A Spatially Downscaled TROPOMI SIF Product at 0.005° Resolution With Bias Correction

Jiaochan Hu, Jia Jia, Zihan Ma, Keyu Yuan, Haoyang Yu, *Member, IEEE*, and Liangyun Liu

***Abstract*—Solar-induced chlorophyll fluorescence (SIF) provides a valuable tool for gross primary production (GPP) monitoring. However, the spatial resolution of satellite SIF products is lower than the kilometer level, which hinders their potential for carbon cycle study at regional scales. This work reconstructed a 0.005° SIF in China during 2019 and 2020 from the 5-km level TROPOMI SIF by a proposed downscaling strategy that corrected the predicted bias when statistics-based machine learning models, such as random forest (RF), were used. Our bias-corrected downscaled SIF (named BCSIF) had an improved capacity of preserving the information of the original TROPOMI SIF than the directly predicted SIF from RF. The BCSIF showed better consistencies with the tower-based SIF than the 0.05° TROPOMI SIF with an averaged R^2 increased from 0.590 to 0.798 at two sites since it has a more comparable spatial scale with spectral observations. For the spatial–temporal correlations with FLUXCOM GPP at different biomes in China, BCSIF outperformed the original SIF with the averaged R^2 increased from 0.472 to 0.877 due to its reduced noise, also outperformed the near-infrared radiation reflected by vegetation (NIRvP), especially for the savanna type with the R^2 increased from 0.828 to 0.889. For the temporal correlations with FLUXCOM GPP, BCSIF gives comparable R^2 values as NIRvP in more than half of China (around 65% pixels), not including the needleleaf forest region in the southern Tibetan Plateau and savanna region in Yunnan province where BCSIF greatly outperformed, as well as some alpine meadows regions in Inner Mongolia and Tibetan Plateau where NIRvP outperformed.**

***Index Terms*—Bias correction, gross primary production (GPP), random forest (RF), solar-induced chlorophyll fluorescence (SIF), spatial downscaling.**

Manuscript received 11 April 2024; revised 10 July 2024; accepted 20 July 2024. Date of publication 25 July 2024; date of current version 5 August 2024. This work was supported in part by the Open Research Fund of Key Laboratory of Digital Earth Science, Chinese Academy of Sciences, under Grant 2022LDE008, in part by the Open Fund of State Key Laboratory of Remote Sensing Science under Grant OFSLRSS202319, and in part by the National Natural Science Foundation of China under Grant 42371353 and Grant 42001280. *(Corresponding author: Jia Jia.)*

Jiaochan Hu is with the Green Shipping and Carbon Neutrality Lab, College of Environmental Sciences and Engineering, Dalian Maritime University, Dalian 116026, China, and also with the State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China (e-mail: hujc@dlmu.edu.cn).

Jia Jia, Zihan Ma, and Haoyang Yu are with the Information Science and Technology College, Dalian Maritime University, Dalian 116026, China (e-mail: jia2015@yeah.net; mzhoyo@dlmu.edu.cn; yuhy@dlmu.edu.cn).

Keyu Yuan is with the College of Environmental Sciences and Engineering, Dalian Maritime University, Dalian 116026, China (e-mail: yuankeyu@dlmu.edu.cn).

Liangyun Liu is with the State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and also with the International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China. (e-mail: liuly@radi.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3433371

I. Introduction

SOLAR-INDUCED chlorophyll fluorescence (SIF) is a long-wave signal emitted by plant chlorophyll molecules during the photosynthesis process under natural sunlight [1]. Compared with vegetation indices (VIs), SIF has unique physiological information that can better track the dynamics in gross primary production (GPP) under environmental stress [2], [3], [4]. Over the past decades, numerous satellite sensors have the potential for retrieving SIF datasets, including SCIAMACHY [5], GOSAT/-2 [6], GOME-2 [7], OCO-2/3 [8], TROPOMI [9], [10], Tansat [11], TECIS-1 [12], etc. With the successful SIF retrieval from satellite sensors, SIF provides an innovative insight for monitoring spatial and temporal patterns of plant-based terrestrial carbon fixation (i.e., GPP) using remote sensing data.

