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., 2012a). 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.

   Modifications: To enhance clarity regarding the contribution of long-term changes of geophysical drivers in the global mean reflectance trend, we now include a brief discussion in the revised manuscript, in lines 134–138 of the revised manuscript with tracked changes. Furthermore, we added the word “major” in line 128 to avoid implying that no long-term trends of geophysical origin are expected in general.

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).

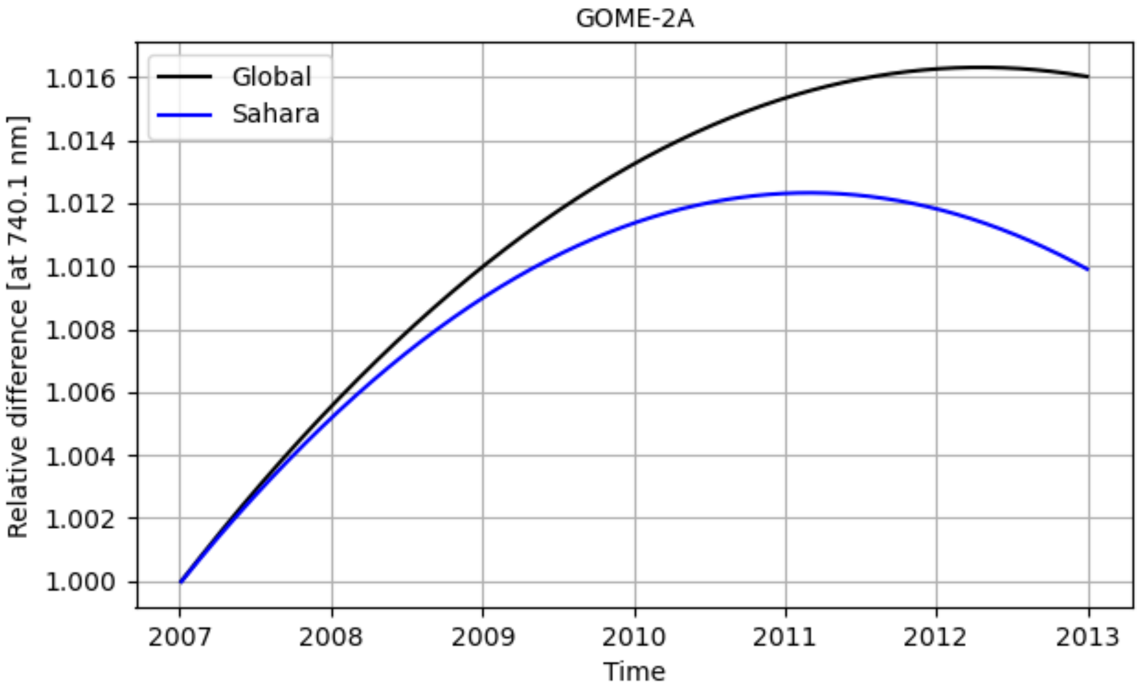

Figure 1. Reflectance degradation correction of GOME-2A as a function of time at λ=740.1 nm over the Global region (60° S, 60° N) (in black) and the Sahara region (16—30° N, 8° W—29° E) (in blue). The degradation is shown as relative difference from the reference date 5 January 2007. Both degradation 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 suggests that the impact on SIF is similar to the 0.06 mW m^{-2} sr^{-1} 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 adjusted 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^{-2} sr^{-1} nm^{-1}]	SIF uncertainties [mW m^{-2} sr^{-1} nm^{-1}]
q=6, p=5	0.996	1.24 ± 0.68	0.69 ± 0.08
q=6, p=2	1.001	1.30 ± 0.67	0.68 ± 0.08

Max (ΔSIF value) = 0.06 mW m^{-2} sr^{-1} nm^{-1}

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_2O 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.

   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 the scope of this study. Future work will be needed to disentangle these contributions.

