https://doi.org/10.5194/essd-2023-329

Response to Reviewer 2 Comments:

Zou et al. explored a pseudo-invariant method to resolve the temporal degradation issue with GOME-2A data and tested how the corrected SIF product improved. The study is of great significance in providing more reliable long-term SIF data. The manuscript was well written and the messages were well delivered. See my detailed comments below.

Thank you for taking the time to provide a thorough and insightful review of our manuscript. We appreciate your effort in providing valuable feedback. We carefully considered each of your comments and suggestions. Please find the point-by-point responses below.

Major comment 1: Not enough information has been given regarding why the GOME-2A is subject to degradation and where it occurred except in lines 43-44 that "GOME-2A is an optical spectrometer that measures reflected sunlight and is therefore sensitive to instrument degradation". If degradation is a problem for all the optical spectrometers, it should also be a problem for MODIS VIS/NIR bands. Identifying the reason for the degradation is crucial for determining the correction method to apply. The correction method applied in the present study assumes the GOME-2A radiance is uniformly downscaled at all wavelengths. This would not be correct if the degradation is caused by a dirty lens. Some results and discussions are required to validate this assumption.

Response: Thanks for this comment. We have investigated the reason for GOME-2's degradation in more detail. ESA's inspection of the GOME-2 sensor degradation indicates that the most likely cause is contamination of the optical lens (A. Hahne et al.,2012). The contamination was mainly caused by the Arathane conformal coating which is used in various places within the optical bench enclosure. This coating releases volatile products, which then deposit onto the optical elements, the cooled detectors, and probably the scan mirror. This material had been used already in GOME 1, where as well a degradation of the optical throughput was observed. The difference to GOME 1 is that in GOME 2, the board surface area has increased by a factor of more than 6, with accordingly more coating amount to act as a source for the off-gassing products. We have modified the description in the revised manuscript as follows:

Line 44:

Yet, the volatile coating used within GOME-2's optical bench enclosure makes the optical lens more susceptible to contamination, which eventually leads to instrument degradation (A. Hahne; Munro et al., 2016).

https://doi.org/10.5194/essd-2023-329

Degradation caused by lens contamination may also occur in MODIS VIS/NIR bands, but as far as the analysis we conducted (as shown in Fig 1b), it does not obviously affect MCD43C4 products.

Previous studies have shown wavelength dependence in the degradation caused by coating volatilization. However, the main manifestation is that the degradation is more obvious in the ultraviolet band and weaker in the visible and near-infrared bands. In our manuscript, only band-4 (593–790 nm) of GOME-2 was utilized, in which no obvious wavelength heterogeneity was found in previous studies. Specifically, for the fitting window we used, we added more analysis in the discussion (Sec. 5.1) of the revised manuscript to show the effect of wavelength on the degradation function as below:

Line 362:

In addition, the degradation at different wavelengths may also differ. Degradation functions fitted by different wavelengths in the 735–758 nm are compared. A difference of less than 1% was found in the degradation from 2007 to 2021 fitted at different wavelengths (Figure 13a). Similar conclusions can be drawn from Figure 13b. In Figure 13b, it is shown that inconsistency mainly occurs at the Fraunhofer line, which is inherently unstable in time. On the other hand, SIF retrieval relies on the filling of absorption lines. Extremely high fitting accuracy must be ensured if wavelengths are considered an influencing factor of the degradation function. Otherwise, the accuracy of SIF retrieval will be greatly affected. Therefore, in this study, the wavelength dependence of the degradation within the 735–758 nm window is ignored.

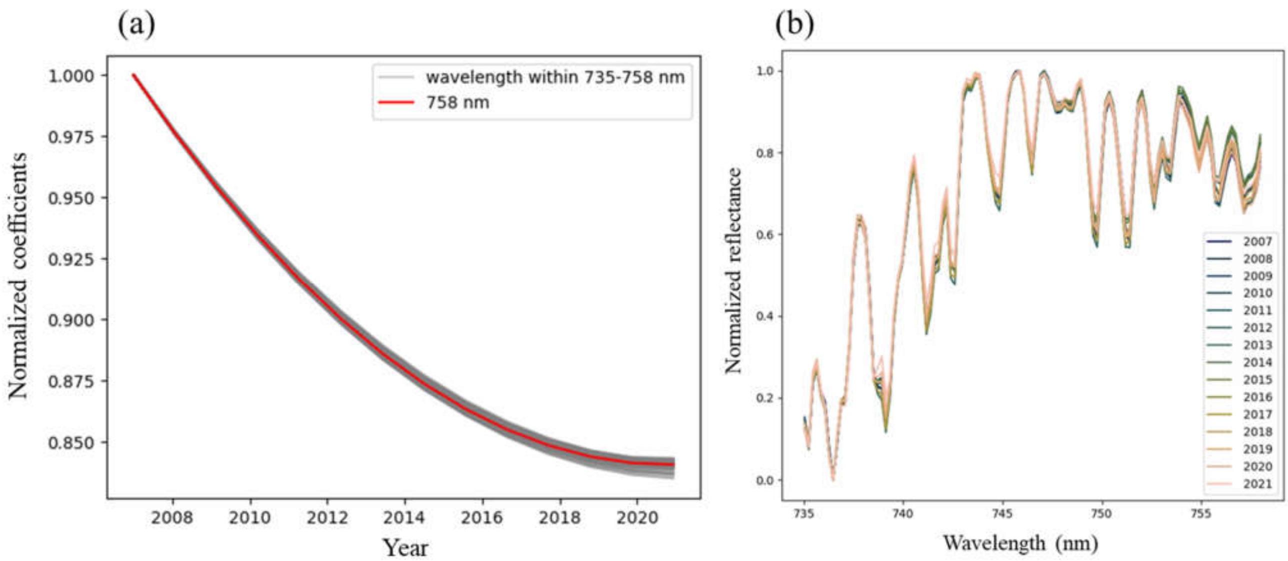


Figure 13 (a) The degradation function fitted using reflectance at different wavelengths in 735–758 nm fitting window. The red line is the result obtained at 758 nm, while the degradation functions fitted by other wavelengths were shown in gray. (b) Normalized NIR reflectance spectra in the 735–758 nm fitting window for different years in 2007–2021.

https://doi.org/10.5194/essd-2023-329

The results show that the inconsistency mainly occurs in the Fraunhofer lines, and the maximum fluctuation is 1%. On the other hand, fitting a wavelength-dependent function can result in huge uncertainties, since SIF retrieval is mainly based on the filling effect of SIF to the radiance/reflectance spectra. Slight inaccuracy in the fitting function can change the original depth of the absorption line, which will bring huge errors to SIF retrievals. Therefore, the slight difference in the degradation fitted by different wavelengths was neglected in this manuscript.

