



Development of the Long-term Harmonized multi-satellite SIF (LHSIF) dataset at 0.05° resolution (1995–2023)

Chu Zou^{1,2,3}, Shanshan Du^{1,2}, Xinjie Liu^{1,2}, Liangyun Liu^{1,2,3}

¹ Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100094, China

5 ² University of Chinese Academy of Sciences, Beijing 100049, China

³ International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China

Correspondence to: Liangyun Liu (liuly@radi.ac.cn)

Abstract. Solar-induced chlorophyll fluorescence (SIF) is a crucial proxy of photosynthetic processes in vegetation. In recent decades, advancements in remote sensing technology have facilitated long-term global SIF monitoring, significantly enhancing our understanding of vegetation dynamics on a global scale. Despite this progress, current SIF datasets face major challenges, including temporal inconsistencies among various satellite-derived products and a lack of long-term, high-resolution observations. In this study, we developed a “Long-term Harmonized SIF” (LHSIF) dataset spanning 1995 to 2023 with a fine spatial resolution of 0.05° by coordinating SIF satellite observations from GOME, SCIAMACHY, GOME-2, and OCO-2. Light use efficiency (LUE)-based spatial downscaling models were employed for each SIF product to generate fine-resolution global SIF maps. The long-term dataset was constructed using temporally corrected GOME-2A SIF (TCSIF) as a benchmark and was combined with a moment-matching normalization method for far-red SIF harmonization across satellite sensors from GOME, SCIAMACHY, and OCO-2. The resulting harmonization dataset exhibits a 45% reduction in overall error and a stable interannual increase ($0.42 \pm 0.13\% \text{ yr}^{-1}$) compared with a fluctuating decline ($-0.57 \pm 0.27\% \text{ yr}^{-1}$) of the original observations. This result strongly aligns with the growth rate of gross primary production (GPP, $0.47 \pm 0.03\% \text{ yr}^{-1}$) and is consistent with ground-based SIF observations ($R > 0.60$). Therefore, the long-term harmonized SIF dataset with a fine 0.05° resolution is a valuable tool for estimating global photosynthesis over extended periods. The LHSIF dataset is available at <https://doi.org/10.5281/zenodo.14854185>(Zou et al., 2025).

Keywords: Solar-induced chlorophyll fluorescence; Multi-satellite harmonization; Spatial downscaling

1 Introduction

25 Solar-induced chlorophyll fluorescence (SIF) is an optical signal naturally released by plants, closely linked to their photosynthetic dynamics (Damm et al., 2015; Mohammed et al., 2019; Porcar-Castell et al., 2014; Rascher et al., 2015; Zhang et al., 2016; Zhang and Peñuelas, 2023; Zhu et al., 2024). SIF has garnered significant attention due to its potential as a novel proxy for gross primary productivity (GPP) (Ryu et al., 2019), bridging the gap in our understanding of global photosynthetic processes (Beer et al., 2010; Anav et al., 2015; Chen et al., 2024).



30

Field-based studies have demonstrated strong correlations between SIF and GPP, highlighting its viability for large-scale ecosystem monitoring (Cui et al., 2017; Rossini et al., 2010). Following the publication of the initial global SIF map from the Greenhouse Gases Observing Satellite (GOSAT), interest in the SIF-GPP association greatly increased (Frankenberg et al., 2011; Guanter et al., 2012; Joiner et al., 2011). Subsequent satellite-based analyses have consistently revealed strong spatial and temporal correlations between SIF and GPP, showcasing remarkable alignment between SIF and GPP in terms of spatial distribution and seasonal variability (Anav et al., 2015; Li et al., 2018; Verma et al., 2017; Yang et al., 2015; Guanter et al., 2014; Zheng et al., 2024).

35

Studies conducted in diverse ecosystems, including needleleaf-dominated woodlands across the western United States (Zuromski et al., 2018) and arid landscapes of southwestern North America (Smith et al., 2018) have shown that SIF provides a superior approach compared to conventional vegetation indices, such as Enhanced Vegetation Index (EVI) and Photochemical Reflectance Index (PRI), in detecting interannual variations in GPP. Further, SIF has managed to capture subtle forest disturbances undetectable by traditional satellite products. However, these results are based on coarse-resolution SIF datasets from the Global Ozone Monitoring Experiment (GOME)-2, leading to inaccuracies related to spatial scale mismatch. Additionally, the SIF-GPP link is dependent on vegetation type, emphasizing the critical need for SIF datasets with higher spatial resolution to comprehensively evaluate the SIF-GPP interaction.

40

45

Long-term global SIF observations are important for analyzing the vegetation functions and changes under different climatic conditions. Multiple high-spectral-resolution satellite missions have provided publicly available global SIF products since 1995. The earliest records originated from the GOME sensor on European Remote sensing Satellite (ERS) in 1995, followed by the the SCanning Imaging Absorption spectroMeter for Atmospheric Cartography (SCIAMACHY) onboard Environmental Satellite (EnviSat) in 2003. However, these sensors had relatively short operational lifespans, ceasing operations in 2003 and 2012, respectively. The GOME-2 sensor onboard the MetOp-A satellite, launched in January 2007, operated until November 2021; this sensor provided the longest SIF time series to date (Joiner et al., 2013). The Orbiting Carbon Observatory(OCO)-2 satellite, launched in 2014, features exceptionally high spatial resolution and has been validated through synchronized airborne campaigns (Sun et al., 2017) to ensure the reliability of resulting SIF products. Recent studies highlight the potential of the TROPOMI sensor onboard Sentinel-5P (Koren et al., 2018; Wen et al., 2020), but its SIF products are currently constrained to a relatively short time series (May 2018 to April 2021).

50

55

Despite the availability of multiple satellite SIF products, most have a temporal coverage shorter than 10 years. While the

60



GOME-2 dataset provides the longest observational record to date, it is impacted by sensor degradation, leading to significant uncertainties in photosynthesis analyses based on GOME-2A SIF products (Parazoo et al., 2019). Yang et al. (2018) reported a rising trend in EVI alongside a declining trend in SIF, which they attributed to a reduction in ecosystem photosynthesis. However, Zhang et al. (2018a) argued that this conclusion was impacted by the deterioration of the GOME-2A instrument
65 itself. Further research by Koren et al. (2018) confirmed that the decline in SIF persisted even after correcting for sensor degradation (Koren et al., 2018). The SIFTERv2 product (Van Schaik et al., 2020) employed in Koren's study was simply corrected using linear models; the reliability of SIFTERv2 decreased significantly after 2016, limiting its application for long-term trend analysis.

70 Wang et al. (2022) attempted to create a temporally corrected long-term SIF product (LT_SIFc*) by correcting the degradation trends in gridded GOME, SCIAMACHY, and GOME-2 SIF products. However, the nonlinear characteristics inherent in the retrieval methodology and subsequent processing procedures (e.g., zero-bias correction and quality filtering) presented a challenge. The bias in sensor observations was not linearly transferred to the SIF product. This approach risks introducing inaccuracies, because the trend corrected in the SIF product may not correspond to the true sensor degradation. Recently, the
75 temporally corrected GOME-2A SIF dataset (TCSIF) included a calibration of the radiance measurements of GOME-2A using a pseudo-invariant method (Zou et al., 2024). This correction effectively eliminates the influence of sensor degradation over time, providing a robust benchmark for generating long-term harmonized SIF products.

In addition, large discrepancies have been observed between different SIF products (Parazoo et al., 2019). These temporal
80 inconsistencies may arise from differences in retrieval algorithms, absolute radiometric calibration errors, instrumental artifacts, directional effects, and differences in satellite overpass times and footprint sizes (Zhang et al., 2018c; Bacour et al., 2019). To address these challenges, researchers have used normalization methods based on observations during the overlapping period, such as cumulative distribution function (CDF), to align data between sensors. Wen et al. (2020) harmonized SCIAMACHY and GOME-2 data, creating a harmonized SIF dataset spanning 2002–2018. Wang et al. (2022) extended the time series to
85 1995 by incorporating the GOME SIF dataset.

A precise, reliable, harmonized, and global high-resolution SIF dataset is not yet available for long-term vegetation monitoring. Here, we employed a straightforward approach using a moment-matching method to normalize the SIF datasets from various
90 satellites. The observational discrepancies were normalized using correction coefficients based on the mean and standard deviation of the common time series of the datasets. SIF datasets obtained from GOME, SCIAMACHY, GOME-2, and OCO-2 instruments were synthesized to generate a comprehensive SIF dataset that spans 1995 to the present. Compared to previous



studies, we used a GOME-2 dataset with rigorous sensor degradation correction as the benchmark. The result is the longest multi-satellite harmonized SIF dataset to date (1995–2023). Additionally, we performed light use efficiency (LUE)-based spatial downscaling on the coarse spatial resolution dataset derived from the satellite SIF products. This downscaling reduced the spatial difference between satellite-derived SIF and ground-based measurements of SIF and GPP, thereby facilitating our understanding of vegetation photosynthesis at the global scale.

2 Method and materials

2.1 Satellite-based SIF datasets

2.1.1 GOME SIF

GOME, which was launched in 1995 on the ERS-2 satellite of the European Space Agency (ESA), was initially developed to measure the column densities of ozone and nitrogen dioxide (Hahne et al., 1993). GOME's channel 4 operates within a spectral range of 590–790 nm, achieving a spectral resolution of about 0.5 nm for far-red SIF retrieval. Although GOME is characterized by a relatively low spatial resolution of 320 km × 40 km, it provides the earliest available record of SIF data. The GOME SIF product utilized in this research is the daily averaged SIF signal at 740 nm, which is retrieved by data-driven algorithms (Joiner et al., 2019). The dataset spans the period from July 1995 to June 2003.

2.1.2 SCIAMACHY SIF

The SCIAMACHY instrument was in operation from 2002 to 2012 onboard ESA's Envisat satellite, overlapping with the timeframe of GOME. The instrument enhanced GOME's capabilities by offering a finer spatial resolution of 30 km × 60 km. The comparable spectral ranges and spectral resolutions of SCIAMACHY and GOME allowed for the use of analogous techniques for SIF retrievals. The SCIAMACHY SIF products we used were retrieved using the same data-driven algorithms and fitting window (734–758 nm) as those used for GOME. Daily SCIAMACHY SIF datasets at 740 nm from January 2003 to April 2012 were employed (Joiner et al., 2021).