   Modifications: We now discuss the water vapour sensitivity influenced by the selected retrieval window, as well as the reduction of this sensitivity through the use of SIFTER v3. Specifically, this is discussed in lines 162—165 of the revised manuscript with tracked changes.

Furthermore, to more clearly discuss the potential contribution of atmospheric humidity on remaining intersensor differences, we now include an additional discussion in lines 378–381.

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).

   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 relative to their width mid-orbit over the Sahara reference region. 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 show 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.

   <u>Modifications:</u> Lines 211–213 of the revised manuscript with tracked changes, now clarify and expand on the description of the latitude bias correction methodology. Furthermore, to show that the latitude bias adjustment resulted in realistic seasonal dynamics, we now include Fig. 3 in the revised supplement as Fig. S5.

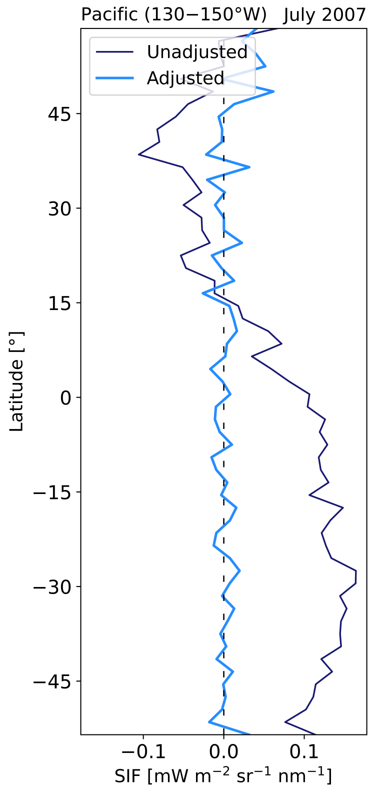

Figure 2. Instantaneous SIF over the Pacific Ocean adjusted (light-blue) and unadjusted (dark-blue) for the zero-level bias offset. SIF is averaged per latitude band of 2°. Figure from Anema et al. (2025).