Major comment 2: A more systematic result session (data validation) is required. The pseudo-invariant method is actually a 2-stage correction to SIF: radiance correction and SIF retrieval. Thus, the method needs to be validated at both stages: comparison of radiance/reflectance to other products such as MODIS, and comparison to other products such as OCO-2 and TROPOMI. The data validation of the radiance/reflectance was missing from the present study.

Response: Thanks for this comment. The 2-stage validation was applied to TCSIF in the revised manuscript.

For the validation of GOME-2's radiance, MODIS may not be the most appropriate validation data source, since there are discrepancies in the overpass time, wavelength, and spatial resolution (as shown in Table R1). Measurements of GOME-2C instead of MODIS were utilized in the revised manuscript. A comparison of MODIS, GOME-2A, and GOME-2C was conducted in Table R1 to show the rationality. GOME-2C was launched in November 2018 onboard MetOp-C, which has a similar orbit altitude and spatial resolution as GOME-2A. The parameters of GOME-2A and GOME-2C are highly consistent, most importantly, with the same specification of the NIR band and the same spectral resolution.

Table R 1. Instrument parameters of GOME-2A, GOME-2C and MODIS

Sensor	Overpass time	NIR band (nm)	Ground pixel resolution
GOME-2A	9:30	590–790	80 km × 40 km/ 40 km × 40 km
GOME-2C	9:30	590–790	80 km × 40 km
MODIS	10:30 / 13:30	841~876	0.05°×0.05°(500m×500m)

At the same time, the degradation of GOME-2C can be neglected if we select the measurements in the early stage of the launch. Therefore, the validation of GOME-2A's radiance was conducted in January 2019 using measurements from GOME-C in the revised manuscript.

The introduction of GOME-2C was added in the Datasets part in Sec. 2.2 as follows:

Line 96:

2.2. Datasets for evaluation and comparison

https://doi.org/10.5194/essd-2023-329

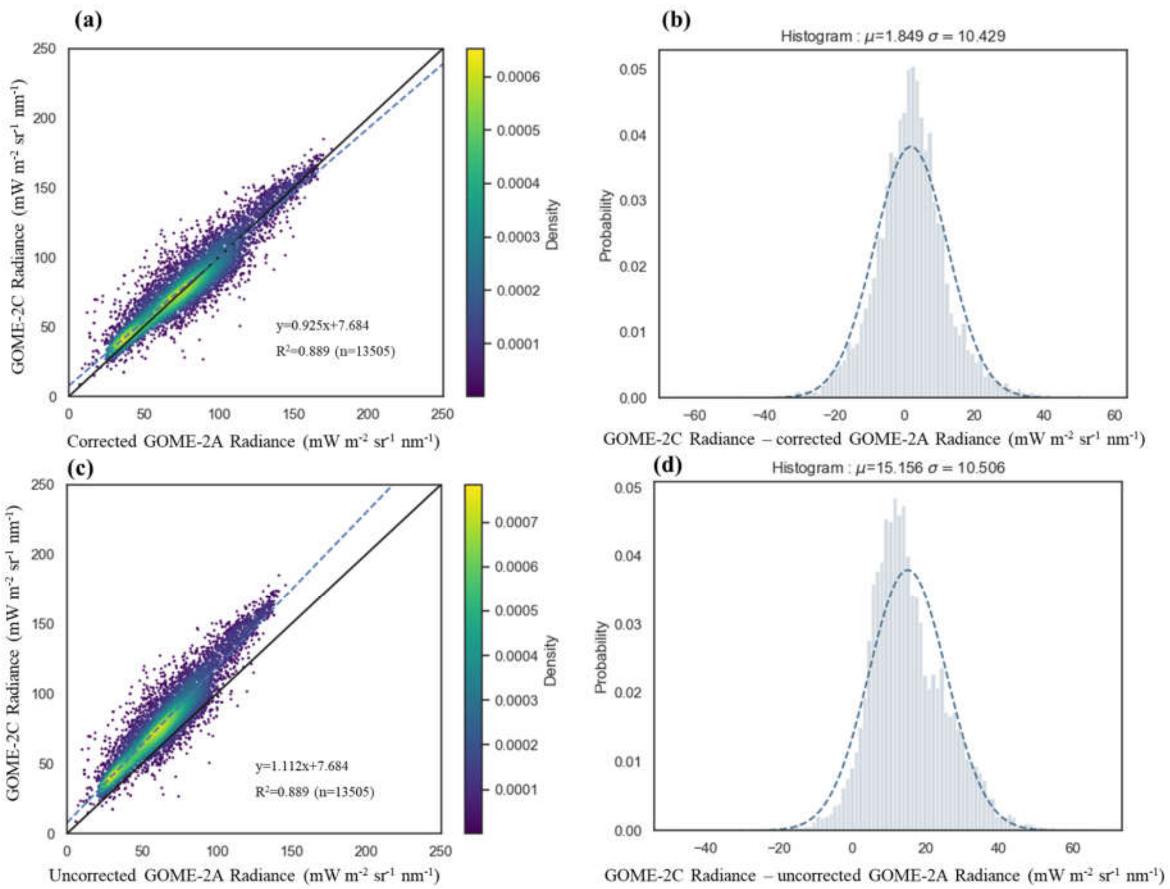
The reliability of this dataset was evaluated using greenness-based vegetation indices and global GPP products, as well as other established long-term SIF products.

Radiance spectra obtained from GOME-2C serve as a benchmark for the radiance spectra. Being a sensor that measures the same bands with the same spectral resolution as GOME-2A, GOME-2C has a later launching time in November 2018. Thus, measurements at the initial launch stage of GOME-2C can be taken as accurate values that are not affected by degradation.

The validation results of the corrected GOME-2A radiance spectra using GOME-2C are shown in Sec.4.1 below:

Line 256:

The temporally corrected GOME-2A NIR radiance was validated using GOME-2C radiance spectra (Figure 4). For the corrected GOME-2A radiance, the scatter plot shows that the majority of points are concentrated near the 1:1 line (Figure 4a). The slight positive offset of the linear regression may be caused by the lower orbit height of GOME-2C (817 km) than GOME-2A (827 km), which results in less atmospheric absorption in GOME-2C. The difference between the two products followed a Gaussian distribution with a small mean value of 1.85 mW m^{-2} sr^{-1} nm^{-1}, which is 2.3% of the mean GOME-2A radiance (Figure 4b). On the contrary, the mean deviation without temporal correction is 15.16W m^{-2} sr^{-1} nm^{-1} (Figure 4d).



https://doi.org/10.5194/essd-2023-329

Figure 4 Comparison between GOME-2A and GOME-2C NIR radiance after (a,b) and before (c,d) the temporal correction on 1 July 2019. The histogram of difference for GOME-C NIR radiance minus the corrected and uncorrected GOME-2A NIR radiance was shown in (b) and (d), respectively. Spatially matched pixels with cloud fraction of lower than 0.3 and SZA of lower than 70° were selected.