2.1.3 GOME-2A SIF

As a successor to GOME, GOME-2 is part of EUMETSAT's MetOp satellite series, with three satellites (MetOp-A, B, and C) launched between 2007 and 2018. GOME-2 improved upon its predecessor by providing enhanced spatial resolution (40 km × 40 km or 80 km × 40 km, contingent upon the specific platform utilized). The GOME-2A SIF datasets were obtained from the MetOp-A satellite, launched in 2007 and operating until 2021.

Research has shown apparent differences between GOME-2A SIF products using different retrieval methods (Parazoo et al.,



120 2019). For instance, SIF retrieval using a fitting window of 720–758 nm and a backward elimination algorithm (Köhler et al.,
2015) yields values up to twice as large as the retrievals using a 734–758 nm window (Joiner et al., 2013). The GOME-2A SIF
dataset used in this study was retrieved using the same data-driven algorithm and fitting window as Joiner et al. (2013),
ensuring consistency with GOME and SCIAMACHY SIF. Furthermore, the GOME-2 SIF we use has undergone correction
for sensor degradation and was found to avoid spurious trends caused by instrument deterioration (Zou et al., 2024). Therefore,
125 this temporal-corrected GOME-2A SIF dataset with degradation correction is used as a benchmark to harmonize the data from
the other three sensors.

2.1.4 OCO-2 SIF

OCO-2 was a satellite mission launched by the National Aeronautics and Space Administration in 2014. Unlike earlier missions,
OCO-2 focuses on small target areas, attaining a considerably greater spatial resolution of approximately 1.3 km × 2.3 km.
130 The spectral range of OCO-2 extends from 757 to 775 nm, facilitating the initial SIF retrievals at 757 nm and 771 nm. Drawing
upon the empirical correlation of SIF across various wavelengths, the product offers daily global SIF at 740 nm (Oco-2/Oco-
3 Science Team, 2020). The SIF datasets from OCO-2 and GOME-2 have eight overlapping years (2014 to 2021). As a result,
a thorough comparison and validation of the consistency can be conducted between the two datasets.

135 The product specifications and sensor information are listed in Table 1. This study resampled the orbital SIF data from different
satellites into global gridded datasets of varying sizes according to the footprint and the global coverage of the satellites.
Satellite-derived SIF measurements from GOME, SCIAMACHY, GOME-2, and OCO-2 were aggregated into monthly maps
with grid sizes of 1°×1°, 1°×1°, 0.5°×0.5° and 1°×1°, respectively.

140 **Table 1 Information on multiple satellite SIF datasets used to construct long-term SIF products.**

Satellite/Sensor	Temporal range	Footprint Size (km ²)	Overpass time	Swath width (km)	Wavelength (nm)	Grid size	Reference
ERS-2/ GOME	1995.07–2003.06	40×320	10: 30	960	740	1°×1°	(Joiner et al., 2019)
Envisat/ SCIAMACHY	2003.01–2012.04	30×240 /30×60	10: 00	960 /240	740	1°×1°	(Joiner et al., 2021)
MetOp-A/ GOME-2	2007.01–2021.11	40×80 /40×40	9: 30	1920 /960	740	0.5°×0.5°	(Zou et al., 2024)
OCO-2	2014.09–2023.12	1.3×2.2	13: 30	10.3	757/771	1°×1°	(Oco-2/Oco-3 Science Team, 2020)

2.2 Spatial downscaling



An LUE-based model was used for downscaling the gridded SIF datasets with coarse spatial resolutions. Assuming that SIF can be represented using the LUE model in a manner that is comparable to GPP (Berry et al., 2012; Guanter et al., 2014; Damm et al., 2015), then:

$$SIF = PAR \times fPAR \times SIF_{yield} \quad (1)$$

where SIF_{yield} is the fluorescence quantum yield, which is influenced by hydric and thermic stresses. $fPAR$ represents the fraction of photosynthetically active radiation (PAR) that is absorbed by vegetation, which exhibits a positive correlation with vegetation indices. Assuming that PAR is uniformly distributed over small areas and can be considered constant, then equation (1) can be further expressed as (Duveiller and Cescatti, 2016; Duveiller et al., 2020) :

$$SIF \approx b_0 \times f(NIRv) \times f(VPD) \times f(AT) \quad (2)$$

where $NIRv$ is the near-infrared reflectance of vegetation, VPD represents vapor pressure deficit (accounts for the effect of hydric stresses), and AT is the air temperature at 2m (accounts for the impact of thermic stresses). A quadratic function, sigmoid function, and Gaussian function with unknown coefficients were used to express $f(NIRv)$, $f(VPD)$ and $f(AT)$, respectively, as follows:

$$SIF \approx b_1 NIRv^{b_2} \times \left[\frac{1}{(1 + \exp(b_3(b_4 - VPD)))} \right] \times \left[\exp\left(-0.5 \left(\frac{AT + b_5}{b_6}\right)^2\right) \right]. \quad (3)$$

The unknown coefficients b_1 to b_6 in Eq. (3) can be determined by a nonlinear iterative approach. Here, we implemented this approach using the “L-BFGS-B” algorithm (Byrd et al., 1995).

$NIRv$, VPD , and AT were the three driving variables of the spatial downscaling model. $NIRv$ datasets characterized by a spatial resolution of 0.05° were partially derived from the Advanced Very High Resolution Radiometer (AVHRR) (Jeong et al., 2024) for 1995–2021, while the $NIRv$ for 2022–2023 were calculated using MODIS MCD43C4 nadir reflectance (Schaaf and Wang, 2021). AT and VPD data for 1995–2023 were obtained from the TerraClimate product with an original spatial resolution of $1/24^\circ$ (Abatzoglou et al., 2018). The driving variables were aggregated to coarse spatial resolutions ($0.5^\circ \times 0.5^\circ$ for GOME-2 and $1^\circ \times 1^\circ$ for other satellites) for training the LUE model described in Eq. (3) and to estimate the coefficients (b_1 to b_6). The 50 closest neighbors of the center pixel were utilized to train the LUE model in an 11×11 sliding window. Subsequently, SIF datasets, characterized by a spatial resolution of 0.05° , were produced by inputting the 0.05° driving variables into the trained model. Since the coefficients (b_1 to b_6) were computed for each coarse-resolution pixel, gridded artifacts may appear in the final product. To further ensure smooth spatial transitions, for each high-resolution pixel within a coarse-resolution pixel, a 3×3 block of coarse-resolution pixels was selected, and nine sets of six coefficients (b_1 to b_6) were computed. Meanwhile, a fine-resolution (0.05°) weighting grid was established within the 3×3 low-resolution pixels using a two-dimensional Gaussian function with a standard deviation of 15 km. The final downscaled result for each high-resolution pixel was obtained as the weighted average of the nine sets of model-predicted values, following the approach of Duveiller



and Cescatti (2016).

175 2.3 Moment-matching method

A moment-matching method was used in this study for the normalization of the SIF data from different satellites. Moment-matching is a statistical technique used to align the distributional characteristics of different datasets by adjusting their statistical moments. We used the average and standard deviation of the original datasets (D_o) and target datasets (D_t) for the normalization of the original time series:

$$180 \quad D_i = \frac{D_{o_i} - \mu_{D_o}}{\sigma_{D_o}} \cdot \sigma_{D_t} + \mu_{D_t} \quad (4)$$

where μ_{D_t} and σ_{D_t} are the mean and standard deviation of the target dataset, respectively. μ_{D_o} and σ_{D_o} are the mean and standard deviation, respectively, of the original dataset. D_i is the corrected value for the i -th temporal point of the corrected dataset. Considering the parameters' spatial representativeness and measurement noise influence, the matching process was not applied globally or on a pixel-by-pixel basis but rather within a sliding window at the same scale used for the downscaling process.

Further, the discrepancy between the two SIF time series can be quantified by the mean squared difference (MSD):

$$\text{MSD} = \frac{1}{n} \sum_{i=1}^n (D_{1_i} - D_{2_i})^2 \quad (5)$$

where D_1 and D_2 are the SIF time series of the two SIF datasets to be compared. i represents the i -th month of the chosen period. Furthermore, Eq. (1) can be broken down into three terms (Bacour et al., 2019):

$$\text{MSD} = (\overline{D_1} - \overline{D_2})^2 + (\sigma_{D_1} - \sigma_{D_2})^2 + 2 \sigma_{D_1} \sigma_{D_2} (1 - r) \quad (6)$$

where $\overline{D_1}$ and $\overline{D_2}$ are the expected values of the time series, while σ_{D_1} and σ_{D_2} signify the respective standard deviations. Additionally, r is the Pearson correlation coefficient that quantifies the relationship between the datasets. The first and second terms in the formula represent the square of the mean deviation (denoted as $bias^2$) and difference in standard deviation (denoted as $variance^2$) between the corrected datasets and the target datasets. The final term quantifies the inconsistency of the linear correlation between the two datasets (denoted as phase).

The GOME-2A SIF dataset with sensor degradation correction is used as a reference for the normalization of all other satellite SIF datasets based on observations during the overlap period. The normalization of GOME data relied on the SCIAMACHY dataset, which had been previously normalized with GOME-2 data. Fig. 1 shows the processing flow of this study.

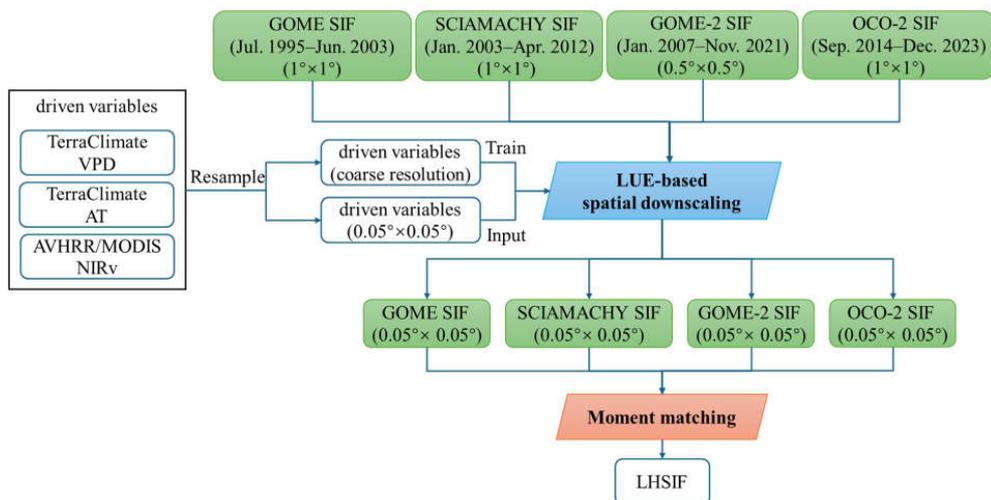


Figure 1. The workflow for spatial downscaling and temporal alignment for generating the LHSIF products from 1995 to 2023.