The satellite SIF datasets have the problems of low spatial resolutions or discontinuous spatial patterns. Thereinto, TROPOMI on Sentinel-5P has relatively good spatial continuities and resolutions, which provides SIF retrievals at a spatial continuous sampling of 3.5–14 km across track and 7.5 km along track (5.5 km since 2018) with daily global coverage. Nevertheless, since TROPOMI SIF still has the spatial resolution lower than the kilometer level, it cannot accurately capture the real variations in photosynthetic capacity at finer scales, particularly in heterogeneous landscapes. The spatial mismatch between satellite TROPOMI SIF and tower-based GPP will limit the understanding of the relationship between SIF and GPP through robust ground validation, which impedes the full capability of it for monitoring region-scale GPP. To improve the finer scale application of TROPOMI SIF in complex heterogeneous landscapes or regional scales, efforts need to be made to obtain SIF at a higher spatial resolution.

To address the problems of low spatial resolution, several studies develop gap-filling or downscaling methods to reconstruct satellite SIF. These methods are mostly based on the concept of light use efficiency (LUE) and can be divided into two ways: semiempirical function [13], [14] and machine-learning (ML) methods [15], [16], [17], [18], [19], [20]. The results of the former depend on the robustness of the semiempirical model, while ML [e.g., random forest (RF), artificial neural network (ANN), or eXtreme gradient boosting (XGBoost)] has a superior ability to learn the nonlinear and complex statistical relationships due to its better flexibility in fitting. Therefore, RF has the outstanding performance in handling regression tasks and reducing the risk of overfitting; thus, it has been widely used in downscaling SIF [15], [16].

To derive sufficient samples for model training, most ML-based methods constructed a universal model at relatively large spatial scales, such as the globe [17], [18], [21], [22], different continents [19], or different land cover types [23], rather than adopting multiple local models. However, the ML models, such as RF, ANN, and XGBoost, are statistics-based, and thus, they tend to learn the relationship (between SIF and driving variables) that occurs more frequently in most samples. Besides, at large spatial scales, a universal model can represent the relationship for most samples when appropriate driving variables are used, but not performing well for a few samples, because this relationship may change in some cases (e.g., certain areas or plant types and environmental stresses). Therefore, the strategy relying solely on ML will inevitably introduce some biases in SIF prediction.

Using spatial downscaling methods, several downscaled SIF products were generated. Most of the downscaled SIF products had a kilometer-level resolution, e.g., 0.05° downscaled SIF from GOME-2 SIF. A few of them had more fined resolution but were limited to regional scales, e.g., 0.005° downscaled SIF from TROPOMI in North America [19] or reconstructed OCO-2 SIF in Henan province, China [20]. Generating a downscaled SIF product with 100-m resolution on a national or continental scale is crucial. Additionally, whether the data-driven reconstructed SIF carried the physiological information of original SIF signals or just canopy optical and structural characters, and how its performance for tracking GPP at large scales needed to be further illustrated. Thus, it is important to generate 0.005° SIF product in a large region and provide proofs for its capability of GPP monitoring.

In this study, we generated a 16-day 0.005° downscaled SIF product [named bias-corrected downscaled SIF (BCSIF)] in China during 2019 and 2020 from the TROPOMI SIF, MODIS, and fifth ECMWF reanalysis (ERA5) data using an improved downscaling strategy with the prediction bias corrected and further evaluated the potential of downscaled SIF for monitoring GLASS GPP by assessing the correlation of SIF-GPP compared with the near-infrared radiation reflected by vegetation (NIRvP). The main contributions of this article were as follows:

1) develop an improved spatial downscaling method correcting the prediction bias when using the RF model;
2) obtain a 0.005° downscaled SIF product over China;
3) evaluate the advantages of the downscaled SIF for monitoring GPP over the original SIF and the NIRvP.