OCO-2 and TROPOMI SIF were used in the comparison with TCSIF. The comparisons were added in Sec.4.3 as shown below:

Line 295:

OCO-2 SIF and TROPOMI SIF were also involved in the validation of TCSIF (Figure 7e,f). To avoid discrepancies in wavelength and the overpassing time, the day-length corrected 740 nm provided by OCO-2 SIF, TROPOMI SIF, and TCSIF were compared. The spatially matched points were selected. TCSIF versus OCO-2 SIF and TCSIF versus TROPOMI SIF comparisons were conducted in July 2019 and July 2021, respectively. Both comparisons show high consistency with $R^2 > 0.65$, and the linear regression results are close to the 1:1 line.

https://doi.org/10.5194/essd-2023-329

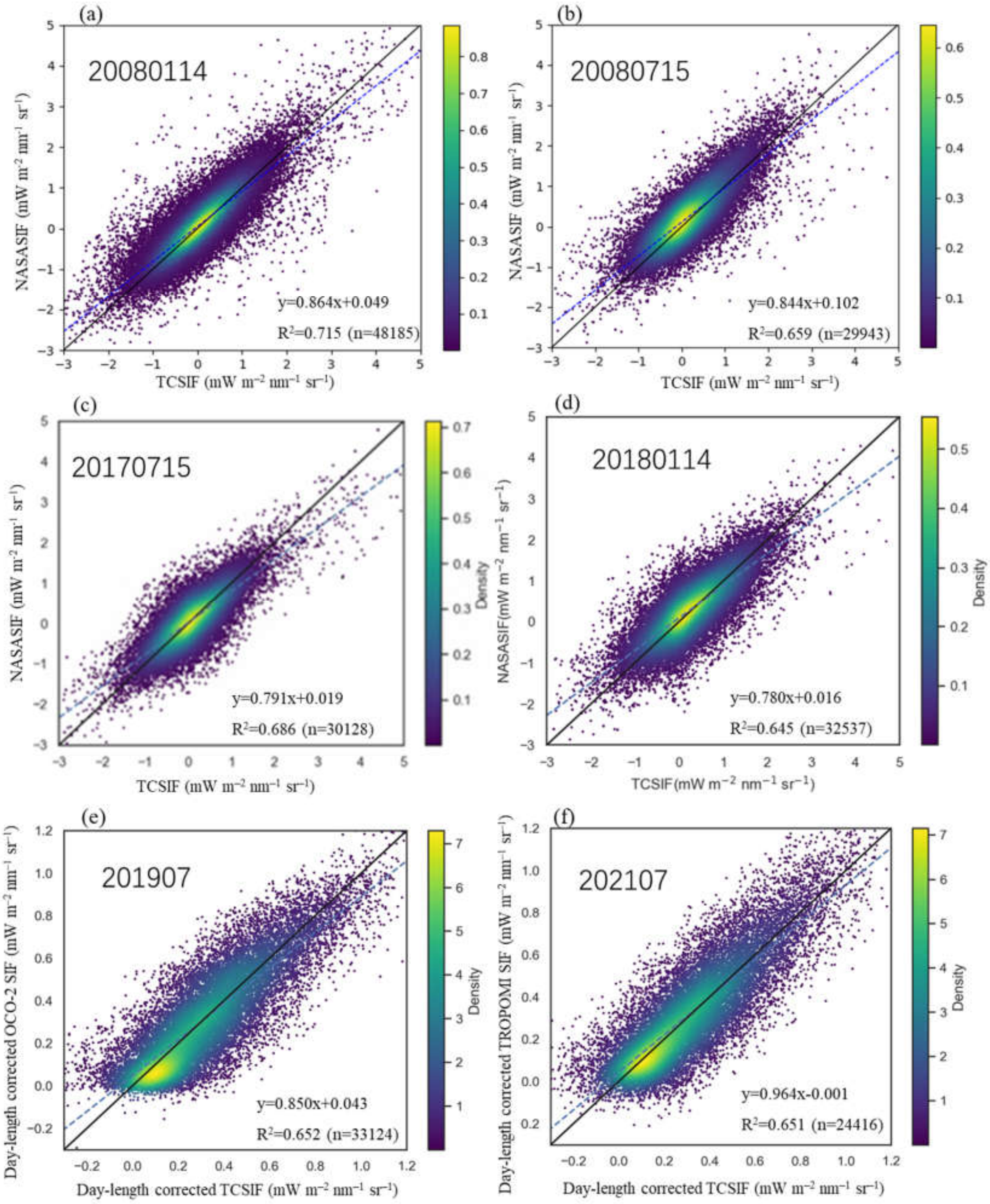


Figure 7. Comparison of TCSIF vs. NASA SIF on 14 January (a) and 15 July (b) in the year 2008, 15 July 2017 (c), and 14 January 2018 (d). Comparison of TCSIF versus OCO-2 in July 2019 (e) and TCSIF versus TROPOMI SIF in July 2021 (f). The comparison was made based on the level 2 product. Co-located pixels over land with a cloud fraction < 0.3 have been selected. The color of the scatter points represents the density of the points. The blue dotted line and the black solid line represent the line fitted based on the scatter points and the 1:1 line, respectively.

Besides, temporal comparisons were made between OCO-2 by demonstrating the trend map (in Sec.4.4 of the revised manuscript) as well as the yearly average curve (in Sec.5.3 of the revised manuscript).

Line 332:

