects, they are expected to affect both ocean and land observations in a similar way. The offsets are relative to the slit function width over the Sahara reference region, over which the principal components are obtained. This is supported by the fact that the zero-level offset over the latitude band corresponding to the Sahara region is close to zero, see Fig. 2.

To demonstrate that the correction preserves realistic seasonal dynamics over vegetated land regions, we include SIF time series adjusted and unadjusted for the zero-level bias in Fig. 3. The time series are shown for three distinct case studies analysed in this manuscript: (a) the Corn Belt, (b) Eastern China, and (c) the Pampas region. These results show that the seasonal cycle is preserved after correction. This is further supported by the strong agreement between the adjusted instantaneous SIF and the independent FluxSat data across all studied regions, as shown in Fig. 13 and Fig. S20 in the manuscript.

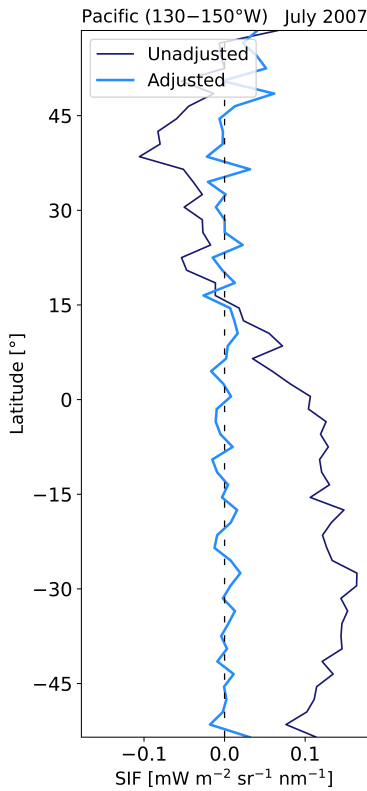


Figure 2. Instantaneous SIF over the Pacific Ocean adjusted (light-blue) and unadjusted (dark-blue) for the zero-level bias offset. SIF is averaged per latitude band of 2°. Figure from Anema et al. (2025).