II. Materials

A. TROPOMI SIF Datasets

We derived the original TROPOMI SIF from the ungridded 743–758 nm L2B SIF datasets at a daily scale from 2019 to 2020 published by Guanter et al. [9].[1] The day-length corrected SIF based on the cosine of the solar zenith angle [cos(SZA)] [24] provided in the L2B datasets was adopted in this study. We aggregated the ungridded SIF to 0.05°, 16-day resolutions with data screening for the downscaling model development and validation of the downscaled SIF. Each SIF grid cell was the average of those footprints, which covered the center of this grid. In order to reduce the uncertainties in individual SIF retrieved soundings due to clouds and noises, data screening was implemented, including two aspects: First, based on the thresholds settings in previous studies [10], [25] and the suggestions by Guanter et al. [9], SIF soundings with the cloud fraction less than 0.2 and 0.8 were aggregated to obtain TROPOMI SIF values under clear-sky and all-sky conditions (hereafter named TROSIF and $\text{TROSIF}_{\text{as}}$, respectively); second, only if the SIF grids contained more than three SIF soundings were averaged, given that the uncertainty in SIF retrieval can be reduced by a factor $\sqrt{n}$ (n is the number of soundings in each grid) [8].

[1][Online]. Available: https://s5p-troposif.noveltis.fr/data-access/

B. Explanatory Variables Datasets

The explanatory variables related to vegetation characters were acquired from the MCD43A4 v006 reflectance product [26]. We collected the BRDF-corrected seven-band reflectance at 500 m, daily resolutions. To ensure the quality of reflectance data, only the pixel values with quality assurance 0 and 1 (i.e., full BRDF inversion and magnitude BRDF inversion, respectively) were retained. The VIs map, i.e., the normalized difference water index (NDWI), was also calculated. The 500 m daily reflectance was aggregated to 0.05°, 16-day and 0.005°, 16-day scales for model development and SIF prediction, respectively.

The explanatory variables related to environmental conditions, including vapor pressure deficit (VPD), air temperature (*Ta*), soil moisture (SM), photosynthetically active radiation (PAR) under all-sky conditions, and PAR under clear-sky conditions, were derived or calculated from the ERA5 datasets at 0.1° hourly resolutions. Thereinto, VPD was calculated utilizing *Ta* and the dewpoint temperature (*DTa*) according to the formula introduced by Allen et al. [27]. The 0.1° hourly meteorological variables were resampled to 0.05°, 16-day and 0.005°, 16-day scales using the nearest neighbor interpolation method for model development and SIF prediction, respectively.

C. Satellite Auxiliary Datasets

To evaluate the potential of downscaled SIF product for GPP monitoring, we collected two GPP products (i.e., GLASS GPP and FLUXCOM GPP) for spatial comparison and correlation analysis, respectively. The FLUXCOM GPP product (RS_V006) (abbreviated as FLUXGPP hereafter) was based on a data-driven approach using multiple ML methods to upscale carbon flux measurements from eddy covariance towers by remote sensing and meteorological data [28], [29]. This dataset was at an eight-day temporal resolution and a 0.0833° spatial resolution, and was resampled to 0.05°, 16-day scale for comparison in this study. The GLASS GPP product (abbreviated as GLASGPP hereafter) has a higher spatial resolution at 500 m. This dataset was retrieved using a collection of eight LUE models based on radiation and environmental components [30]. We aggregated the original data to 0.005°, 16-day resolution.

To evaluate the performance of BCSIF for different biomes, the land cover product with 500 m resolution was obtained from MCD12Q1 v006. The primary land cover scheme based

TABLE I
DETAILS OF THE FLUX TOWER SITES

ID	Location	Period	Height
DM	38.8555°N 100.3722°E	July to September in 2019 May to October in 2020	25 m
AR	38.0473°N 100.4643°E	May and August to October in 2019 June to August in 2020	25 m

on the International Geosphere–Biosphere Program consists of 17 land cover classes [31]. Based on the strategy in [24], we divided the 17 landcover types into 7 classes, including evergreen and deciduous needleleaf forest (NF), evergreen broadleaf forest (EBF), deciduous broadleaf forests (DBFs) and mixed forests, closed and open shrublands (SHR), woody savannas and savannas (SAV), croplands and cropland/natural vegetation mosaic (CRO), and grasslands (GRA). The unvegetated types, such as "permanent wetland," "water," "urban and built-up," "barren or sparsely vegetated," and "snow and ice," were excluded in this study.