https://doi.org/10.5194/essd-2023-329

4.4 Temporal variation in the TCSIF dataset

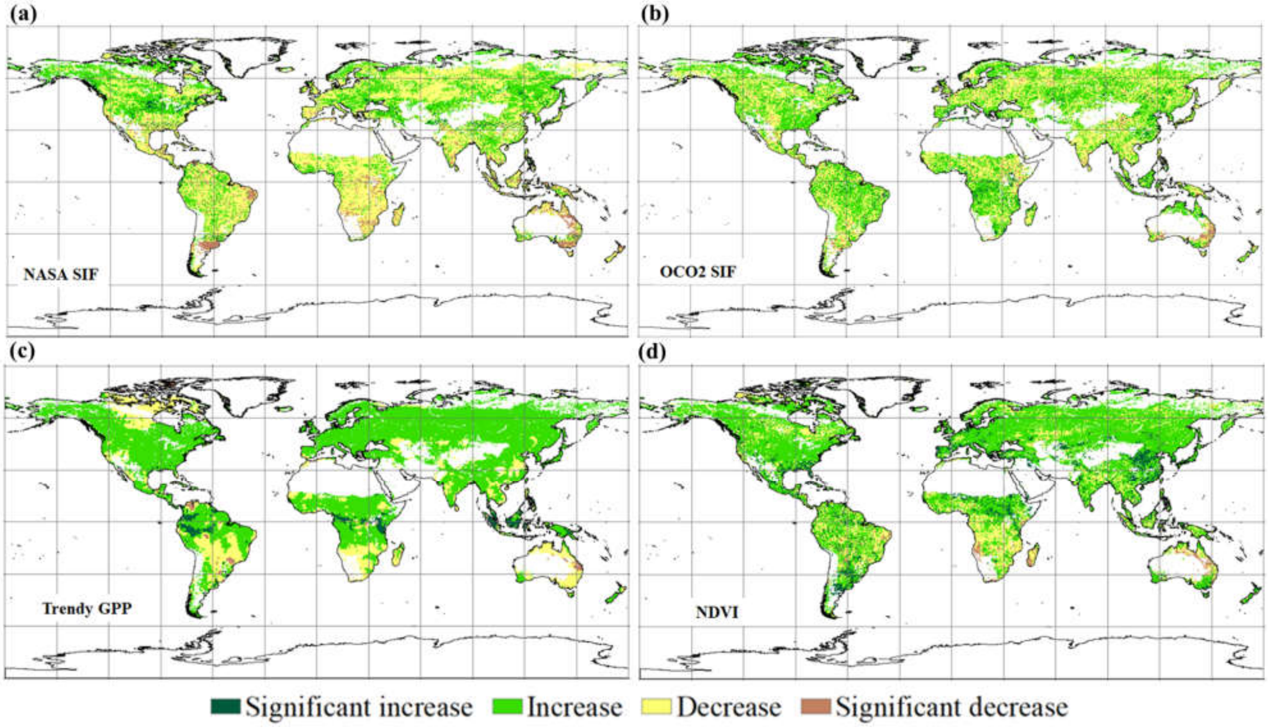


Figure 11. Map of trends in the annual average (a)NASA SIF for 2007–2018, (b) OCO-2 SIF for 2015–2021, as well as (c) trendy GPP and (d) NDVI for 2007–2021. The colors represent four types of trends (significant increase: positive correlation with $p < 0.05$; increase: positive correlation with $p \geq 0.05$; decrease: negative correlation with $p \geq 0.05$; significant decrease: negative correlation with $p < 0.05$).

As shown in Figure 11, 57.11% of vegetation areas are facing a decline in NASA SIF. On the opposite, as seen from OCO-2 SIF, trendy GPP, and NDVI, vegetation is growing over a large area globally (>70%) from 2007 (or 2015) to 2021. The main inconsistency between NASA SIF and the other products occurs in central and southern Africa, eastern Europe, and southern North America, where NASA SIF declines and the others increase. In southeastern China, vegetation greenings were found by TCSIF, OCO-2 SIF, and NDVI, while an insignificant downward trend was shown by trendy GPP. Vegetation growth in southern North America, Europe, the Amazon rainforest, central Africa, and Southeast Asia was detected by all the products apart from NASA SIF.

Line 400:

5.3 Comparison with other long-term SIF products

The annual average values of TCSIF and other long-term SIF products were compared (Figure 15). Importantly, most of the long-term SIF products were in agreement, featuring an increasing trend of SIF from 2007 to 2018, except for NASA SIF and SIFTER(Figure 15a–e). Among the temporal corrected SIF products, the annual curves of TCSIF and LTSIFc* are generally consistent, while LTSIFc* gives a higher growing

https://doi.org/10.5194/essd-2023-329
trend of 1.247% yr^{-1}, and the uncertainty of the growing trend of TCSIF(0.15 % yr^{-1}) is lower. A slightly decreasing trend of -0.08 % yr^{-1} characterized the SIFTER v2 product (Figure 15b), while the annual fluctuation of SIFTER v2 was clearly greatest among all the SIF products shown in Figure 15 (0.37 % yr^{-1}). The yearly trends according to TCSIF (1.06 % yr^{-1}) are close to the results from OCO-2 SIF from 2015 to 2021 (1.23% yr^{-1}, Figure 15f), while GOSIF shows a lower growing trend of 0.50% yr^{-1} during the same period. Compared to GOSIF, which was derived from OCO-2 SIF using a machine learning method, TCSIF is even more consistent with OCO-2 SIF, suggesting the flaws of machine learning methods in maintaining the temporal trend of the original SIF products.

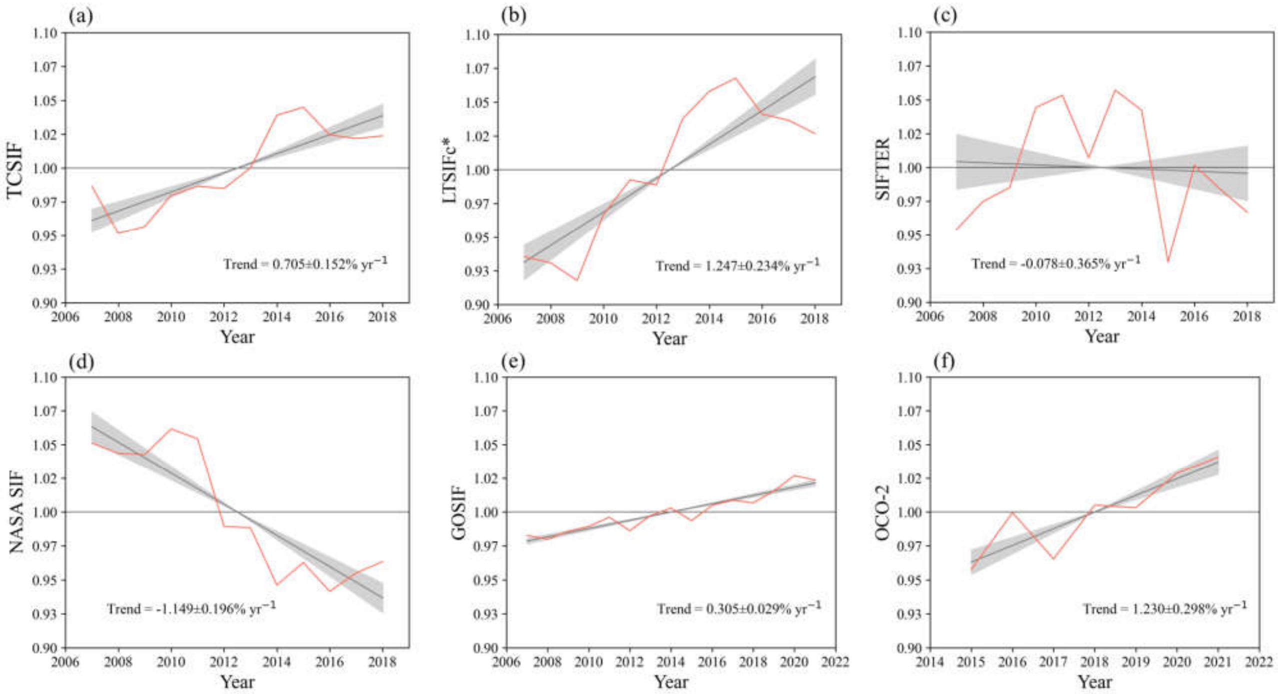


Figure 15. Comparison of temporal trends in the annual SIF average from (a) TCSIF, (b) LT_SIFc*, (c) SIFTER v2, (d) NASA SIF during 2007–2018, as well as (e) GOSIF during 2007–2021, and (f) OCO-2 SIF during 2015–2021. All data shown are normalized to relative values (by dividing the mean). The shaded areas indicated the standard deviations.