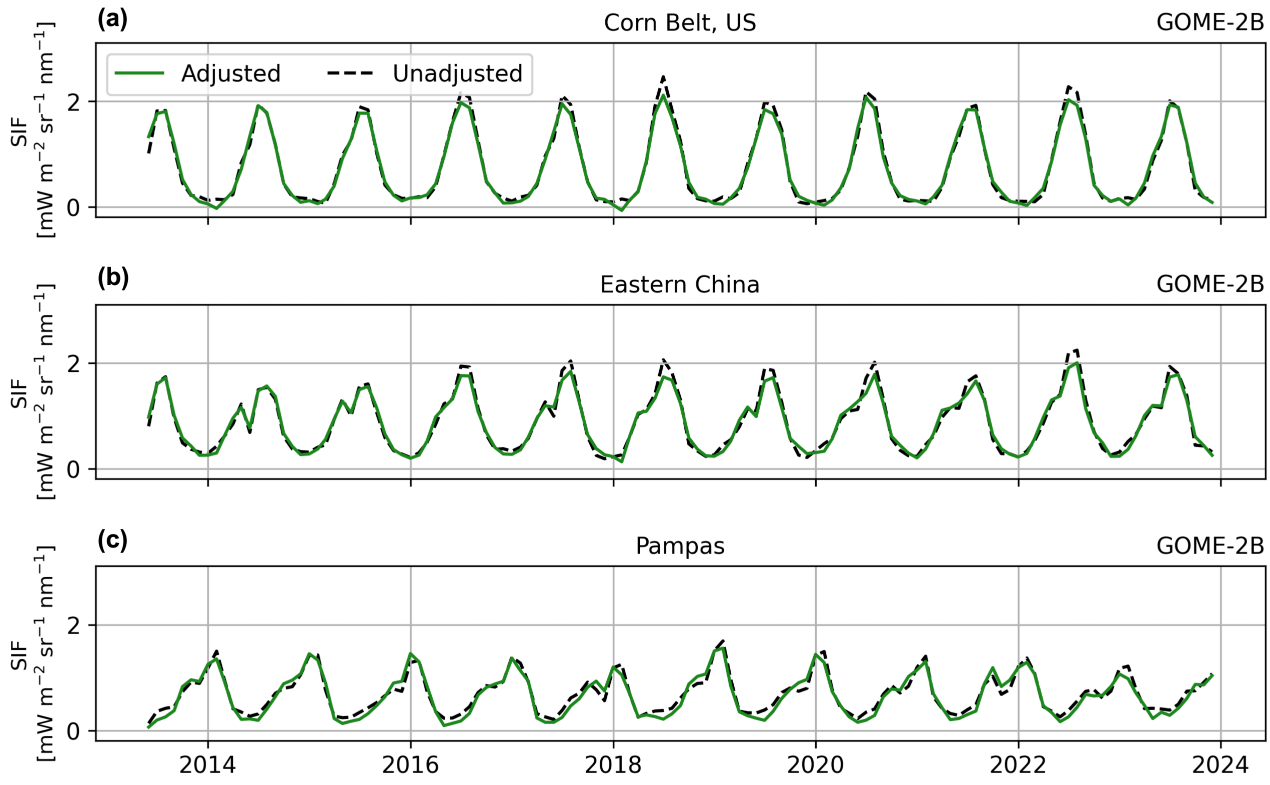

Figure 3. Time series of monthly averaged instantaneous SIF values adjusted (green solid line) and unadjusted (black dashed line) for the zero-level bias offset over (a) the Corn Belt, US, (b) Eastern China, and (c) the Pampas.

Referee #2

We thank Referee 2 for this review of our manuscript. Below, we address the comments with the comments of Referee 2 in bold and our reply in normal font.

This paper presents a harmonised GOME-2A and GOME-2B SIF dataset generated using the SIFTER v3 algorithm, along with an in-depth discussion on SIF biases caused by sensor degradation. This work is of significant value for the application and observation of SIF data. However, I have the following concerns:

1. **The degradation correction is based on the long-term trend of global daily mean reflectance. However, while this trend is attributed primarily to instrument degradation, could it also include contributions from long-term changes in cloud cover, aerosols, or surface albedo?**

   Reply: Reported overall trends in global reflectance are much smaller, with MODIS data estimating a decrease in global albedo of 0.0004 between 2002 and 2016 (Li et al., 2018), compared to the impact of throughput loss of up to +10% in terms of reflectance (EUMETSAT, 2022). Therefore the long-term changes from geophysical factors, such as, changes in cloud cover, aerosols, and greening, will be smaller than the accuracy of the degradation correction method used (Tilstra et al., 2012b).

   Note, that this point was also raised by Referee #1 in his/her first and second commend. In response to these commends we performed a robustness test over the bright Sahara region. The results indicated similar tendencies. This strongly suggests that the degradation of the sensor led to the observed long-term trends.

   Modifications:
   To enhance clarity regarding the contribution of long-term changes of geophysical drivers in the global mean reflectance trend, we now include a brief discussion in the revised manuscript, in lines 134–138 of the revised manuscript with tracked changes. Furthermore, we added the word “major” in line 128 of the original manuscript to avoid implying that no long-term trends of geophysical origin are expected in general.

2. **Figure 3 shows an overall increase in GOME-2A reflectance relative to the reference day across all scan positions during the first six years, followed by a decline. Could you please explain the reason for this initial increase?**

   Reply: The reflectance (*R*) is proportional to the observed radiance (*I*) to solar irradiance (I_0) (I:I_0). Reflectance degradation arises from differences in throughput loss between the observed radiance and solar irradiance over time (Tilstra et al., 2012a). Both follow different optical paths, making them subject to different instrumental effects. The decline in reflectance during the first years of GOME-2A reflects a stronger degradation of the solar irradiance signal compared to the radiance signal. Vice versa, an increase in reflectance, noticed in the later years, represents a stronger degradation of the radiance signal, as described in lines 144—147 of the original manuscript.