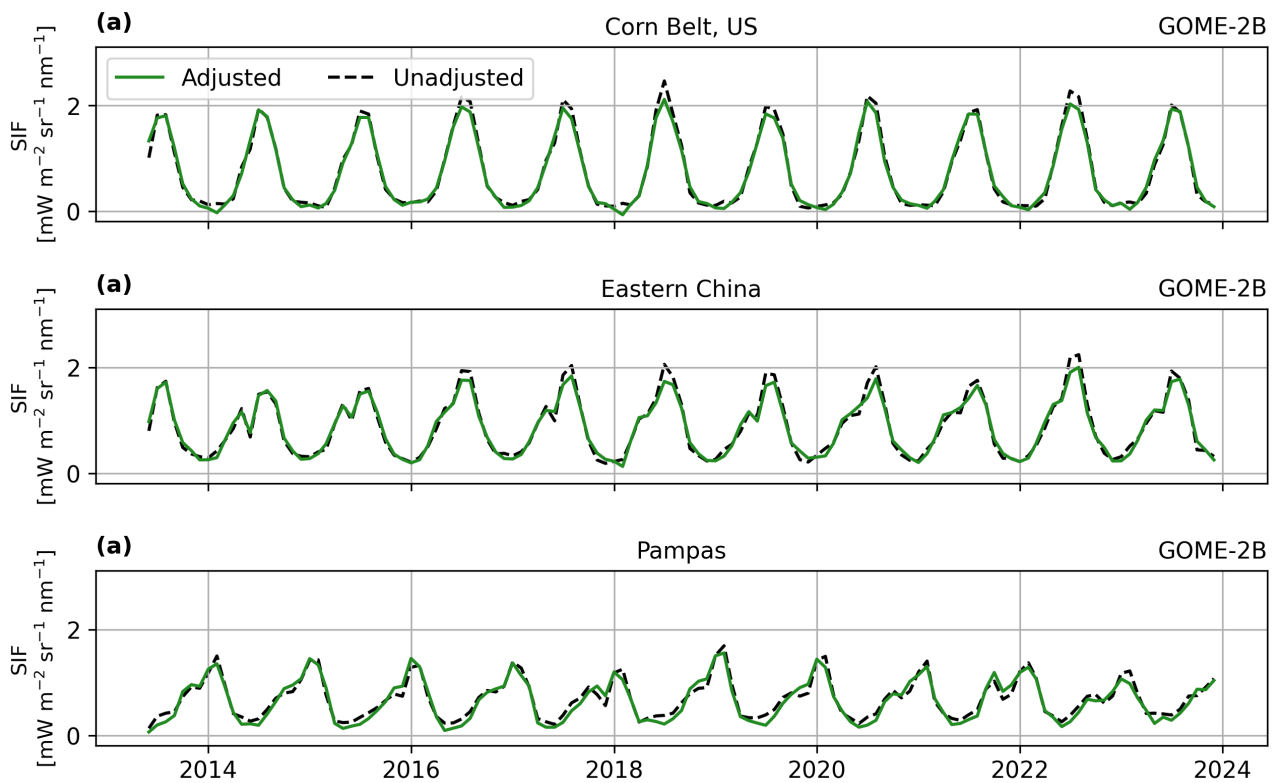


Figure 3. Time series of monthly averaged instantaneous SIF values adjusted (green solid line) and unadjusted (black dashed line) for the zero-level bias offset over (a) the Corn Belt, US, (b) Eastern China, and (c) the Pampas.

## References

- Anema, J. C., Boersma, K. F., Tilstra, L. G., Tuinder, O. N., & Verstraeten, W. W. (2025). Improved consistency in solar-induced fluorescence retrievals from GOME-2A with the SIFTER v3 algorithm. *Atmospheric Measurement Techniques*, 18(8), 1961-1979.
- EUMETSAT (2022). GOME-2 Metop-A and -B FDR Product Validation Report Reprocessing R3, EUM/OPS/DOC/21/1237264.
- Guanter, L., Bacour, C., Schneider, A., Aben, I., van Kempen, T. A., Maignan, F., ... & Zhang, Y. (2021). The TROPOSIF global sun-induced fluorescence dataset from the Sentinel-5P TROPOMI mission. *Earth System Science Data*, 13(11), 5423-5440.
- Köhler, P., Frankenberg, C., Magney, T. S., Guanter, L., Joiner, J., & Landgraf, J. (2018). Global retrievals of solar-induced chlorophyll fluorescence with TROPOMI: First results and intersensor comparison to OCO-2. *Geophysical Research Letters*, 45(19), 10-456.
- Li, Q., Ma, M., Wu, X., & Yang, H. (2018). Snow cover and vegetation-induced decrease in global albedo from 2002 to 2016. *Journal of Geophysical Research: Atmospheres*, 123(1), 124-138.
- Li, X., Qu, Y., & Xiao, Z. (2022). Responses of Land Surface Albedo to Global Vegetation Greening: An Analysis Using GLASS Data. *Atmosphere*, 14(1), 31.
- Mao, K., Yuan, Z., Zuo, Z., Xu, T., Shen, X., & Gao, C. (2019). Changes in global cloud cover based on remote sensing data from 2003 to 2012. *Chinese Geographical Science*, 29(2), 306-315.
- Parazoo, N. C., Frankenberg, C., Köhler, P., Joiner, J., Yoshida, Y., Magney, T., ... & Yadav, V. (2019). Towards a harmonized long-term spaceborne record of far-red solar-induced fluorescence. *Journal of Geophysical Research: Biogeosciences*, 124(8), 2518-2539.
- Tilstra, L. G., De Graaf, M., Aben, I., & Stammes, P. (2012). In-flight degradation correction of SCIAMACHY UV reflectances and Absorbing Aerosol Index. *Journal of Geophysical Research: Atmospheres*, 117(D6).