The analysis of the TROPOMI time series is not shown because the TROPOMI gridded SIF dataset are available only between 2018.4 and 2021.3, which means that there are only two full years of data in 2019 and 2020. Therefore, TROPOMI SIF is currently not able to provide accurate temporal trends.

Major comment 3: Following my comment 1, the GOME-2A TOA reflectance at the calibration site (before calibration) has a very clear seasonality in it. Similar seasonality is also found in MODIS data as shown in Fig. 2. I believe the seasonality is due to the BRF effect caused by sun-sensor geometry. But the variation of GOME-2A reflectance seems to be higher than MODIS, why? Also, the GOME-2A reflectance

https://doi.org/10.5194/essd-2023-329

variation seems to increase with time (Fig. 3a). If the degradation is due to the dirty lens, the radiance/reflectance variation should also scale with the Dfactor, right?

Response: Thanks for this comment. Quantify analyses were made to show the annual standard deviation and annual range (maximum minus minimum) in the NIR reflectance of GOME-2A and MODIS. The results are shown below:

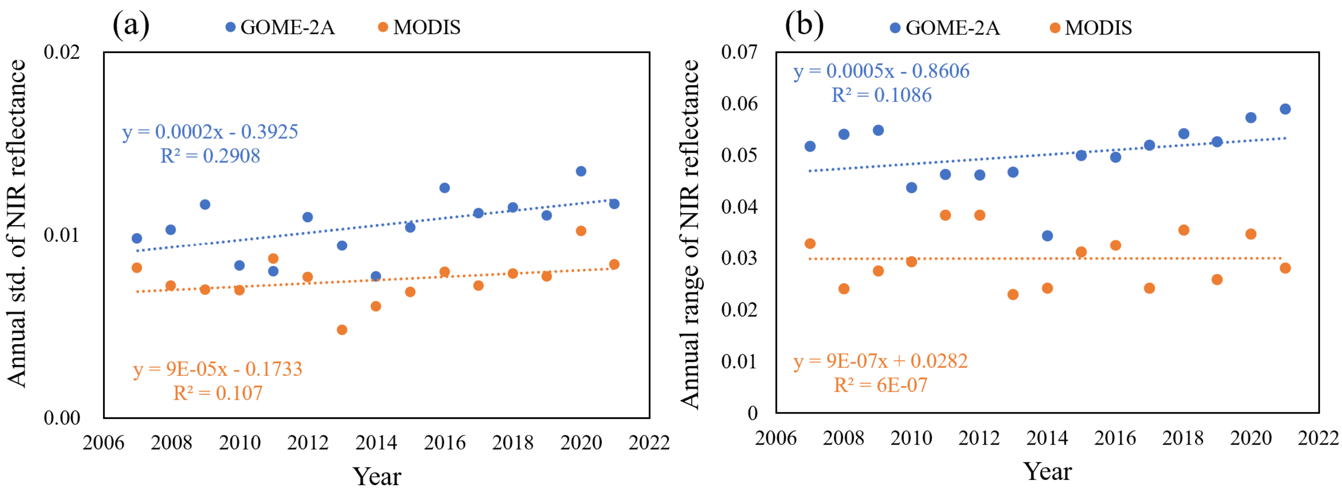


Figure R 1 Annual variation of GOME-2A and MODIS NIR reflectance at the calibration site. (a) The annual standard deviation of NIR reflectance. (b) The annual range of NIR reflectance calculated by the annual maximum minus the annual minimum. Blue and orange dots stand for GOME-2A and MODIS, respectively.

Firstly, the reason for the higher variation of GOME-2A than MODIS may be caused by the effect of the atmosphere. It should be noted that MCD43C4 provides the corrected surface reflectance with atmospheric correction while GOME-2A does not. Although we have reduced the influence of the atmosphere on the TOA reflectance by selecting only pixels with cloud fraction equal to zero, the results may be affected by atmospheric scattering.

Secondly, growing variations were observed in the NIR reflectance of GOME-2A. A similar result was found in GOME-2A chlorine dioxide products (Pinardi et al., 2022), while the authors speculated that the increased noise in GOME-2A after 2011 came from the degradation of the instrument. The results suggested that the instrument degradation may affect the radiance spectra in other forms apart from the contamination of the lens. However, the pattern of such an effect is not easy to find, since the trend is slight (with an average annual change less than 0.05%) and not significant (p-value>0.1).

Compared with the inevitable intra-annual variations, the fitting of the general interannual trend is the main focus of our study. We have added relative analysis to the discussion as follows:

Line 391:

https://doi.org/10.5194/essd-2023-329

Besides, the contamination of the lens may not be the only reason for GOME-2A's degradation. As shown in Figure 3, the annual variations in NIR reflectance do not decrease as the annual average does. On the contrary, the fluctuation is growing with time. A similar phenomenon was found in the chlorine dioxide products (Pinardi et al., 2022) that GOME-2A results are noisier than those of GOME-2B, especially after 2011. These results suggest that in addition to the decline in reflectance over time caused by lens contamination, the temporal degradation is impacting GOME-2A measurements in other forms. However, the pattern of this effect is not clear now, further research is needed on the impact of GOME-2A's degradation on its measurements in more aspects. Therefore, only the interannual decline trend was considered in this study, while the inevitable intra-annual variations caused by other factors such as BRF and atmosphere were neglected.

Minor comment 1: Line 10. SIF cannot provide a "direct way" to monitor photosynthesis. SIF and GPP are not linearly correlated.

Response: Thanks for this comment. The sentence has been revised as follows:

Satellite-based solar-induced chlorophyll fluorescence (SIF) **serves as a valuable proxy for** monitoring the photosynthesis of vegetation globally.

Minor comment 2: Line 11. I thought the TROPOMI and OCO SIF datasets were more popular. It is not necessary to say it is the most popular.

Response: Thanks for this comment. The SIF products of TROPOMI and OCO-2 are indeed popular because of their high spatial resolution. At the same time, GOME-2 is currently the sensor that can provide the longest SIF time series. From this perspective, we believe that GOME-2 SIF would be of great significance without the problem of time decay. We modified the text to:

Global Ozone Monitoring Experiment-2A (GOME-2A) SIF product **has gained widespread popularity, particularly** due to its extensive global coverage since 2007.

Minor comment 3: Line 28. of monitoring -> to proxy.