2.4 Datasets for validation and comparison analysis

2.4.1 Boreal Ecosystem Productivity Simulator GPP

205 The Boreal Ecosystem Productivity Simulator (BEPS) is an ecological process model that integrates vegetation parameters with meteorological data to simulate ecosystem productivity. We used the GPP dataset generated by the BEPS model for 1995–2019 (Weimin Ju and Zhou., 2021). The original spatial resolution of the dataset is $0.072727^\circ \times 0.072727^\circ$, providing fine-scale insights into productivity dynamics. This high-resolution dataset allows for detailed spatiotemporal analysis and facilitates comparisons with downscaled SIF datasets in this study to help our understanding of ecosystem carbon dynamics.

2.4.2 Long-term satellite SIF products

210 The LT_SIFc* dataset provides long-term SIF retrievals corrected for temporal inconsistencies between GOME, SCIAMACHY, and GOME-2 SIF datasets (Wang et al., 2022). A CDF method was employed for the harmonization of different SIF datasets, and the LUE-based model was used for spatial downscaling. The LT_SIFc* dataset spans 1995 to 2018 at a spatial resolution of $0.05^\circ \times 0.05^\circ$.

215

The SIF_005 dataset is a SIF product spanning 2003 to 2017, with a spatial resolution of $0.05^\circ \times 0.05^\circ$ (Wen et al., 2020). This product integrates data from SCIAMACHY and GOME-2 SIF datasets, and it is downscaled using a machine learning-based method. The v2.2 (trend_corrected) version was utilized in this study; the original SIF dataset used for this version has been preliminarily corrected for temporal degradation.

2.4.3 AVHRR vegetation indices

220



Global NDVI and NIRv datasets derived from the AVHRR sensors were utilized in this study. These datasets have been rigorously corrected for inconsistencies among different AVHRR sensors to ensure a temporally consistent and reliable data product from 1995 to 2021 (Jeong et al., 2024). The correction process involved addressing sensor-specific calibration differences, orbital drift, and atmospheric effects, thereby enabling robust comparisons across time and space.

225 2.4.4 Ground-based observations

Ground-based SIF and GPP observations were integrated into this study to validate and enhance the interpretation of satellite-derived datasets. Specifically, FLUXNET GPP observations were employed, which are based on in-situ measurements from a global network of flux towers distributed across diverse ecosystems (Pastorello et al., 2020). FLUXNET sites with more than five years of data were grouped into climate zones and vegetation functional types (see Fig. S1 for site distribution and types).

230 The field “GPP_DT_VUT_REF” was used.

In addition, tower-based SIF observations from the ChinaSpec network, including sites such as DM, GC, HL, XTS, and AR (Zhang et al., 2021), were used to validate the accuracy and spatiotemporal consistency of the long-term SIF dataset generated in this study. The locations and cover types of the ChinaSpec sites used are listed in Table S1. These datasets provided robust
235 benchmarks for assessing the consistency and reliability of the long-term SIF dataset.

3 Results

3.1 Downscaled SIF dataset

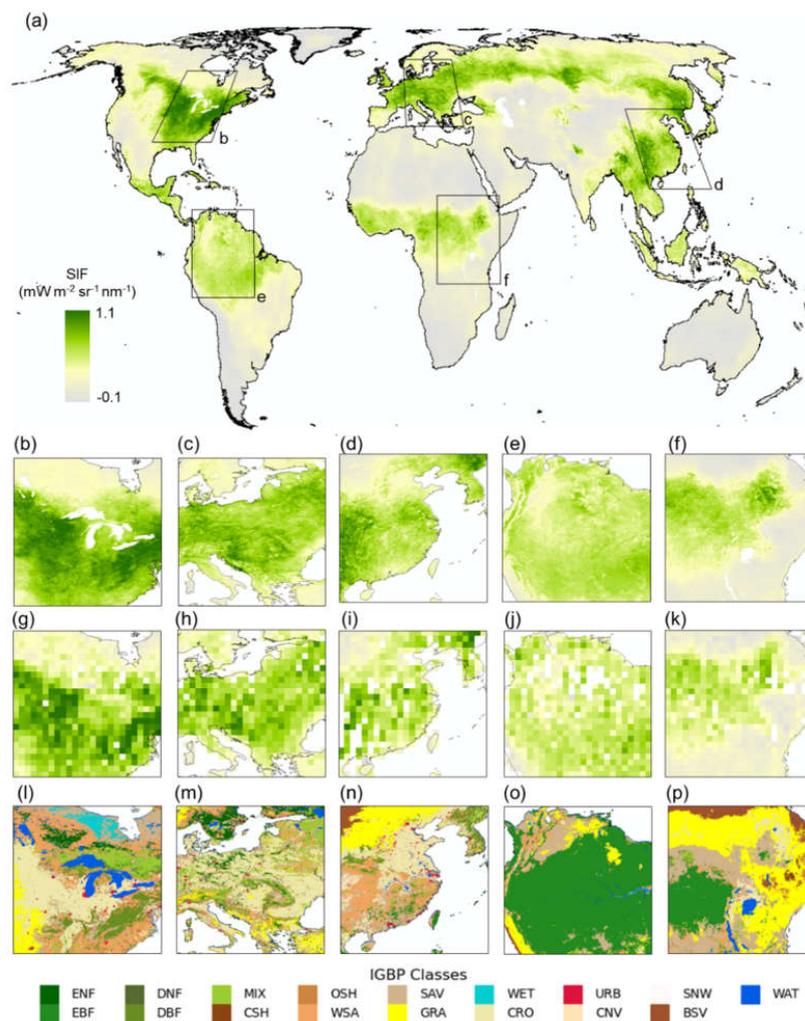
Figure 2 provides a comparative analysis of downscaled SIF datasets at fine spatial resolution (0.05°) and their coarse-resolution (1°) counterparts (derived from GOME). The top two rows (panels a–f) illustrate the enhanced spatial variability achieved through the downscaling process, revealing finer vegetation patterns and distinct intensity gradients. The downscaled
240 SIF datasets render subtle patterns in SIF more apparent compared to the original coarse-resolution data (panels g–k). Additionally, the downscaling method, which incorporates neighborhood-based pixel searching, effectively fills in data gaps in the original data while preserving spatial continuity. These enhanced spatial details align well with land cover types (panels l–p), highlighting the capability to capture sub-grid heterogeneity.

245

Figure 3 further demonstrates the distribution of monthly SIF before and after spatial downscaling (using GOME as an example). Results for other satellites are provided in Figs. S2–S4. The spatially downscaled SIF ($0.05^\circ \times 0.05^\circ$) was reaggregated to $1^\circ \times 1^\circ$ for a lower resolution comparison with the original SIF. The results show that most of the SIF values of the reaggregated pixels align well with the original SIF values, i.e., they are concentrated near the 1:1 line. The range of SIF



250 values remains mostly unchanged, proving that the training of the LUE downscaling model successfully established the relationship between the driving variables and SIF.



255 **Figure 2.** Spatially downscaled SIF maps (a-f) compared to the original SIF maps at a coarser resolution (g-k). SIF data from GOME observations in July 1996 are shown as an example. The bottom row (l-p) shows the corresponding International Geosphere-Biosphere Programme land cover classifications obtained from the MCD12C1 product. Panels b, g, and l depict North America; c, h, and m focus on Europe; d, i, and n depict East Asia (centered on China); e, j, and o represent the Amazon Basin; and f, k, and p show Sub-Saharan Africa.

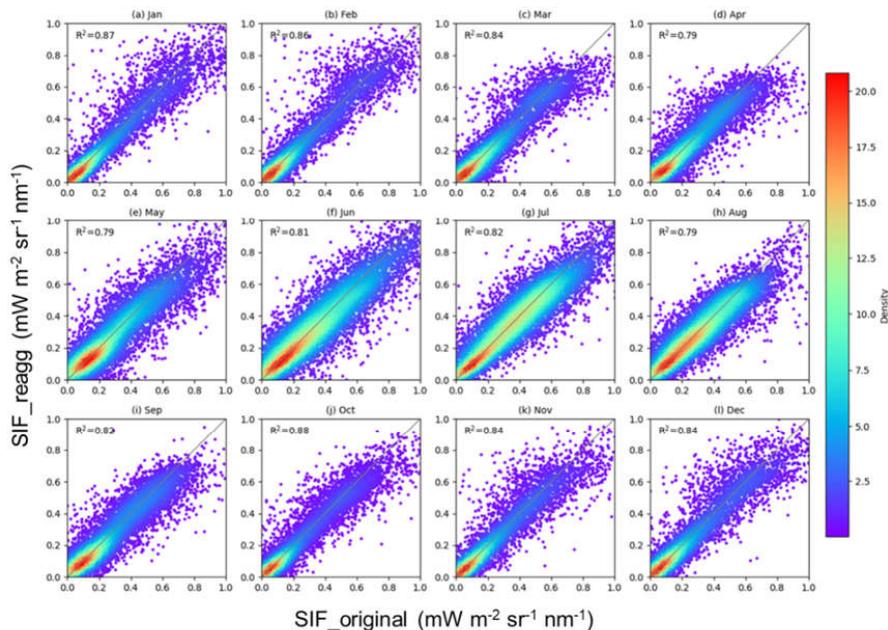


Figure 3. The relationship between the reagggregated GOME SIF (SIF_{reagg}) and the original GOME SIF ($SIF_{original}$) for 1998 (by month).

260

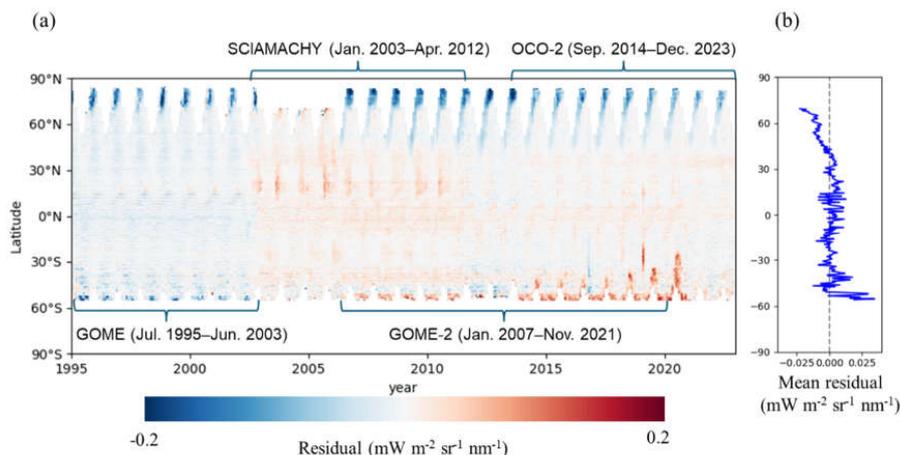


Figure 4. The (a) time series and (b) temporal average of the latitudinally distributed residual generated by the LUE-based downscaling model. The residuals are calculated as the difference between the reagggregated SIF (SIF_{reagg}) and the original SIF ($SIF_{original}$). For consistency, the coarse spatial resolution SIF datasets were uniformly resampled to $0.5^\circ \times 0.5^\circ$. Only data for latitudes below $70^\circ N$ are shown in (b).