To investigate the potential of BCSIF for GPP monitoring, we compared it with the performance of NIRvP [32]. NIRvP denotes the radiation reflected by the vegetation at the near-infrared band. Several studies have reported that NIRvP has a capability for GPP estimation on certain ecosystems, including grassland, crops, wetland, and savanna [33], [34]. The good relationship between the satellite-based NIRvP and GPP has also been studied [35], [36]. In this study, we calculated the NIRvP using the collected MODIS reflectance and PAR data, as shown in the following equation:

$$\mathrm{NIRvP} = \mathrm{PAR} \times \mathrm{NIRv} = \mathrm{PAR} \times \mathrm{NDVI} \times R_{\mathrm{NIR}} \quad (1)$$

where NIRvP is the combination of PAR and near-infrared reflectance of vegetation (NIRv). NIRv is calculated by the normalized difference vegetation index multiplied by the near-infrared band reflectance (R_{NIR}).

D. Tower-Based Datasets

We used tower-based datasets at two sites to validate the downscaled SIF. The details of these sites were shown in Table I. The sites were components of the ChinaSpec [37], including DaMan (DM) site predominated by single-cropping maize in Gansu province at northwest China and Arou (AR) site predominated by the alpine meadow in Qinghai province at northwest China. More details about the site conditions, measurement system, and data process can be found in [37].

The SIF at 760 nm was retrieved based on the three-band Fraunhofer line depth [38] and the singular value decomposition method [39]. To compare with the 740 nm satellite SIF, we converted the 760 nm tower-based SIF into 740 nm by multiplying a wavelength scaling coefficient of 1.5 [40]. To be consistent with the TROPOMI SIF, the tower-based SIF was averaged into 16-day scale if there were more than 8-day valid values in a 16-day period.

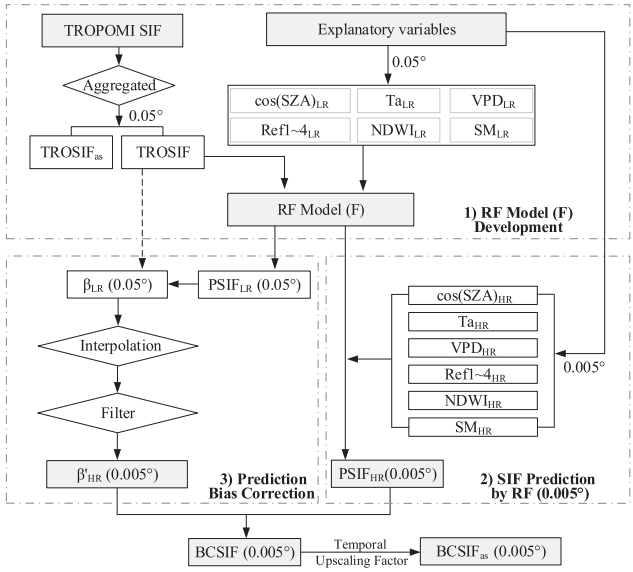


Fig. 1. Flowchart of the process of SIF downscaling.

III. METHODOLOGY

A. Selection of Explanatory Variables

Similar to the concept of LUE model for GPP, SIF can be defined as follows [41]:

$$\mathrm{SIF} = \mathrm{PAR} \times \mathrm{fPAR} \times \varepsilon \times \Phi_{\mathrm{SIF}} \quad (2)$$

where fPAR is the fraction of PAR absorbed by vegetation, ε is the fraction of SIF photons escaping from the photosystem level to the canopy level, and Φ_{SIF} is the fluorescence quantum yield. According to the linear correlation between PAR and cos(SZA) under clear-sky conditions [24], [42], we utilized cos(SZA) to express the information in clear-sky incoming PAR. The effects of ε and fPAR on SIF variations were mainly dominated by the leaf optical, canopy structural, and geometric information, which could be denoted by the MODIS reflectance [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23]. The effects of Φ_{SIF} on SIF changes were mostly governed by the environmental conditions related to sunlight, temperature, and water [43]. Based on the accuracy evaluation of different explanatory variables in our previous study [44] and related data-driven SIF downscaling pieces of literature, we selected nine explanatory variables to drive the downscaling model in total, including the first four bands of MODIS reflectance, cos(SZA), *Ta*, VPD, SM, and NDWI.

B. Improved Downscaling Approach

The improved downscaling procedure was divided into three parts: RF model development, SIF prediction by RF, and prediction bias correction (see Fig. 1). The subscript "LR" or "HR" denoted as the data was at a low/high resolution as 0.05° or 0.005°, respectively. The SIF data (i.e., BCSIF and TROSIF) with the subscript "as" or not denoted it were under all-sky or clear-sky conditions, respectively.