3. **Figure 4(a) shows that the spectral features (peaks/troughs) in PC#1 for GOME-2B are less coherent and shallower than those for GOME-2A. The authors attribute this to "higher reflectance uncertainty." However, could this also indicate a broadening of the Instrument Slit Function in GOME-2B? If the PCs fail to capture the sharp absorption features seen in GOME-2A, this could lead to systematic biases in SIF retrieval.**

   Reply: We agree with the referee that the slight difference in the slit functions could, in principle, contribute to differences in the principal components (PCs). While GOME-2A and GOME-2B share an identical instrumental design, small differences in their instrumental properties are unavoidable in practice. GOME-2A has a spectral resolution of 0.48 nm, whereas GOME-2B has a spectral resolution of 0.50 nm; corresponding to a 4% difference in effective slit function width. Such a difference could indeed lead to minor smoothing of spectral features in GOME-2B.

Overall, GOME-2A and GOME-2B show highly consistent spectral patterns in the PCs, with only minor difference in feature sharpness. The consistency indicates that the dominant transmittance structures are well captured in both sets of PCs. Nonetheless, slight differences in instrumental characteristics may cause systematic biases in the SIF retrieval, as discussed in the manuscript (Lines 297—298). The overall strong agreement between GOME-2A and GOME-2B SIF over the analysed regions suggest that residual differences do not significantly impact the robustness of the retrieved SIF.

Modifications: We now include the difference in spectral resolution as a potential source of the observed difference in #PC1, see lines 190–191 of the revised manuscript with tracked changes.

4. **Line 175：The PCs are derived from different time periods (2007–2012 vs. 2013–2018). Given potential interannual variability in atmospheric conditions (e.g., water vapour) over the Sahara, this temporal mismatch might introduce non-instrumental differences into the PCs. I suggest verifying the PCs using the overlapping period (2013–2017) for both sensors to isolate purely instrumental differences.**

   Reply: We appreciate the referee's suggestion to use the overlapping period for verification of the PCs to isolate instrumental differences. However, such a comparison would not isolate the instrumental differences in the case of GOME-2A and GOME-2B. The overlapping period is characterised by different viewing geometry sampling between the two instruments. During the overlapping period, GOME-2A operated under reduced swath configuration, limiting the viewing zenith angle range to [-34°, +34°], whereas GOME-2B operated under nominal swath configuration, ranging from [-54°, +54°]. Therefore, the derived PCs over the overlapping period will not be 1:1 comparable. In addition, the overlapping period coincides with the period of strongest instrumental degradation for GOME-2A, further complicating the isolation of purely instrumental differences.

   We computed both sets of PCs from a large selection of spectra, spanning 6 years of data, to ensure robust sampling of a wide range of atmospheric conditions. Furthermore, the Sahara region is considered to be radiometrically stable. Therefore, we do not expect substantial differences in the derived PCs related to atmospheric variability.

5. **Line 225：The observation that GOME-2B is more negative over barren areas but more positive over high-vegetation regions implies a multiplicative bias (gain difference) rather than a simple offset. If so, is the "simple additive correction" proposed later sufficient to address this amplitude-dependent discrepancy?**

   Reply: We thank the referee for this comment. While the observed pattern of more negative values over barren area and more positive values of SIF may suggest a multiplicative component, this pattern is not uniquely indicative of a multiplicative gain difference. In theory, the SIF signal over barren areas should be 0. However, several factors, including instrumental and atmospheric, can impact the apparent relative in-filling of the Fraunhofer lines, and thus the retrieved SIF. For example, recent work (Sanghavi et al., 2025), shows that atmospheric inelastic scattering causes in-filling of Fraunhofer lines which can introduce significant biases in SIF over vegetation-free regions (up to 0.35 mW m^{-2} sr^{-1} nm^{-1}), which may contribute to the slight difference in SIF between GOME-2A and GOME-2B over barren regions. Additionally, when accounting for the different viewing geometry sampling, the noted difference in GOME-2A and GOME-2B is not observed consistently across all regions. A true gain difference would be expected to produce a systematic scaling between the sensors. Hence, this would suggest that the discrepancies cannot be attributed to a simple multiplicative bias. Given this, the additive correction proposed in this study (Section 5.1) is designed based on the observed differences between GOME-2A and GOME-2B SIF over vegetative regions, where the SIF signal is most relevant.