Response: Thanks, the sentence has been revised as follows:

Solar-induced chlorophyll fluorescence (SIF) retrieved from satellite-based hyperspectral data provides a new way **to proxy** the photosynthesis of vegetation globally.

Minor comment 4: Line 45. The contamination assumption does not seem to be able to explain the variations in TOA reflectance

https://doi.org/10.5194/essd-2023-329

Response: Thanks for this comment. Previous research has attributed sensor degradation to lens contamination with volatile coatings (A. Hahne; Munro et al., 2016). The sentence in Line 45 has been revised as follows to show the details:

Yet, the volatile coating used within GOME-2's optical bench enclosure makes the optical lens more susceptible to contamination, which eventually leads to instrument degradation (A. Hahne; Munro et al., 2016).

As analyzed in major comment 3, BRF, atmospheric scattering, as well as other forms of sensor degradation are affecting the intra-annual variation of GOME-2A's spectra. However, such effects are irregular and have little effect on fitting the general trend (as shown in major comment 3). Therefore, these inevitable variations are ignored in the manuscript, and the limitation was added in the discussion as shown in major comment 3.

Minor comment 5: Line 91-95. The use of external PAR to rescale SIF makes the SIF data more prone to errors in external data. Also, since the GOME-2A radiance is corrected using external NIR data from MODIS, should it be considered an L3 product? My understanding is that the L2 product is purely inferred from the L1 product without any external data correction.

Response: Thanks for this comment. It is inevitable to introduce external signals (such as PAR, EVI, or fPAR, as shown by Eq.(10) in the manuscript.) while extending the instantaneous SIF signal to the daily average. PAR instead of cos(SZA) was used in the temporal extend model, which indeed adds an external variable. However, as shown in the manuscript, using PAR is a better solution to compensate for the significant effects of diurnal weather changes due to cloud and atmospheric scattering. The results using PAR are shown to be better in the research of Hu et al. (2018). In addition, the instantaneous SIF and daily SIF were provided simultaneously in the Level 2 product. Using instantaneous SIF will avoid the impact of external data.

Also, there are several reasons for defining the un-gridded TCSIF products as Level 2 products:

Firstly, the MCD43C4 products were only used to prove the reliability of the pseudo-invariant site, and not directly used for the correction. The degradation function was fitted merely based on GOME-2A NIR reflectance at the pseudo-invariant site.

Secondly, such a definition makes it easier to distinguish the un-gridded data and the gridded data and keep consistent with other satellite SIF products such as NASA SIF.

Therefore, the un-gridded TCSIF is still defined as Level 2 products in the revised manuscript.

Minor comment 6: Line 96. Should it be better to use MODIS data to validate the corrected GOME-2A

https://doi.org/10.5194/essd-2023-329
radiance/reflectance and other SIF data to validate the SIF product? The NDVI/GPP can only be used as indirectly supporting results.

Response: Thanks for this comment, we agree that NDVI and GPP can only serve as comparable parameters, rather than benchmarks for validations. We have moved these comparisons to discussion and added a more direct comparison in the results with OCO-2 SIF, and TROPOMI SIF following the previous suggestions, **as detailed in major comment 2**.

It is also necessary to validate the accuracy of the temporally corrected radiance before SIF retrieval. However, MODIS will not be an ideal choice for the validation since there are differences in the spatial resolution, spectral band, and spectral resolution between MODIS and GOME-2A. Instead, we find that GOME-2C measurements can serve as a benchmark. Please find the detailed analysis in the major comment 2.

Minor comment 7: Fig. 1. How about other bands? Are they also very stable?

Response: Thanks for this comment. In Figure R 2, we show the MCD43C4 Red reflectance (Band 1) as a comparison for the NIR Band. Reflectance at the red band shows similar yearly cycles and interannual variations with NIR reflectance.

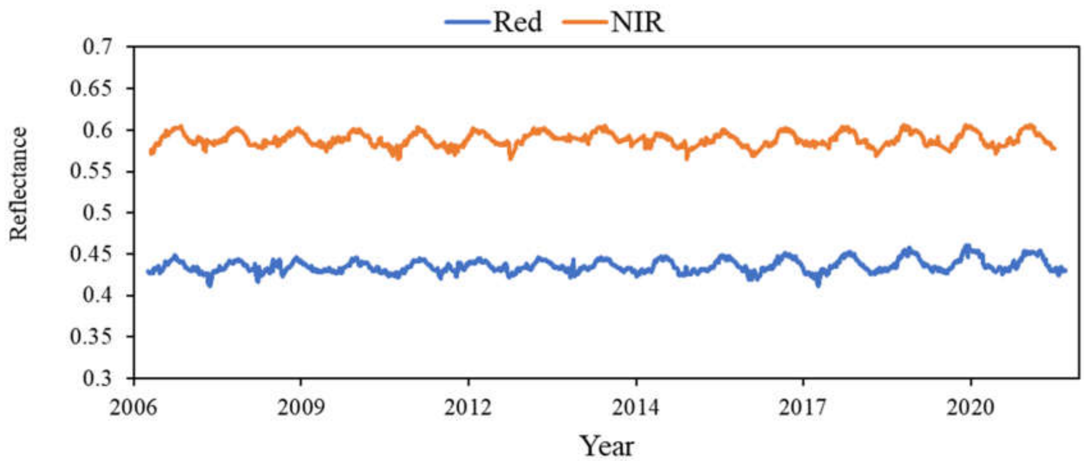


Figure R 2 The surface reflectance at the red and NIR band of the MCD43C4 product.

Yes, the other bands of MODIS also show stable interannual trends. Since we only use 735–758 nm wavelengths for SIF retrieval in this study, only the NIR band of MCD43C4 was manifested in the manuscript.

Minor comment 8: Fig. 2. Since there is no vegetation in the calibration site, the spectral curves should align to the same standard curve when rescaled properly? If you rescale them, is it the case? If so, the Dfactor method is fine; if not, the Dfactor needs to be a wavelength-dependent function!

https://doi.org/10.5194/essd-2023-329

Response: Thanks for this advice, we have normalized the radiance spectra by the difference normalization method, the results are added in Figure 13 of the revised manuscript (**as detailed in Major comment 1**). The results show little discrepancies between different wavelengths in our selected fitting window for SIF retrieval (735–758 nm). The main difference appears at the Fraunhofer absorption line between 745 and 758 nm, while the normalized reflectances almost overlap at other wavelengths. Detailed explanations for not considering wavelength in the degradation function were given in major comment 1. Firstly, the deviation caused by wavelength is neglectable. Secondly, to ensure the accuracy of the SIF retrieval, it is better not to include wavelength as an additional independent variable to fit the degradation function without understanding the principle of change.

Minor comment 9: Line 221. Is the EVI used as a f_APAR here?