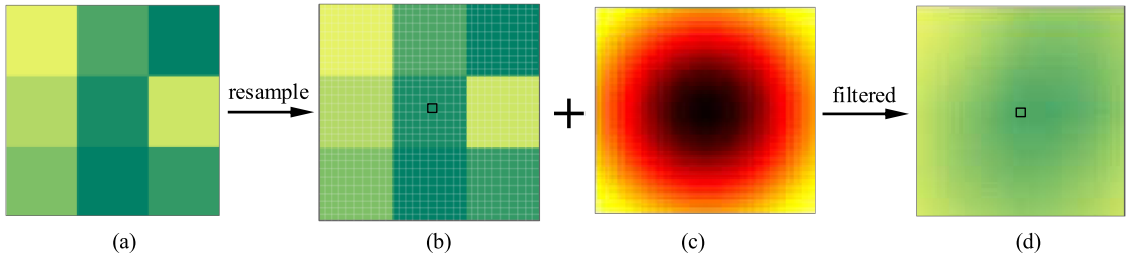


Fig. 2. Schematic representation of the calculation of the prediction bias. (a) β_{HR} (0.05°). (b) β_{HR} (0.005°). (c) Weights (0.005°). (d) β'_{HR} (0.005°).

1) RF Model Development: For the model development, we used a strategy of threefold cross validation, i.e., each of the three individual experiments selected 30% of the whole samples for testing and the remaining 70% for training. All the samples were normalized by the average and standard deviation of training samples. We established the RF model (F) for each 16-day period over China using the training samples derived from the low-resolution original TROPOMI SIF under clear-sky conditions and the corresponding explanatory variables, as shown in the following equation:

$$\mathrm{TROSIF} = F\left(\cos(\mathrm{SZA})_{\mathrm{LR}}, Ta_{\mathrm{LR}}, \mathrm{VPD}_{\mathrm{LR}}, \ldots\right). \quad (3)$$

For the processing of RF model, the setting of the number of trees (*n*trees) and minimum leaf size (*m*leaf, i.e., the minimum number of samples required in each training subset for splitting the decision tree at each node) was determined according to multiple experiments with testing samples. With the value of *n*trees increasing, both the model accuracy and computing time increased, reaching saturation at 100 trees. Additionally, the model accuracy markedly decreased with the value of *m*leaf larger than 5. The setting of *m*leaf as 5 produced comparable accuracy to that of 2 while taking less computing time. Therefore, in the RF model, we set the values of *n*trees and *m*leaf as 100 and 5, respectively.

2) SIF Prediction by RF: For SIF prediction, the established model F was applied to obtain the high-resolution SIF predictions by RF (named $\mathrm{PSIF}_{\mathrm{HR}}$) under clear-sky conditions using explanatory variables at 0.005° resolution, as (4). The explanatory variables were also normalized by the same mean value and standard deviation as training samples over each 16-day period

$$\mathrm{PSIF}_{\mathrm{HR}} = F\left(\cos(\mathrm{SZA})_{\mathrm{HR}}, Ta_{\mathrm{HR}}, \mathrm{VPD}_{\mathrm{HR}}, \ldots\right). \quad (4)$$

However, the SIF predicted solely by the RF model would inevitably contain biases in a few samples due to the reasons mentioned before that one statistics-based universal model cannot well represent the relationship (between SIF and driven variables) for some cases in China. Therefore, in this study, the SIF predicted by RF was postprocessed using a bias correction method.

3) Prediction Bias Correction: The idea of bias correction in this article is to constrain the ML-based SIF predictions by the value of local original SIF. We first calculated the low-resolution SIF predictions by the developed RF model (named $\mathrm{PSIF}_{\mathrm{LR}}$) using explanatory variables at 0.05° resolutions, as shown in the following equation:

$$\mathrm{PSIF}_{\mathrm{LR}} = F\left(\cos(\mathrm{SZA})_{\mathrm{LR}}, Ta_{\mathrm{LR}}, \mathrm{VPD}_{\mathrm{LR}}, \ldots\right). \quad (5)$$

Then, we calculated the difference between the TROSIF and $\mathrm{PSIF}_{\mathrm{LR}}$ as the 0.05° low-resolution prediction bias (named β_{LR}), as shown in the following equation:

$$\beta_{\mathrm{LR}} = \mathrm{TROSIF} - \mathrm{PSIF}_{\mathrm{LR}}. \quad (6)$$