   To further assess whether the discrepancy between GOME-2A and GOME-2B includes a multiplicative bias, we extend the tested model (Eq. 4 of the manuscript) by including an additional term allowing the seasonal amplitude to change after the transition at T_0, resulting in the following equation:

$y_t = \mu + \alpha t + S_t^* + \delta U_t + \epsilon_t$ (Eq. 1)

With S_t^* as:

$S_t^* = \beta_1 \sin(\frac{2\pi(t-\phi)}{12}) + \beta_2 \cos(\frac{2\pi(t-\phi)}{12}) + \zeta_1 U_t \sin(\frac{2\pi(t-\phi)}{12}) + \zeta_2 U_t \cos(\frac{2\pi(t-\phi)}{12})$ (Eq. 2)

We fit this model (Eq. 1) to the regions of Eastern China and the Amazon, where we noticed small discrepancies between GOME-2A and GOME-2B in the manuscript. Table 1 shows the regression results of this model fit. To allow easy comparison against the model with solely an additive factor, we've included the regression results from Eq. 4 of the manuscript within brackets in blue.

Table 1. Regression results of the model (Eq. 1) for Eastern China and the Amazon region. The results of the model used in the manuscript (Eq. 4 of the manuscript) are shown within brackets and in blue for comparison.

The results in Table 1 show that the inclusion of the multiplicative terms (ζ_1, ζ_2) only leads to a

Regression results	Eastern China	Amazon
R^2	0.90 (0.89)	0.89 (0.88)
$R^2_{adj.}$	0.89 (0.88)	0.89 (0.88)
RMSE [mW m^{-2} sr^{-1} nm^{-1}]	0.16 (0.16)	0.11 (0.11)
ϕ	0.02 (0.02)	0.97 (0.97)
μ [mW m^{-2} sr^{-1} nm^{-1}]	0.67** (0.68**)	1.21** (1.22**)
α [mW m^{-2} sr^{-1} nm^{-1}]	4.11e-4 (4.06e-4)	1.99e-4 (2.02e-4)
β_1	0.23** (0.24**)	-0.24** (-0.30**)
β_2	-0.51** (-0.59**)	0.30** (0.30**)
ζ_1	0.02	-0.01**
ζ_2	-0.13**	0.01
δ [mW m^{-2} sr^{-1} nm^{-1}]	0.13** (0.13**)	-0.10** (-0.10**)

* p<0.1, ** p<0.05

minor improvement in model performance by a slightly increase in R^2, but with no change in RMSE. Although in both regions, one of the interaction terms (ζ_1 or ζ_2) is statistically significant, the magnitude of these coefficients is smaller than the corresponding seasonal coefficients (β_1 or β_2). Furthermore, the other model parameters do not change substantially by the inclusion of the multiplicative terms, indicating that the model's structure is not fundamentally altered by these terms. While a multiplicative component cannot be fully excluded, the dominant discrepancy is captured by an additive offset. Finally, we advise caution in using the corresponding fit coefficients ζ_1 and ζ_2 for bias correction between GOME-2A and GOME-2B SIF, as this could distort genuine seasonal patterns.

Modifications: We now include an additional analysis and discussion regarding the use of multiplicative correction in the supplementary S3 and Table S7.