Response: Thanks for this comment. Yes, the EVI is used as a substitution for f_APAR here. Previous studies (Hu et al., 2021) have demonstrated that the FAPAR can be directly quantified by spectral vegetation indices such as NDVI and EVI, and can be estimated using a VI-based linear stretching model (Sellers et al., 1994; Jiang et al., 2002; Liu et al., 2017, 2020). Therefore, the retrieved parameter f_APAR was replaced with EVI as a directly observable parameter.

Minor comment 10: Fig. 3. The degradation looks to be an exponential curve, why use a polynomial function to fit it?

Response: Thanks for this comment. We have conducted quantified experiments, and the results show that the polynomial function can better fit the data (with higher R^2, as shown in Figure R 3). Meanwhile, we select the polynomial function in this study because it is simple to provide robust estimation.

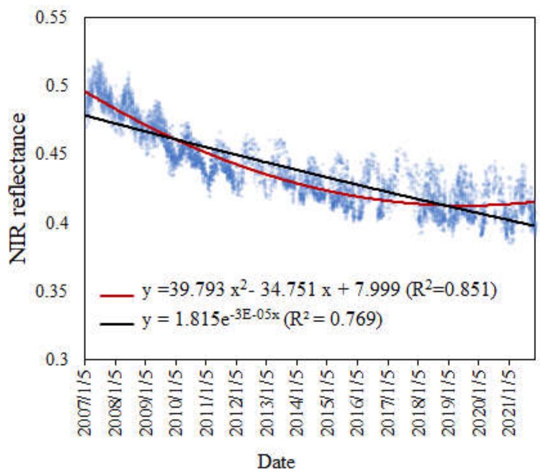


Figure R 3 Degradation function fitted for GOME-2A NIR reflectance using exponential and

https://doi.org/10.5194/essd-2023-329

polynomial models

Minor comment 11: Fig. 6. Combine it with Fig. 5.

Response: Thanks for this comment. The two figures have been combined into the new Figure 6 in the revised manuscript as follows:

Line 292:

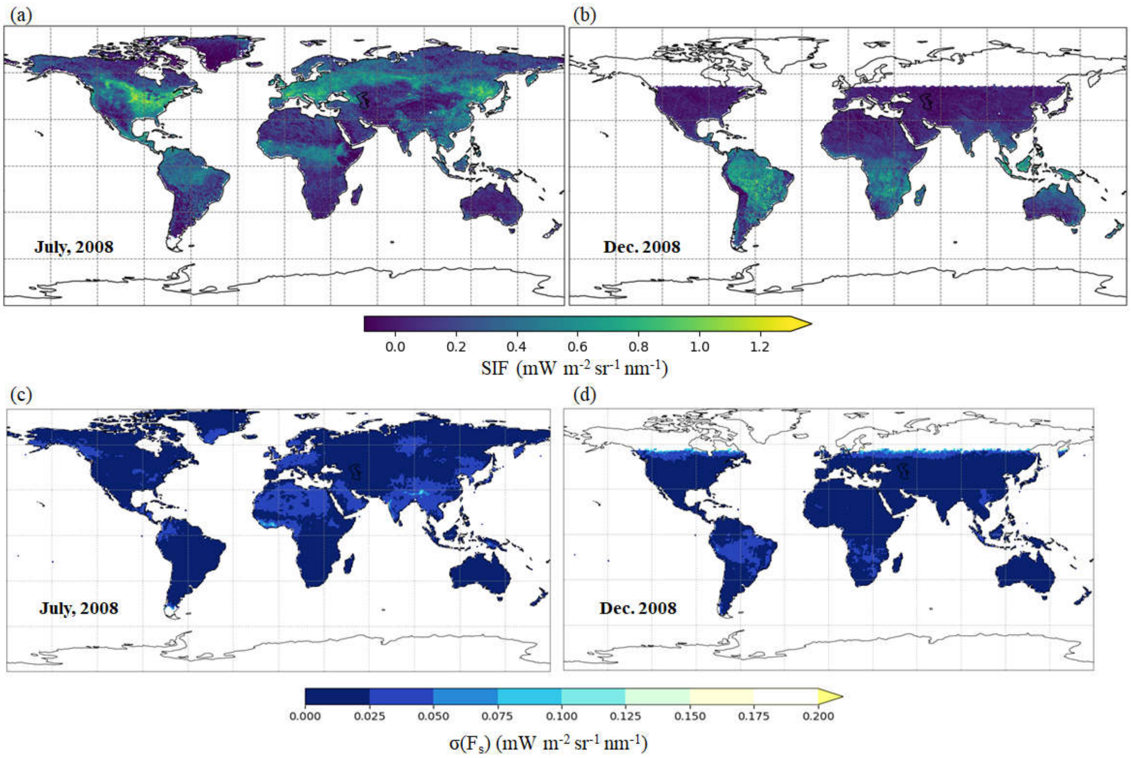


Figure 6. Global patterns in the upscaled monthly TCSIF(a, b) and standard error of the weighted mean (σ(F_s))(c, d) in July (a, c) and December (b, d) in the year 2008.

Minor comment 12: Fig. 7. You also need to show examples of how the TCSIF and NASA SIF differ in the year 2021 (or more recently) to show the degradation effects.

Response: Thanks for this comment. Comparisons between NASA SIF and TCSIF in recent years are better to show the effects of degradation. However, NASA SIF did not provide data beyond 2019, thus, the comparisons in July 2017 and January 2018 were conducted instead. The revision is as below:

Line 286:

We also compared the spatially matched TCSIF and NASA SIF pixels **in January and July 2008, July 2017, and January 2018** (Figure 7a–d). The linear relationships between the two SIF products revealed these to be strongly correlated ($R^2 > 0.65$), significant (p-value < 0.05), and close to the 1:1 correspondence line

https://doi.org/10.5194/essd-2023-329

(slope > 0.84) for either season in 2008 (Figure 7a, b). **For comparison in 2017 and 2018 (Figure 7c, d), there are still good linear relationships between the TCSIF and NASA SIF (R^2 > 0.64). However, it is worth noticing that the regression line deviates from the 1:1 line in both 2017 and 2018(slope < 0.80), which was caused by the degradation in NASA SIF.**

Line 300:

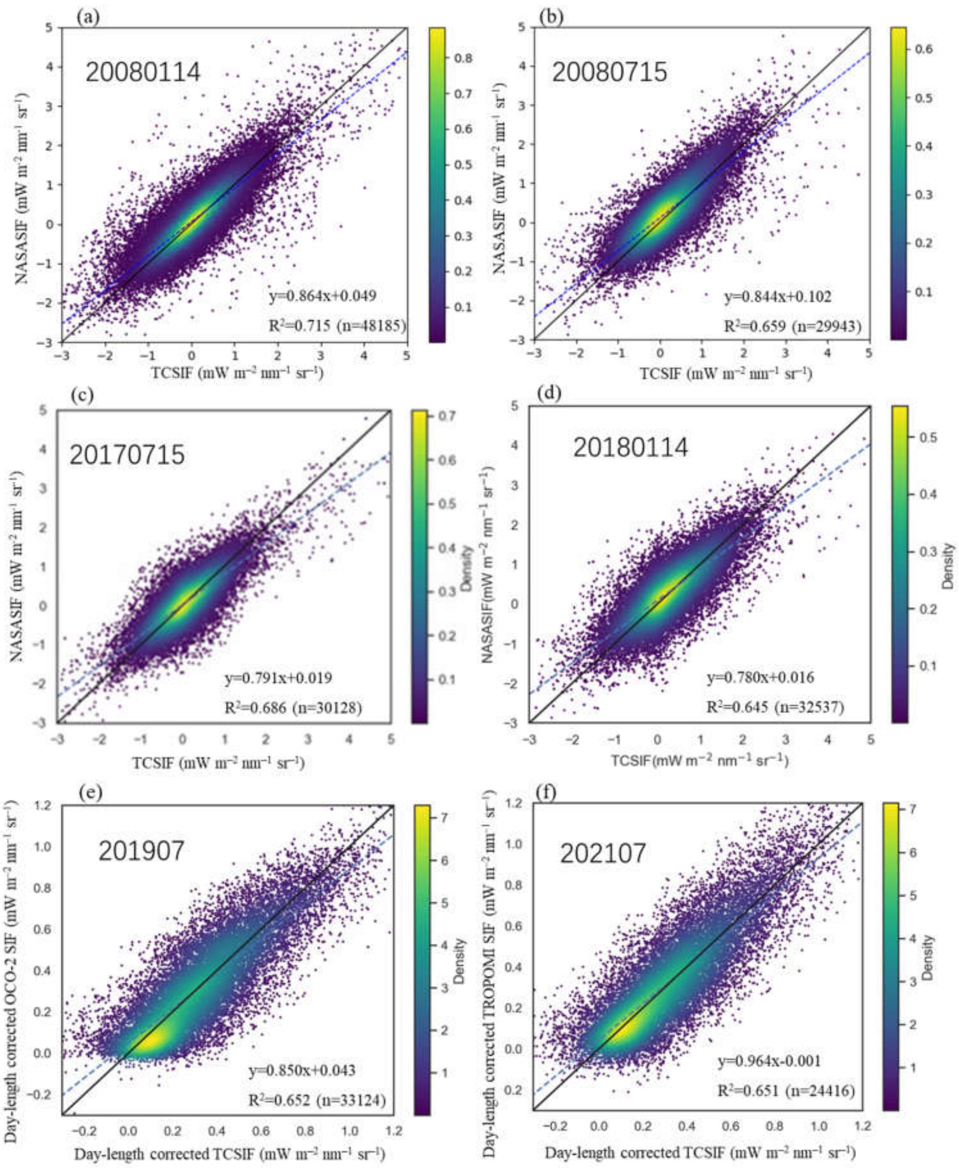


Figure 7. Comparison of TCSIF vs. NASA SIF on 14 January (a) and 15 July (b) in the year 2008**, 15 July 2017 (c), and 14 January 2018 (d).** Comparison of TCSIF vs. OCO-2 in July 2019(e) and TCSIF vs. TROPOMI SIF in July 2021 (f). The comparison was made based on the level 2 product. Co-located pixels over land with a cloud fraction < 0.3 have been selected. The color of the scatter points represents the density of the points. The blue dotted line and the black solid line represent the line fitted based on the scatter points and the 1:1 line, respectively.

https://doi.org/10.5194/essd-2023-329

It can already be seen that the discrepancy between NASA SIF and TCSIF increases with time in the results. To compensate for the lack of NASA SIF data, we have also added comparisons between TCSIF, OCO-2 SIF, and TROPOMI SIF in July 2019 and July 2021 to validate the latest results. **As detailed in major comment 2.**

Minor comment 13: Line 307. Consider moving these indirect results to a separate section to the very end or discussion.

Response: Thanks for this advice, we have moved the indirect comparison with GPP and NDVI to the discussion of the revised manuscript (Sec. 5.4), the comparison with other SIF products is moved to Sec.5.3. Instead, validation with GOME-2C, OCO-2, and TROPOMI was added in the result.

More direct validation of SIF to OCO/TROPOMI and reflectance to MODIS are required.

Thanks again. Please see the response for **major comment 2** for further validation.

Hahne, A.: Investigation on GOME-2 throughput degradation, 2012.

Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview, Atmos. Meas. Tech., 9, 1279–1301, https://doi.org/10.5194/amt-9-1279-2016, 2016.

Hu, J., Liu, L., Guo, J., Du, S., and Liu, X.: Upscaling Solar-Induced Chlorophyll Fluorescence from an Instantaneous to Daily Scale Gives an Improved Estimation of the Gross Primary Productivity, Remote Sensing, 10, 1663, 2018.

Hu, Jiaochan, et al. "Upscaling GOME-2 SIF from clear-sky instantaneous observations to all-sky sums leading to an improved SIF–GPP correlation." Agricultural and Forest Meteorology 306 (2021): 108439.

Jiang, D., et al. "Dynamic properties of absorbed photosynthetic active radiation and its relation to crop yield." Syst. Sci. Compr. Stud. Agric 18 (2002): 51-54.

Liu, Liangyun, Linlin Guan, and Xinjie Liu. "Directly estimating diurnal changes in GPP for C3 and C4 crops using far-red sun-induced chlorophyll fluorescence." Agricultural and Forest Meteorology 232 (2017): 1-9.

Liu, L., Liu, X., Chen, J., Du, S., Ma, Y., Qian, X., ... & Peng, D. (2020). Estimating maize GPP using near-infrared radiance of vegetation. Science of Remote Sensing, 2, 100009.

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop

https://doi.org/10.5194/essd-2023-329
series of satellites: instrument design, calibration, and level 1 data processing – an overview, Atmos. Meas. Tech., 9, 1279-1301, https://doi.org/10.5194/amt-9-1279-2016, 2016.

Pinardi, G., Van Roozendael, M., Hendrick, F., Richter, A., Valks, P., Alwarda, R., Bognar, K., Frieß, U., Granville, J., Gu, M., Johnston, P., Prados-Roman, C., Querel, R., Strong, K., Wagner, T., Wittrock, F., and Yela Gonzalez, M.: Ground-based validation of the MetOp-A and MetOp-B GOME-2 OClO measurements, Atmos. Meas. Tech., 15, 3439-3463, 10.5194/amt-15-3439-2022, 2022.

Sellers, P. J., Tucker, C. J., Collatz, G. J., Los, S. O., Justice, C. O., Dazlich, D. A., & Randall, D. A. (1994). A global 1 by 1 NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. International Journal of remote sensing, 15(17), 3519-3545.