To be consistent with the spatial resolution of $\mathrm{PSIF}_{\mathrm{HR}}$, the bias β_{LR} was processed as follows, which was illustrated by the schematic representation in Fig. 2. First, we resampled β_{LR} to the 0.005° high resolution using the nearest neighbor interpolation method (namely β_{HR}). But large differences between adjacent pixels in the original 0.05° TROSIF would be propagated to the low-resolution β_{LR}, and eventually resulted in some artificial abrupt transitions in the high-resolution β_{HR} that seemed as unnatural "large square" patterns [see Fig. 2(b)]. To minimize the effects of this issue, we filtered the β_{HR} image using a running window with a 2-D Gaussian function (namely weights) to calculate the filtered high-resolution bias (namely β'_{HR}), as shown in (7). Specifically, each β'_{HR} pixel [the central pixel in Fig. 2(d)] is calculated as a spatially weighted average of the adjacent β_{HR} pixels with the Gaussian running window in Fig. 2(c).

The standard deviation (σ) of the Gaussian function was set as 5, according to the optimal validation accuracy of BCSIF predictions from multiple experiments. With the value of σ increasing, the consistency between reaggregated BCSIF and TROSIF improved, reaching saturation at σ as 5 (see Fig. 10). Accordingly, the size of running window (w) was set as 29×29 because the values of Gaussian function out of the range $[-3\sigma, 3\sigma]$ were close to zero. Through this process, the artificial abrupt transitions in β_{HR} could be smoothed, which is more reasonable in accordance with the actual distribution of high-resolution bias map

$$\beta'_{\mathrm{HR}} = \sum_{i}^{n=29\times 29} \mathrm{Weights} \times \beta_{\mathrm{HR}}. \quad (7)$$

Thus, the downscaled BCSIF product was generated by adding the resampled and filtered bias obtained at a coarse resolution to the $\mathrm{PSIF}_{\mathrm{HR}}$ directly predicted by the RF model, as follows:

$$\mathrm{BCSIF} = \mathrm{PSIF}_{\mathrm{HR}} + \beta'_{\mathrm{HR}}. \quad (8)$$

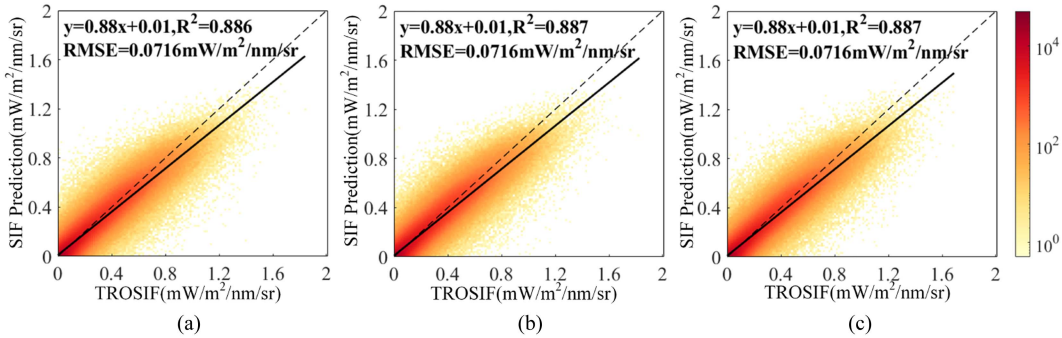


Fig. 3. Scatter diagrams between the 0.05° TROSIF and predicted SIF by RF model for the testing samples of three cross-validation experiments in 2019. (a) First experiment. (b) Second experiment. (c) Third experiment.

C. All-Sky Downscaled SIF Generation

Based on the whole improved downscaling approach, the 0.005°, 16-day BCSIF at clear-sky scale in China over the periods 2019 and 2020 was reconstructed by using nine explanatory variables at 0.005°, 16-day resolution. The spatial gaps of the original TROSIF or the high-resolution explanatory variables for predicting $PSIF_{HR}$ would result in no BCSIF in some pixels. If there is no TROSIF, $PSIF_{HR}$ will be preserved as BCSIF. If the explanatory variables were absent, the value of BCSIF would be the same as TROSIF.