6. **Line 247 : suggests that eastern pixels might have higher SIF due to lower SZA. However, Figure 8 clearly shows that western pixels (VZA > 0) exhibit significantly higher SIF than eastern pixels during peak seasons (e.g., US Corn Belt in JJA). This indicates that the azimuthal anisotropy (sunlit/hot-spot effect on the west) dominates over the SZA timing effect.**

   Reply: We thank the referee for this observation. The intention of Section 4.2 of the manuscript is to discuss the various different mechanisms contributing to the viewing geometry dependence of SIF. The relative importance of each factor can vary depending on vegetation type, canopy structure, region, and time of observation. We first describe the effect that each mechanism can have individually. Figure 8 of the manuscript shows the combined influence of multiple mechanisms. The SZA timing effect, discussed in line 247, refers to the increase in illumination for eastern pixels (relative to nadir and western pixels), which can lead to higher SIF. Additionally, the hot-spot effect can lead to even larger enhancements towards the western side of the swath, as clearly shown in Figure 8.

   Modifications: The referee is right to observe the dominance of the hot-spot effect in the shown regions during the peak seasons. To enhance the clarity of the text, we now clearly state that we first discuss the various mechanisms affecting SIF, rather than the combined effect of all mechanisms, specifically in lines 247—249.

7. **Fig 10 shows that restricting viewing angles reduces the bias to 2.1%. However, as noted, the sampling is geographically limited to Northern latitudes. Given that tropical rainforests (e.g., Amazon) have distinct canopy structures and much higher SIF intensities—and showed the largest biases in Fig. 6—can validation based primarily on high-latitude data sufficiently demonstrate global consistency? Is there a risk that the sensors agree well at low/moderate SIF levels (high latitudes) but still diverge at high SIF levels (tropics)**

   Reply: The reduced swath configuration of GOME-2A limits the ability to perform strict co-sampled comparisons with GOME-2B exclusively to high latitudes. We acknowledge that this restricts a full validation of the consistency over tropical regions.

   However, our analysis presented in the manuscript indicates that the dominant driver of inter-sensor differences is the viewing geometry. In Fig. 8 of the manuscript, we show the relationship between GOME-2A and GOME-2B SIF as a function of VZA. This relationship is highly consistent across different regions, including tropical regions, such as the Amazon. This indicates that the VZA-dependence of the bias is not fundamentally different in the tropics compared to higher latitudes.

   Furthermore, confidence in the consistency of the sensors over high SIF levels is provided by the similar relationships observed between GOME-2A and FluxSat and between GOME-2B and FluxSat across several high-SIF regions, such as Eastern China and the Corn Belt (Fig. S2.21 and Fig. S2.22 of the original manuscript). These results do not indicate systematic divergence at higher SIF levels. While we can't fully exclude region-specific effects, the available evidence does not suggest a systematic divergence between GOME-2A and GOME-2B at higher SIF levels.

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). Reponses 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. (2012a). In-flight degradation correction of SCIAMACHY UV reflectances and Absorbing Aerosol Index. Journal of Geophysical Research: Atmospheres, 117(D6).

Tilstra, L. G., Tuinder, O. N., & Stammes, P. (2012b). A new method for in-flight degradation correction of GOME-2 Earth reflectance measurements, with application to the Absorbing Aerosol Index. In Proceedings of the 2012 EUMETSAT Meteorological Satellite Conference (Vol. 570).

Sanghavi, S., Frankenberg, C., Nelson, R. R., O'Dell, C. W., Rosenberg, R., & Joiner, J. (2025). Impact of Raman scattering on SIF retrievals from hyperspectral satellite observations. *Geophysical Research Letters*, *52*(15), e2024GL112777.

Dear editor,

While revising the manuscript, we noticed that we did not explicitly clarify that the FluxSat GPP data, shown in Figure 13, represents daily averaged GPP. To make this clearer to the reader, we added the following sentence to the caption of Figure 13:
"The used FluxSat data reflects daily averaged GPP".

The underlying data and analysis remain unchanged. This modification only provides additional clarification for the reader.