265

The temporal and spatial distributions of the spatial downscaling residuals were analyzed (Fig. 4). The residual was calculated as the difference between SIF_{reagg} and $SIF_{original}$. As shown by the temporally averaged residuals (Fig. 4 (b)), the absolute mean residuals for most locations are less than $0.05 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$ for regions below $70^\circ N$. This result highlights the global applicability of the LUE-based downscaling method.



270 **3.2 Temporal harmonization**

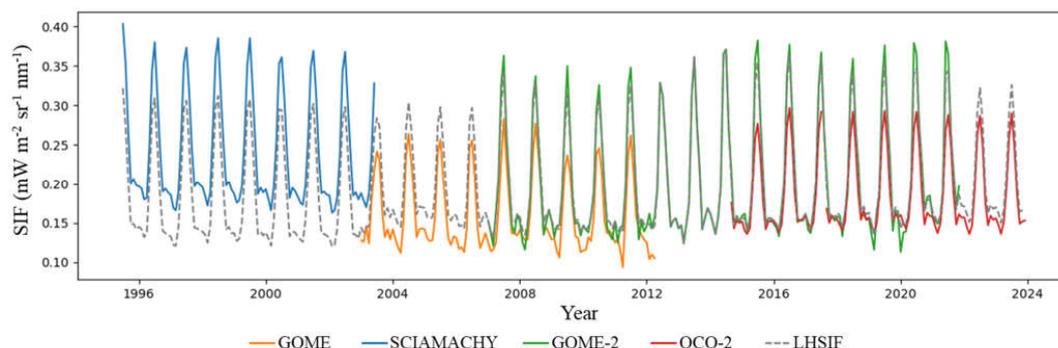
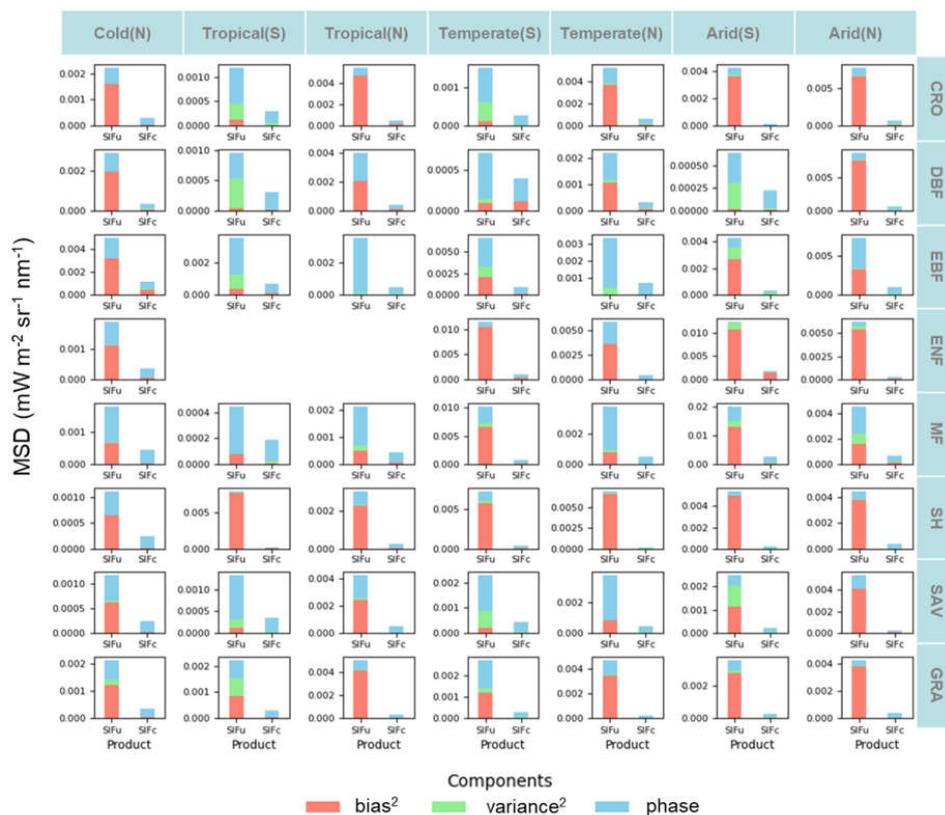


Figure 5. Global-averaged SIF time series derived from GOME (blue), SCIAMACHY (orange), GOME-2 (green), and OCO-2 (red), along with the long-term harmonized SIF time series (LHSIF, gray dotted line), which aligns the satellite datasets based on overlapping periods.

275 Figure 5 illustrates the original SIF datasets from individual satellites and the long-term harmonized SIF dataset spanning 1995–2023 after normalization. Moment matching was conducted in the overlapping periods to normalize the differences among the SIF datasets. Discrepancies in the value range across the original data sources, while after normalization, the systematic bias between the SIF measurements from different satellites was effectively eliminated. The abrupt changes in the interannual maximum and minimum values across sensors were corrected without altering the SIF trends provided by
280 individual sensors. The result is a more accurate and reliable temporal trend in global vegetation.

Error analyses were performed for different climatic zones and plant functional types. Fig. 6 displays the comparison between GOME-2 and the SCIAMACHY SIF before and after moment matching. The overall error (MSD) decreased by more than 45% after moment matching. For most cases, the difference in the average (bias, in red) is the dominant component of MSD between
285 GOME-2 SIF and SCIAMACHY SIF before moment matching. Throughout the Southern Hemisphere’s temperate and tropical zones, the discrepancies were primarily attributed to variations in variance (in green) and weak correlations (in blue). The MSD was significantly reduced after temporal correction, with a notable decrease in the proportion of bias². Only a small proportion of phase-related errors remain in the corrected dataset.



290 **Figure 6.** The mean squared difference (MSD) between GOME-2 SIF and SCIAMACHY SIF before (SIF_u) and after (SIF_c) moment matching. The results show the average conditions across different climatic zones and vegetation functional categories during 2007.

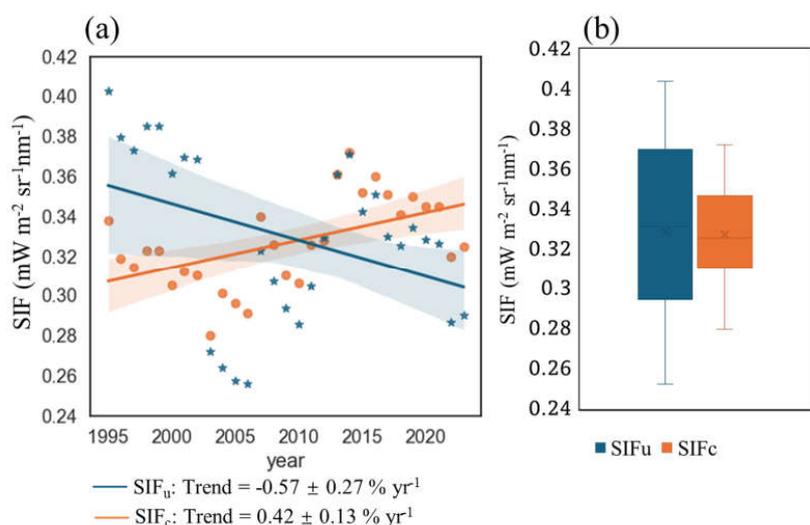
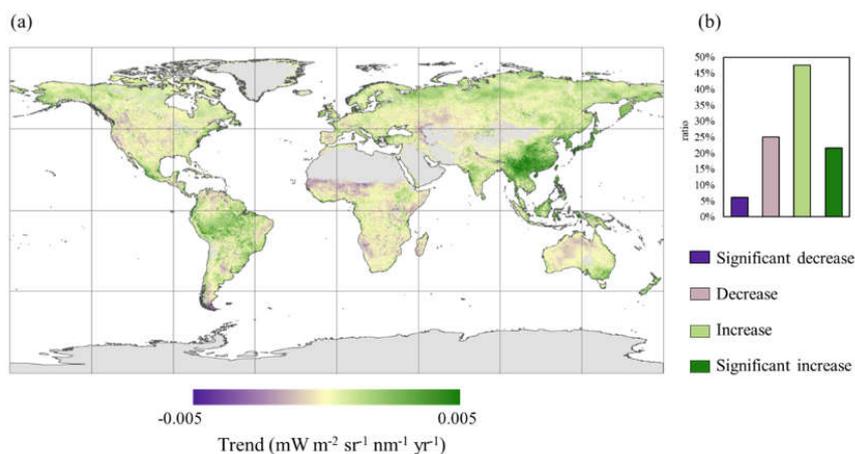


Figure 7. (a) Trend and (b) box plot of the yearly maximum global-averaged SIF of the combined time series before (SIF_u) and after (SIF_c) normalization during 1995–2023.



295 The annual maximums of the global-averaged SIF were used to investigate the fluctuation of the worldwide vegetation from 1996 to 2023. Significant interannual fluctuations were found for the SIF time series without moment matching, with an overall decline (blue line in Fig. 7 a). After normalization, the uncertainty in the interannual trend was reduced by half (from 0.27% to 0.13%), revealing a growth rate of $0.42\% \text{ yr}^{-1}$. The boxplot in Fig. 7 b further indicates a relatively small range in SIF values after temporal matching.