Finally, to convert BCSIF from clear-sky conditions to all-sky conditions (hereafter named $BCSIF_{as}$), we multiplied BCSIF to a temporal upscaling factor (i.e., the ratio of PAR under all-sky conditions to PAR under clear-sky conditions from ERA5) [45], as (9). Thus, the final 0.005° $BCSIF_{as}$ product at all-sky scale in China during 2019 and 2020 was obtained

$$BCSIF_{as} = \frac{BCSIF}{PAR_{cs}} \times PAR_{as} \tag{9}$$

where PAR_{as} and PAR_{cs} are PAR under all-sky and clear-sky conditions, which were derived from the ERA5 data.

D. Validation Approaches

The BCSIF was validated in three ways. First, to validate the downscaled SIF for preserving information of the original SIF and to evaluate the effects of bias correction, we compared the consistency of BCSIF and $PSIF_{HR}$ with the original TROPOMI SIF values under clear-sky conditions (i.e., TROSIF). On the one hand, we analyzed the correlations between TROSIF with the reaggregated BCSIF and $PSIF_{HR}$ at 0.05° scales. On the other hand, we compared the spatial patterns of 0.05° TROSIF with the 0.005° BCSIF and 0.005° $PSIF_{HR}$.

Second, we assessed the reliability of BCSIF using the continuous tower-based SIF retrievals and analyzed its advantages over the low-resolution original SIF by comparing the time series and linear correlations between tower-based SIF with the 0.05° $TROSIF_{as}$, the 0.05° reaggregated $BCSIF_{as}$, and the 0.005° $BCSIF_{as}$ at two sites under all-sky conditions.

Third, we explored the potential of BCSIF for GPP monitoring and its performance compared with the original SIF and the NIRvP. Considering that the data-driven GPP (e.g., FLUXCOM GPP) can avoid the uncertainties in the estimation of LUE, we used the FLUXCOM GPP as a reference for analyzing the correlations of SIF and NIRvP with GPP. Since GLASS has a higher spatial resolution, it was used to assess the spatial details of the 0.005° downscaled SIF. Specifically, we compared the spatial patterns of the annual maps of 0.05° $TROSIF_{as}$ and 0.005° $BCSIF_{as}$ with 0.005° GLAGPP under all-sky conditions in 2019. Then, we further evaluated both the correlations with FLUXGPP at different biomes and the spatial distributions of temporal correlations at pixel level among $TROSIF_{as}$, $BCSIF_{as}$, and NIRvP at 0.05° scales.

IV. Results

A. Performance of the RF Model

The performance of the RF model for 0.05° SIF predicting was shown in Fig. 3. It displayed the scatterplots between SIF directly predicted by the RF and the original TROSIF for testing samples of three cross-validation experiments in 2019.

In general, the RF model produced a relatively satisfactory accuracy: the averaged coefficient of determination (R^2) of three experiments was around 0.887, and the averaged root-mean-square error (RMSE) was 0.0716 mW/m^2/nm/sr. The predicted SIF was consistent with TROSIF: the scatters were distributed closely to the 1:1 line (the dashed lines). Note that the SIF predictions from RF had a slight underestimated bias at the high-value level probably because some changes in SIF cannot be fully expressed by the model of driving variables.

B. Validation of Downscaled SIF With TROPOMI SIF

To evaluate the capability of downscaled SIF for preserving key information in the original SIF, we reaggregated the high-resolution BCSIF and $PSIF_{HR}$ to 0.05°, namely $\overline{BCSIF}$ and $\overline{PSIF}_{HR}$, and quantitatively assessed the consistency between them with TROSIF, as shown in Fig. 4. Through the bias correction, BCSIF had better correlations with TROSIF than PSIF: the scatters were closer to the 1:1 line with the R^2 improved from 0.853 to 0.882 and the RMSE decreased from 0.0898 to 0.0810 mW/m^2/nm/sr. Particularly, with bias correction, some scatters for the PSIF, which were obviously off