300 **Figure 8. (a) Map of trends in LHSIF for 1996–2023. (b) Percentage of areas in global vegetation covered by four different trend types (significant decrease: negative correlation and $p < 0.05$; decrease: negative correlation and $p \geq 0.05$; increase: positive correlation and $p \geq 0.05$; significant increase: positive correlation and $p < 0.05$).**

The interannual variability map of LHSIF from 1996 to 2023 was obtained by fitting SIF trends pixel-by-pixel (as shown in
305 Fig. 8 a). Significant increases in SIF are observed in vegetation-rich regions, such as the Amazon rainforest, the eastern and southern United States, Southeast Asia, and southeastern China. Tropical grasslands south of the Sahara have expanded, while SIF in the Congo Basin remains largely unchanged. Southern Africa exhibited a decline in SIF. Slight declines were also found in vegetated zones, such as high-latitude grasslands of the Northern Hemisphere and western United States. In regions with sparse vegetation, like Central Australia and Northwest China, SIF slightly declined or remained unchanged. Overall, SIF
310 growth occurred in over 69% of the world's vegetation, with 22% of the global vegetated area exhibiting notable increases between 1996 and 2023 (Fig. 8 b).

3.3 Validation and comparative analysis

315 The annual maximum LHSIF exhibited a positive trend ($0.42 \pm 0.13\% \text{ yr}^{-1}$), with data points clustering around the trendline (Fig. 9 a). The growth rate of LHSIF ($0.42\% \text{ yr}^{-1}$) closely aligns with that of BEPS GPP (Fig. 9 f, $0.47\% \text{ yr}^{-1}$), demonstrating LHSIF's stability in capturing long-term trends of GPP. LT_SIFc* also shows a positive trend but with a lower growth rate (Fig. 9 b). The SIF_005 product exhibits a negative trend during 2003–2017 in stark contrast to all other datasets (Fig. 9 c).



Although the spurious trends have been largely corrected for the original SIF products used by SIF_005, the long-term trend
320 remains suboptimal.

From 1995 to 2021, less pronounced trends were shown by AVHRR NDVI (Fig. 9 d, $0.18 \pm 0.02\% \text{ yr}^{-1}$) and NIRv (Fig. 9 e,
0.34 \pm 0.02% yr^{-1}) compared with SIF and GPP. Compared to SIF-based products, NDVI is more susceptible to interference
from vegetation canopy structure and non-photosynthetic processes; thus, it is less effective at capturing photosynthetic activity.
325 In this regard, LHSIF provides a more direct indication of photosynthesis and can supplement NDVI and NIRv in detecting
changes in GPP.

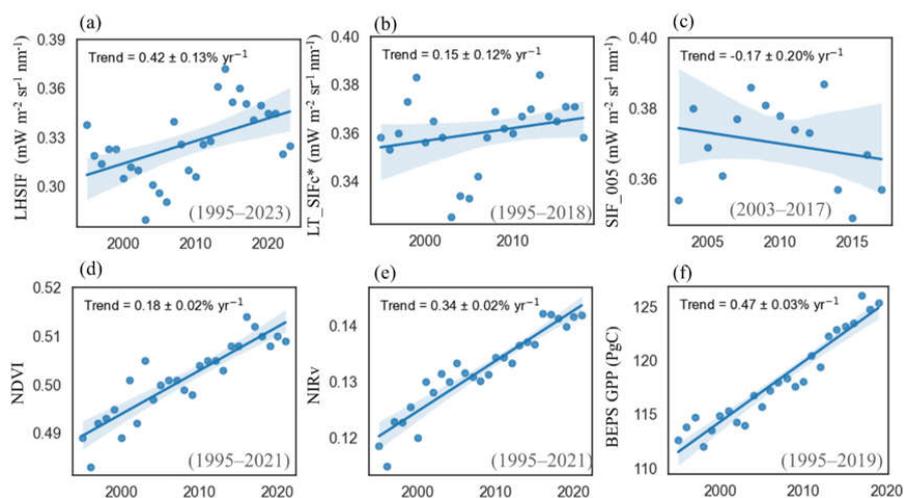


Figure 9. The yearly maximum (a) LHSIF compared with (b) LT_SIFc*, (c) SIF_005, (d) AVHRR NDVI, (e) AVHRR NIRv, as well as the annual total (f) BEPS GPP.

330 Figure 10 illustrates the contrasting relationships between LHSIF and GPP and between AVHRR NDVI and GPP. The LHSIF
product shows a strong ability to track GPP, especially for cropland and mixed forest types (Fig. 10a). In contrast, NDVI
consistently exhibits lower R^2 values (Fig. 10 b) and a more pronounced nonlinear relationship with GPP due to saturation
effects. Apart from a few groups in the Southern Hemisphere (such as grasslands in tropical and arid areas), where only a small
number of sites are available (see Fig. S1), SIF outperforms NDVI in most cases.

335

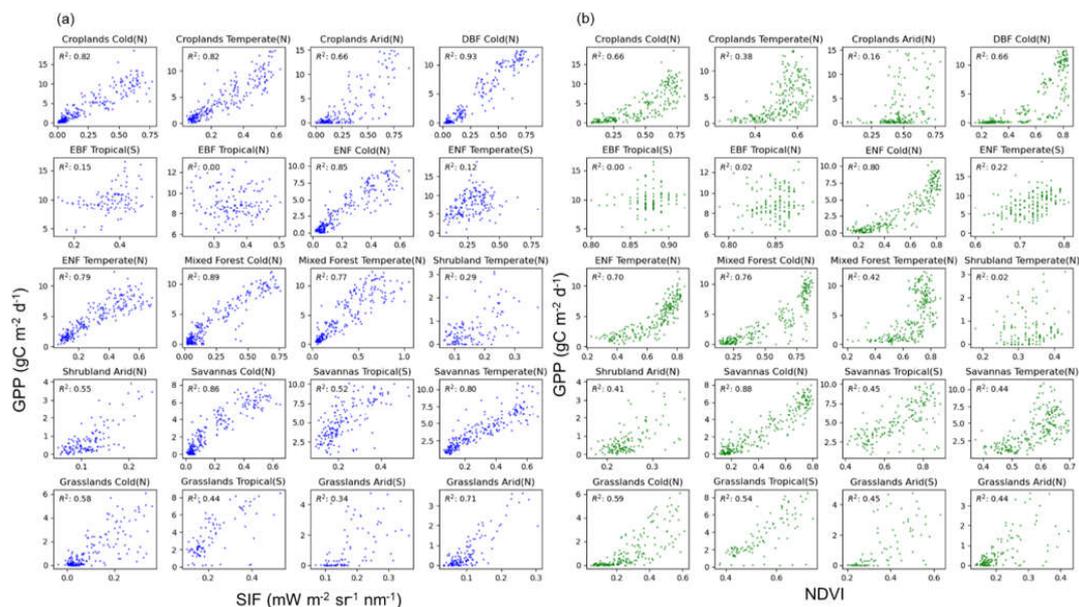
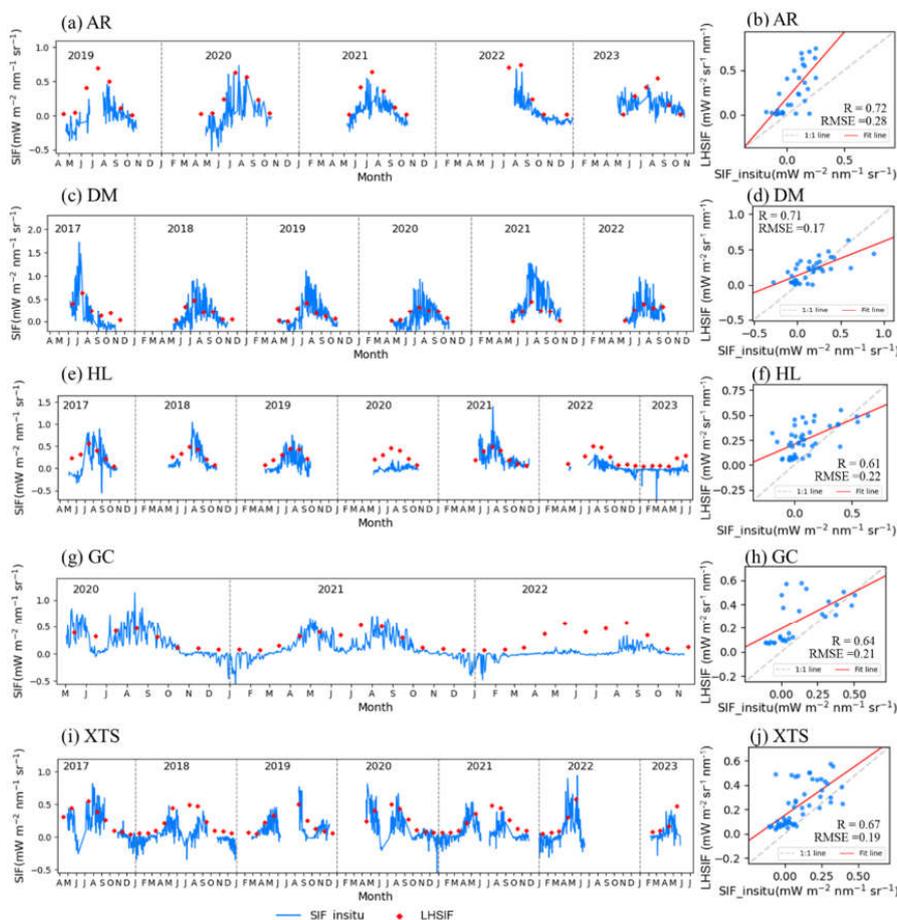


Figure 10. Comparison of long-term relationships between SIF vs. GPP and NDVI vs. GPP at FLUXNET sites. Only those flux tower sites that have accumulated over a decade of data were chosen and subsequently categorized according to their respective climate zones and vegetation types.

340 Additionally, comparisons were conducted between LHSIF and the tower-based SIF measurements at five ChinaSpec sites: AR, DM, GC, HL, and XTS (details of the locations and land cover types are provided in Table S1). To ensure consistent comparisons, the tower-based SIF at 760.6 nm was converted to 740 nm using an empirical correction factor of 1.48. Additionally, the original half-hourly tower-based SIF data were temporally upscaled to daily and monthly values with the aid of PAR and NDVI, following the method described by Hu et al. (2018). As a result, LHSIF demonstrated strong agreement with tower-based SIF measurements both temporally and spatially (Fig. 11). The consistency of the intra-annual variations was evident between LHSIF and the in-situ measurements for each site. As shown in the right panel, the monthly composite values are highly correlated, with most points clustering near the 1:1 line and correlation coefficients generally exceeding 0.6.

350 However, some deviations were observed. For example, at the Gucheng (GC) and Xiaotangshan (XTS) sites, which are characterized as wheat-maize rotation croplands, discrepancies occurred in June. During this month, tower-based SIF measurements recorded a trough when wheat was harvested and maize had yet to emerge. Due to spatial heterogeneity, LHSIF was unable to capture this phenomenon, resulting in a reduced correlation between LHSIF and in-situ measurements at these two sites. To highlight the overall correlation, the June data for these two sites were removed from the scatter plot (Fig. 11 h, j).



355

Figure 11. Comparison between LHSIF with the tower-based SIF measurements (SIF_insitu) at (a, b) AR, (c, d) DM, (e, f) HL, (g, h) GC, and (i, j) XTS sites. The temporal pattern of LHSIF compared with daily and monthly averaged SIF_insitu is shown in the left and right columns, respectively. The horizontal axis of the chart in the left column represents the first letters of the month names. Data points in June were excluded from the scatter plot for GC and XTS (h, j).

360 4 Discussion

4.1 Improvements in cross-sensor harmonization

In this study, we applied a moment-matching normalization method to harmonize cross-sensor SIF measurements. The steadily increasing interannual trend in the multi-sensor synthesized SIF time series demonstrates the validity of the temporal matching approach. This finding aligns with previous research (Wen et al., 2020; Wang et al., 2022) and further reinforces the reliability of the initial SIF products and the feasibility of the moment-matching normalization process for cross-sensor harmonization.

365

Unlike previous research, the temporal-corrected GOME-2A SIF (TCSIF) was used as a benchmark for harmonizing the time series (Zou et al., 2024). The TCSIF product systematically addressed the known pseudo-trends in GOME-2A data, and it



underwent a rigorous two-step validation process for both radiance and SIF. The interannual variation of TCSIF closely
370 resembles GPP, making it a reliable data benchmark. In contrast, the original SIF products used by SIF_005 (Wen et al., 2020)
may still suffer from incomplete trend correction of GOME-2 SIF and lead to a general decrease and significant interannual
fluctuations in SIF_005 during the 2003–2017 period (Fig. 9 c).

Additionally, adopted a different approach from that of Wang et al. (2022) to coordinate SIF data from multi-sensors.
375 Specifically, we used the GOME-2A SIF dataset, which is centrally positioned within the time ranges covered by the various
sensors, as the benchmark for harmonization. The temporal coverage of GOME-2A (2007–2021) includes overlapping periods
of more than five years with both SCIAMACHY and more than eight years with OCO-2. As a result, a single normalization
process is sufficient for aligning SIF of both instruments with GOME-2A SIF, effectively reducing propagation errors. In
contrast, the LT_SIFc* dataset uses GOME, the earliest time series, as the benchmark. This approach relies on a six-month
380 overlap between GOME and SCIAMACHY for the normalization of SCIAMACHY SIF, which is then used to calibrate
GOME-2A SIF. Such a multi-step calibration process may lead to cumulative errors.

The method used for temporal matching between different sensors is another critical factor that influenced the results. Improper
approaches can introduce discrepancies during the matching process. Previous research commonly used the CDF matching
385 method for temporal harmonization (Wang et al., 2022; Wen et al., 2020). While this method theoretically ensures precise
cross-sensor alignment in the absence of observational errors and discrepancies between satellites, it is not a suitable solution
when training data are insufficient, the SIF dataset is noisy, or sensors use different distribution functions. These observations
are particularly relevant for sensors with short overlap, such as the SIF data from GOME and SCIAMACHY (six-month
overlap). This short duration results in an insufficient amount of high-quality data to effectively support such transformations.

390

The approach proposed adopted in this study is robust compared to CDF matching. By modifying only the mean and standard
deviation of the SIF data distribution, the method avoids overfitting and enhances stability. This less stringent approach ensures
harmonization while preserving the inherent variability of the datasets. Therefore, the LHSIF dataset provides an
unprecedented long-term harmonized SIF dataset that is theoretically reliable and demonstrates a reasonable temporal trend.

395 4.2 Limitations and future perspectives

Our investigation shows that the moment-matching normalization approach is an excellent way to reduce the disparities across
sensors, providing a unified reference framework with the longest time series to date. Despite its high performance, it is crucial
to remember that this statistical methodology is not a rigorous calibration procedure. The most robust approach for cross-
sensor calibration is based on pseudo-invariant calibration sites (PICs) located in non-vegetated areas (Markham and Helder,



400 2012; Khakurel et al., 2021). This method has been successfully applied to the normalization and long-term monitoring of reflectance data and vegetation index products (Angal et al., 2013; Mishra et al., 2014; Jeong et al., 2024; Tavora et al., 2023). However, in commonly used PICs, such as deserts and water surfaces, SIF signals are inherently weak and highly susceptible to noise that can cause significant uncertainty in PIC-based calibration for SIF applications.

405 Additionally, our downscaling approach follows the methodology proposed by Duveiller et al. (2016, 2020), where NIR_v is used in the LUE model instead of NDVI to enhance the model's interpretability for SIF (Badgley et al., 2017). Nevertheless, some explanatory variables remain unaccounted for. For instance, incorporating PAR could improve model interpretability under cloudy conditions (Ryu et al., 2018). However, discrepancies in overpass time and scale effects can cause inconsistencies between PAR and SIF products, which may increase uncertainties in the downscaling model. Furthermore, incorporating
410 fluorescence escape efficiency (Ryu et al., 2019) and topographic factors (Chen et al., 2022; Tao et al., 2024) into the downscaling model could further enhance its performance.

Model selection is ultimately more critical than the choice of input variables (Duveiller et al., 2020). Previous research on spatial downscaling was predominantly using purely empirical machine-learning approaches (Gentine and Alemohammad,
415 2018; Wen et al., 2020; Hong et al., 2022; Lu et al., 2024). An alternative strategy redistributes the initial global downscaling results based on the original coarse-resolution SIF values and retains the characteristics of the original observational signal to the greatest extent possible (Ma et al., 2022; Chen et al., 2025). Our experimental results confirm that our downscaled SIF products also remain consistent with the original signals (Fig. 4). In addition, the LUE-based approach incorporates physiological constraints, ensuring that the downscaled SIF values remain within a reasonable range compared to traditional
420 machine-learning models.

Another type of long-term SIF datasets have been generated by temporally extrapolating SIF observations based on machine-learning methods. These datasets provide more than two decades of high-temporal-resolution data beyond the monthly scale (Zhang et al., 2018b; Li and Xiao, 2019). However, such datasets predominantly depend on model-driven predictions
425 constrained by satellite observation periods, rather than being based on actual observational data, which is fundamentally different from the approach we employed. Currently, the temporal resolution of purely observation-based enhanced SIF products that span longer than 20 years remains constrained at the monthly scale, largely due to noise in the satellite SIF products. Overcoming this limitation will require further refinement of existing downscaling models, paving the way for future products to achieve a resolution of 16 days or higher.



In this study, we developed a long-term harmonized SIF dataset (LHSIF) spanning 1995 to 2023. SIF datasets from various satellites were harmonized using multi-sensor SIF observations through a moment-matching normalization approach, using the temporally corrected GOME-2A SIF dataset as a benchmark. An LUE-based model was used for spatial downscaling, yielding a fine resolution of 0.05° with an absolute mean residual less than $0.05 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$.

435

Our analysis demonstrated that the harmonized dataset significantly reduced overall errors by more than 45% and exhibited a stable interannual increase ($0.42 \pm 0.13\% \text{ yr}^{-1}$); in contrast, the original dataset displayed a variable decline ($-0.57 \pm 0.27\% \text{ yr}^{-1}$). The interannual trend of LHSIF closely aligns with the growth of GPP ($0.47 \pm 0.03\% \text{ yr}^{-1}$) and demonstrates superior temporal and spatial consistency compared to NDVI. Validation against ground-based SIF observations ($R > 0.6$) further underscores the reliability of the harmonization approach and the dataset's utility in global vegetation studies.

440

By focusing on the harmonization of satellite-derived SIF products, the LHSIF dataset offers a unified framework for integrating multi-sensor SIF data to enable long-term monitoring of global photosynthesis. This contribution provides an essential tool for understanding vegetation responses to environmental changes and advancing the field of Earth system science.

445 **6 Data availability statement**

The LHSIF dataset generated in this study is publicly available at <https://doi.org/10.5281/zenodo.14854185> (Zou et al., 2025). Additional information regarding the data and methods is available upon request from the corresponding author.

7 Author contribution

CZ and LL designed the experiments, CZ carried them out. CZ, SD and XL developed the model code and generated the products. CZ prepared the manuscript with contributions from all co-workers.

450

8 Competing interests

The contact author has declared that none of the authors has any competing interests.

9 Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant number: 42425001).

455



References

- Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C.: TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958 – 2015, *Scientific Data*, 5, 170191, <https://doi.org/10.1038/sdata.2017.191>, 2018.
- 460 Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., and Zhao, M.: Spatiotemporal patterns of terrestrial gross primary production: A review, *Reviews of Geophysics*, 53, 785-818, <https://doi.org/10.1002/2015RG000483>, 2015.
- Angal, A., Xiong, X., Choi, T., Chander, G., Mishra, N., and Helder, D. L.: Impact of Terra MODIS Collection 6 on long-term trending comparisons with Landsat 7 ETM+ reflective solar bands, *Remote sensing letters*, 4, 873-881, 465 <https://doi.org/10.1080/2150704X.2013.809496>, 2013.
- Bacour, C., Maignan, F., Peylin, P., MacBean, N., Basting, V., Joiner, J., Köhler, P., Guanter, L., and Frankenberg, C.: Differences Between OCO-2 and GOME-2 SIF Products From a Model-Data Fusion Perspective, *Journal of Geophysical Research: Biogeosciences*, 124, 3143-3157, <https://doi.org/10.1029/2018JG004938>, 2019.
- Badgley, G., Field, C. B., and Berry, J. A.: Canopy near-infrared reflectance and terrestrial photosynthesis, *Science Advances*, 3, e1602244, <https://doi.org/10.1126/sciadv.1602244>, 2017.
- 470 Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luysaert, S., Margolis, H., Oleson, K. W., Rouspard, O., . . . Papale, D.: Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate, *Science*, 329, 834-838, <https://doi.org/10.1126/science.1184984>, 2010.
- 475 Berry, J. A., Frankenberg, C., Wennberg, P., Baker, I., Bowman, K. W., Castro-Contreas, S., Cendrero-Mateo, M. P., Damm, A., Drewry, D., and Ehlmann, B.: New methods for measurement of photosynthesis from space, *Geophys. Res. Lett.*, 38, L17706, <https://doi.org/10.26206/9NJP-CG56>, 2012.
- Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C.: A Limited Memory Algorithm for Bound Constrained Optimization, *SIAM Journal on Scientific Computing*, 16, 1190-1208, <https://doi.org/10.1137/0916069>, 1995.
- 480 Chen, R., Liu, L., Liu, X., and Rascher, U.: CMLR: A Mechanistic Global GPP Dataset Derived from TROPOMIS SIF Observations, *Journal of Remote Sensing*, 4, 0127, <https://doi.org/10.34133/remotesensing.0127>, 2024.
- Chen, S., Liu, L., Sui, L., Liu, X., and Ma, Y.: An improved spatially downscaled solar-induced chlorophyll fluorescence dataset from the TROPOMI product, *Scientific Data*, 12, 135, <https://doi.org/10.1038/s41597-024-04325-6>, 2025.
- Chen, X., Huang, Y., Nie, C., Zhang, S., Wang, G., Chen, S., and Chen, Z.: A long-term reconstructed TROPOMI solar-induced 485 fluorescence dataset using machine learning algorithms, *Scientific Data*, 9, 427, <https://doi.org/10.1038/s41597-022-01520-1>, 2022.
- Cui, T., Sun, R., Qiao, C., Zhang, Q., Yu, T., Liu, G., and Liu, Z.: Estimating Diurnal Courses of Gross Primary Production for Maize: A Comparison of Sun-Induced Chlorophyll Fluorescence, Light-Use Efficiency and Process-Based Models, *Remote Sensing*, 9, 1267, <https://doi.org/10.3390/rs9121267>, 2017.
- 490 Damm, A., Guanter, L., Paul-Limoges, E., van der Tol, C., Hueni, A., Buchmann, N., Eugster, W., Ammann, C., and Schaepman, M. E.: Far-red sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An assessment based on observational and modeling approaches, *Remote Sensing of Environment*, 166, 91-105, <https://doi.org/10.1016/j.rse.2015.06.004>, 2015.
- Duveiller, G. and Cescatti, A.: Spatially downscaling sun-induced chlorophyll fluorescence leads to an improved temporal 495 correlation with gross primary productivity, *Remote Sensing of Environment*, 182, 72-89, <https://doi.org/10.1016/j.rse.2016.04.027>, 2016.
- Duveiller, G., Filippini, F., Walther, S., Köhler, P., Frankenberg, C., Guanter, L., and Cescatti, A.: A spatially downscaled sun-induced fluorescence global product for enhanced monitoring of vegetation productivity, *Earth Syst. Sci. Data*, 12, 1101-1116, <https://doi.org/10.5194/essd-12-1101-2020>, 2020.
- 500 Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., Toon, G. C., Butz, A., Jung, M., Kuze, A., and Yokota, T.: New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross



- primary productivity, *Geophysical Research Letters*, 38, <https://doi.org/10.1029/2011GL048738>, 2011.
- Gentine, P. and Alemohammad, S. H.: Reconstructed Solar-Induced Fluorescence: A Machine Learning Vegetation Product Based on MODIS Surface Reflectance to Reproduce GOME-2 Solar-Induced Fluorescence, *Geophysical Research Letters*, 45, 3136-3146, <https://doi.org/10.1002/2017GL076294>, 2018.
- 505 Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gómez-Dans, J., Kuze, A., Suto, H., and Grainger, R. G.: Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements, *Remote Sensing of Environment*, 121, 236-251, <https://dx.doi.org/10.1016/j.rse.2012.02.006>, 2012.
- Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., Frankenberg, C., Huete, A. R., Zarco-Tejada, P., Lee, J. E., Moran, M. S., Ponce-Campos, G., Beer, C., Camps-Valls, G., Buchmann, N., Gianelle, D., Klumpp, K., Cescatti, A., Baker, J. M., and Griffis, T. J.: Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence, *Proc Natl Acad Sci U S A*, 111, E1327-1333, <https://doi.org/10.1073/pnas.1320008111>, 2014.
- 510 Hahne, A., Lefebvre, A., and Callies, J.: Global ozone monitoring experiment on board ERS-2, *Environmental Sensing '92*, SPIE, <https://doi.org/10.1117/12.140233>, 1993.
- 515 Hong, Z., Hu, Y., Cui, C., Yang, X., Tao, C., Luo, W., Zhang, W., Li, L., and Meng, L.: An Operational Downscaling Method of Solar-Induced Chlorophyll Fluorescence (SIF) for Regional Drought Monitoring, *Agriculture*, 12, 547, <https://doi.org/10.3390/agriculture12040547>, 2022.
- Jeong, S., Ryu, Y., Gentine, P., Lian, X., Fang, J., Li, X., Dechant, B., Kong, J., Choi, W., Jiang, C., Keenan, T. F., Harrison, S. P., and Prentice, I. C.: Persistent global greening over the last four decades using novel long-term vegetation index data with enhanced temporal consistency, *Remote Sensing of Environment*, 311, 114282, <https://doi.org/10.1016/j.rse.2024.114282>, 2024.
- 520 Joiner, J., Yoshida, Y., Koehler, P., Frankenberg, C., and Parazoo, N. C.: L2 Daily Solar-Induced Fluorescence (SIF) from ERS-2 GOME, 1995-2003, <https://doi.org/10.3334/ORNLDAAC/1758>, 2019.
- Joiner, J., Yoshida, Y., Koehler, P., Frankenberg, C., and Parazoo, N. C.: L2 Solar-Induced Fluorescence (SIF) from SCIAMACHY, 2003-2012, <https://doi.org/10.3334/ORNLDAAC/1871>, 2021.
- 525 Joiner, J., Yoshida, Y., Vasilkov, A. P., Yoshida, Y., and Corp, L. A.: First observations of global and seasonal terrestrial chlorophyll fluorescence from space, *Biogeosciences*, 8, 3(2011-03-08), 8, <https://doi.org/10.5194/bg-8-637-2011>, 2011.
- Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., Huemmrich, K. F., Yoshida, Y., and Frankenberg, C.: Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2, *Atmospheric Measurement Techniques*, 6, 2803-2823, <https://doi.org/10.5194/amt-6-2803-2013>, 2013.
- 530 Khakurel, P., Leigh, L., Kaewmanee, M., and Pinto, C. T.: Extended pseudo invariant calibration site-based trend-to-trend cross-calibration of optical satellite sensors, *Remote Sensing*, 13, 1545, <https://doi.org/10.3390/rs13081545>, 2021.
- Köhler, P., Guanter, L., and Joiner, J.: A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data, *Atmospheric Measurement Techniques*, 8, 2589-2608, <https://doi.org/10.5194/amt-8-2589-2015>, 2015.
- 535 Koren, G., van Schaik, E., Araújo, A. C., Boersma, K. F., Gärtner, A., Killaars, L., Kooreman, M. L., Kruijt, B., van der Laan-Luijkx, I. T., von Randow, C., Smith, N. E., and Peters, W.: Widespread reduction in sun-induced fluorescence from the Amazon during the 2015/2016 El Niño, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373, 20170408, <https://doi.org/10.1098/rstb.2017.0408>, 2018.
- 540 Li, X. and Xiao, J.: A Global, 0.05-Degree Product of Solar-Induced Chlorophyll Fluorescence Derived from OCO-2, MODIS, and Reanalysis Data, *Remote Sensing*, 11, 517, <https://doi.org/10.3390/rs11050517>, 2019.
- Li, X., Xiao, J., and He, B.: Chlorophyll fluorescence observed by OCO-2 is strongly related to gross primary productivity estimated from flux towers in temperate forests, *Remote Sensing of Environment*, 204, 659-671, <https://doi.org/10.1016/j.rse.2017.09.034>, 2018.
- 545 Lu, X., Cai, G., Zhang, X., Yu, H., Zhang, Q., Wang, X., Zhou, Y., and Su, Y.: Research on downscaling method of the enhanced TROPOMI solar-induced chlorophyll fluorescence data, *Geocarto International*, 39, 2354417, <https://doi.org/10.1080/10106049.2024.2354417>, 2024.



- Ma, Y., Liu, L., Liu, X., and Chen, J.: An improved downscaled sun-induced chlorophyll fluorescence (DSIF) product of GOME-2 dataset, *European Journal of Remote Sensing*, 55, 168-180, <https://doi.org/10.1080/22797254.2022.2028579>, 2022.
- 550 Markham, B. L. and Helder, D. L.: Forty-year calibrated record of earth-reflected radiance from Landsat: A review, *Remote Sensing of Environment*, 122, 30-40, <https://doi.org/10.1016/j.rse.2011.06.026>, 2012.
- Mishra, N., Haque, M. O., Leigh, L., Aaron, D., Helder, D., and Markham, B.: Radiometric Cross Calibration of Landsat 8 Operational Land Imager (OLI) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+), *Remote Sensing*, 6, 12619-12638, <https://doi.org/doi:10.3390/rs61212619>, 2014.
- 555 Mohammed, G. H., Colombo, R., Middleton, E. M., Rascher, U., van der Tol, C., Nedbal, L., Goulas, Y., Pérez-Priego, O., Damm, A., Meroni, M., Joiner, J., Cogliati, S., Verhoef, W., Malenovsky, Z., Gastellu-Etchegorry, J.-P., Miller, J. R., Guanter, L., Moreno, J., Moya, I., . . . Zarco-Tejada, P. J.: Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress, *Remote Sensing of Environment*, 231, 111-177, <https://doi.org/10.1016/j.rse.2019.04.030>, 2019.
- 560 OCO-2/OCO-3 Science Team, V. P., Abhishek Chatterjee: OCO-2 Level 2 bias-corrected solar-induced fluorescence and other select fields from the IMAP-DOAS algorithm aggregated as daily files [dataset], <https://doi.org/10.5067/OTRE7KQS8AU8>, 2020.
- Parazoo, N. C., Frankenberg, C., Köhler, P., Joiner, J., Yoshida, Y., Magney, T., Sun, Y., and Yadav, V.: Towards a Harmonized Long-term Spaceborne Record of Far-red Solar-induced Fluorescence, *Journal of Geophysical Research: Biogeosciences*, 124, <https://doi.org/10.1029/2019JG005289> 2019.
- 565 Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Ribeca, A., van Ingen, C., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö, J., . . . Papale, D.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, *Scientific Data*, 7, 225, <https://doi.org/10.1038/s41597-020-0534-3>, 2020.
- 570 Porcar-Castell, A., Tyystjärvi, E., Atherton, J., Tol, C., Flexas, J., Pfündel, E., Moreno, J., Frankenberg, C., and Berry, J.: Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges %J *Journal of experimental botany*, 65, <https://doi.org/10.1093/jxb/eru191>, 2014.
- Rascher, U., Alonso, L., Burkart, A., Cilia, C., Cogliati, S., Colombo, R., Damm, A., Drusch, M., Guanter, L., Hanus, J., Hyvärinen, T., Julitta, T., Jussila, J., Kataja, K., Kokkalis, P., Kraft, S., Kraska, T., Matveeva, M., Moreno, J., . . . Zemek, F.: Sun-induced fluorescence – a new probe of photosynthesis: First maps from the imaging spectrometer HyPlant, *Global Change Biology*, 21, 4673-4684, <https://doi.org/10.1111/gcb.13017>, 2015.
- 575 Rossini, M., Meroni, M., Migliavacca, M., Manca, G., Cogliati, S., Busetto, L., Picchi, V., Cescatti, A., Seufert, G., and Colombo, R.: High resolution field spectroscopy measurements for estimating gross ecosystem production in a rice field, *Agricultural and Forest Meteorology*, 150, 1283-1296, <https://doi.org/10.1016/j.agrformet.2010.05.011>, 2010.
- 580 Ryu, Y., Berry, J. A., and Baldocchi, D. D.: What is global photosynthesis? History, uncertainties and opportunities, *Remote Sensing of Environment*, 223, 95-114, <https://doi.org/10.1016/j.rse.2019.01.016>, 2019.
- Ryu, Y., Jiang, C., Kobayashi, H., and Detto, M.: MODIS-derived global land products of shortwave radiation and diffuse and total photosynthetically active radiation at 5 km resolution from 2000, *Remote Sensing of Environment*, 204, 812-825, <https://doi.org/10.1016/j.rse.2017.09.021>, 2018.
- 585 Schaaf, C. and Wang, Z.: MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF-Adjusted Ref Daily L3 Global 0.05Deg CMG V061. [dataset], <https://doi.org/10.5067/MODIS/MCD43C4.061>, 2021.
- Smith, W. K., Biederman, J. A., Scott, R. L., Moore, D. J. P., He, M., Kimball, J. S., Yan, D., Hudson, A., Barnes, M. L., MacBean, N., Fox, A. M., and Litvak, M. E.: Chlorophyll Fluorescence Better Captures Seasonal and Interannual Gross Primary Productivity Dynamics Across Dryland Ecosystems of Southwestern North America, *Geophysical Research Letters*, 45, 748-757, <https://doi.org/10.1002/2017GL075922>, 2018.
- 590 Sun, Y., Frankenberg, C., Wood, J. D., Schimel, D. S., Jung, M., Guanter, L., Drewry, D. T., Verma, M., Porcar-Castell, A., Griffis, T. J., Gu, L., Magney, T. S., Köhler, P., Evans, B., and Yuen, K.: OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence, *Science*, 358, eaam5747, <https://doi.org/10.1126/science.aam5747>, 2017.
- Tao, S., Chen, J. M., Zhang, Z., Zhang, Y., Ju, W., Zhu, T., Wu, L., Wu, Y., and Kang, X.: A high-resolution satellite-based solar-induced chlorophyll fluorescence dataset for China from 2000 to 2022, *Scientific Data*, 11, 1286,



- <https://doi.org/10.1038/s41597-024-04101-6>, 2024.
- Tavora, J., Jiang, B., Kiffney, T., Bourdin, G., Gray, P. C., de Carvalho, L. S., Hesketh, G., Schild, K. M., Faria de Sousa, L., Brady, D. C., and Boss, E.: Recipes for the Derivation of Water Quality Parameters Using the High-Spatial-Resolution Data from Sensors on Board Sentinel-2A, Sentinel-2B, Landsat-5, Landsat-7, Landsat-8, and Landsat-9 Satellites, *Journal of Remote Sensing*, 3, 0049, <https://doi.org/10.34133/remotesensing.0049>, 2023.
- 600 van Schaik, E., Kooreman, M. L., Stammes, P., Tilstra, L. G., Tuinder, O. N. E., Sanders, A. F. J., Verstraeten, W. W., Lang, R., Cacciari, A., Joiner, J., Peters, W., and Boersma, K. F.: Improved SIFTER v2 algorithm for long-term GOME-2A satellite retrievals of fluorescence with a correction for instrument degradation, *Atmospheric Measurement Techniques*, 13, 4295-4315, <https://doi.org/10.5194/amt-13-4295-2020>, 2020.
- 605 Verma, M., Schimel, D., Evans, B., Frankenberg, C., Beringer, J., Drewry, D. T., Magney, T., Marang, I., Hutley, L., Moore, C., and Eldering, A.: Effect of environmental conditions on the relationship between solar-induced fluorescence and gross primary productivity at an OzFlux grassland site, *Journal of Geophysical Research: Biogeosciences*, 122, 716-733, <https://doi.org/10.1002/2016JG003580>, 2017.
- Wang, S., Zhang, Y., Ju, W., Wu, M., Liu, L., He, W., and Peñuelas, J.: Temporally corrected long-term satellite solar-induced fluorescence leads to improved estimation of global trends in vegetation photosynthesis during 1995–2018, *ISPRS Journal of Photogrammetry and Remote Sensing*, 194, 222-234, <https://doi.org/10.1016/j.isprsjprs.2022.10.018>, 2022.
- 610 Weimin Ju and Zhou., Y.: Global Daily GPP Simulation Dataset from 1981 to 2019 [dataset], <https://doi.org/10.12199/nesdc.ecodb.2016YFA0600200.02.001>, 2021.
- Wen, J., Köhler, P., Duveiller, G., Parazoo, N. C., Magney, T. S., Hooker, G., Yu, L., Chang, C. Y., and Sun, Y.: A framework for harmonizing multiple satellite instruments to generate a long-term global high spatial-resolution solar-induced chlorophyll fluorescence (SIF), *Remote Sensing of Environment*, 239, 111644, <https://doi.org/10.1016/j.rse.2020.111644>, 2020.
- 615 Yang, J., Tian, H., Pan, S., Chen, G., Zhang, B., and Dangal, S.: Amazon drought and forest response: Largely reduced forest photosynthesis but slightly increased canopy greenness during the extreme drought of 2015/2016, *Global Change Biology*, 24, 1919-1934, <https://doi.org/10.1111/gcb.14056>, 2018.
- 620 Yang, X., Tang, J., Mustard, J. F., Lee, J.-E., Rossini, M., Joiner, J., Munger, J. W., Kornfeld, A., and Richardson, A. D.: Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest, *Geophysical Research Letters*, 42, 2977-2987, <https://doi.org/10.1002/2015GL063201>, 2015.
- Zhang, Y. and Peñuelas, J.: Combining Solar-Induced Chlorophyll Fluorescence and Optical Vegetation Indices to Better Understand Plant Phenological Responses to Global Change, *Journal of Remote Sensing*, 3, 0085, <https://doi.org/10.34133/remotesensing.0085>, 2023.
- 625 Zhang, Y., Joiner, J., Gentine, P., and Zhou, S.: Reduced solar-induced chlorophyll fluorescence from GOME-2 during Amazon drought caused by dataset artifacts, *Global Change Biology*, 24, 2229-2230, <https://doi.org/10.1111/gcb.14134>, 2018a.
- Zhang, Y., Joiner, J., Alemohammad, S. H., Zhou, S., and Gentine, P.: A global spatially contiguous solar-induced fluorescence (CSIF) dataset using neural networks, *Biogeosciences*, 15, 5779-5800, <https://doi.org/10.5194/bg-15-5779-2018>, 2018b.
- 630 Zhang, Y., Guanter, L., Berry, J. A., van der Tol, C., Yang, X., Tang, J., and Zhang, F.: Model-based analysis of the relationship between sun-induced chlorophyll fluorescence and gross primary production for remote sensing applications, *Remote Sensing of Environment*, 187, 145-155, <https://doi.org/10.1016/j.rse.2016.10.016>, 2016.
- Zhang, Y., Zhang, Q., Liu, L., Zhang, Y., Wang, S., Ju, W., Zhou, G., Zhou, L., Tang, J., Zhu, X., Wang, F., Huang, Y., Zhang, Z., Qiu, B., Zhang, X., Wang, S., Huang, C., Tang, X., and Zhang, J.: ChinaSpec: A Network for Long-Term Ground-Based Measurements of Solar-Induced Fluorescence in China, *Journal of Geophysical Research: Biogeosciences*, 126, e2020JG006042, <https://doi.org/10.1029/2020JG006042>, 2021.
- 635 Zhang, Z., Zhang, Y., Joiner, J., and Migliavacca, M.: Angle matters: Bidirectional effects impact the slope of relationship between gross primary productivity and sun-induced chlorophyll fluorescence from Orbiting Carbon Observatory-2 across biomes, *Global Change Biology*, 24, 5017-5020, <https://doi.org/10.1111/gcb.14427>, 2018c.
- 640 Zheng, X., Zhao, W., Zhu, Z., Wang, Z., Zheng, Y., and Li, D.: Characterization and Evaluation of Global Solar-Induced Chlorophyll Fluorescence Products: Estimation of Gross Primary Productivity and Phenology, *Journal of Remote Sensing*, 4, 0173, <https://doi.org/10.34133/remotesensing.0173>, 2024.



- Zhu, W., Xie, Z., Zhao, C., Zheng, Z., Qiao, K., Peng, D., and Fu, Y. H.: Remote sensing of terrestrial gross primary productivity: a review of advances in theoretical foundation, key parameters and methods, *GIScience & Remote Sensing*, 61, 2318846, <https://doi.org/10.1080/15481603.2024.2318846>, 2024.
- 645
- Zou, C., Du, S., Liu, X., and Liu, L.: TCSIF: a temporally consistent global Global Ozone Monitoring Experiment-2A (GOME-2A) solar-induced chlorophyll fluorescence dataset with the correction of sensor degradation, *Earth Syst. Sci. Data*, 16, 2789-2809, <https://doi.org/10.5194/essd-16-2789-2024>, 2024.
- Zou, C., Liu, L., Du, S., and Liu, X.: LHSIF: the Long-Term Harmonized Multi-Satellite SIF dataset with a resolution of 0.05° spanning 1995 to 2023 [dataset], <https://doi.org/10.5281/zenodo.14854185>, 2025.
- 650
- Zuromski, L. M., Bowling, D. R., Köhler, P., Frankenberg, C., Goulden, M. L., Blanken, P. D., and Lin, J. C.: Solar-Induced Fluorescence Detects Interannual Variation in Gross Primary Production of Coniferous Forests in the Western United States, *Geophysical Research Letters*, 45, 7184-7193, <https://doi.org/10.1029/2018GL077906>, 2018.