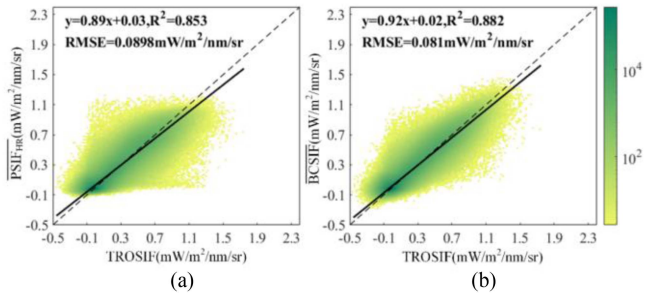


Fig. 4. Scatter diagrams between TROSIF and two reaggregated downscaled SIF. (a) $\overline{\text{PSIF}_{\text{HR}}}$. (b) $\overline{\text{BCSIF}}$ at the 0.05° scale in 2019.

the linear trend (see light yellow dots in the figure), were greatly more concentrated.

We further compared the capability of PSIF_{HR} and BCSIF for preserving spatial variations in the original SIF, as shown in Fig. 5. Overall, the spatial patterns of two downscaled SIF (PSIF_{HR} and BCSIF) in China were similar to that of TROSIF. Specifically, in August 2019, the relatively high SIF values were exhibited in the cropland and forests' areas (e.g., the north China plain, the Xiaoxing'an Mountains, and Changbaishan Mountain in the northeast of China), whereas the relatively low values were distributed in grass and barren or sparsely vegetated areas (e.g., Tibet, Qinghai Province, Xinjiang Autonomous Region, and part areas of Neimeng Autonomous Region in China). Compared with PSIF_{HR}, BCSIF was more consistent with the TROSIF in some regions. Particularly, in parts of the cropland or DBF types of the northeastern China (gray square and blue ellipse regions in Fig. 5) and parts of the grassland in Qinghai Province (the blue circle region in Fig. 5), the PSIF_{HR} underestimated the TROSIF with significant positive values of bias (referring to the bias maps in Fig. 11), while BCSIF corrected this bias. Moreover, the 0.005° downscaled SIF exhibited finer spatial details in regional scales, such as the identification of Nenjiang river basin in the enlarged maps, whereas the 0.05° TROSIF showed obvious patchy structures.

C. Validation of Downscaled SIF With Tower-Based SIF

The tower-based SIF retrievals were also used to validate the reliability of downscaled SIF products. We investigated the consistency between the tower-based SIF with the all-sky original TROPOMI SIF at 0.05° resolution (i.e., $\text{TROSIF}_{\text{as}}$), the reaggregated all-sky BCSIF at 0.05° resolution (named as $\overline{\text{BCSIF}_{\text{as}}}$), and the all-sky BCSIF at 0.005° resolution (i.e., BCSIF_{as}) at two sites (see Fig. 6).

All the correlations between 0.005° BCSIF_{as} and tower-based SIF at two sites were high with the R^2 larger than 0.68 and all the RMSE values were less than 0.15 mW/m^2/nm/sr, particularly, the R^2 at DM site was 0.911, which verified the reliability of the 0.005° downscaled BCSIF_{as}. The 0.005° BCSIF_{as} was obviously more consistent with tower-based SIF than the two 0.05° low-resolution products, with the averaged R^2 improved from 0.590/0.689 ($\text{TROSIF}_{\text{as}}$/$\overline{\text{BCSIF}_{\text{as}}}$) to 0.798, and the corresponding averaged RMSE decreased from 0.201/0.234 to 0.141 mW/m^2/nm/sr for two sites. It indicated that the spatial downscaling could improve the ability of satellite SIF for tracking tower-based SIF due to its more comparable spatial footprint with spectral observations. In addition, although at the same 0.05° scale, the $\overline{\text{BCSIF}_{\text{as}}}$ tracked the tower-based SIF better than the $\text{TROSIF}_{\text{as}}$ probably because the noises in the original SIF had been reduced through the reconstruction by the data-driven model.

D. Comparison of Downscaled SIF With GPP Products

To investigate the potential of BCSIF for monitoring GPP, we compared the spatial patterns of the annual maps of 0.05° $\text{TROSIF}_{\text{as}}$ and the downscaled 0.005° BCSIF_{as} with high-resolution GLAGPP in 2019, as shown in Fig. 7. In general, $\text{TROSIF}_{\text{as}}$ and BCSIF_{as} had a similar spatial pattern over China, which was also consistent with the distribution of GLAGPP. However, for the regional scale, compared with BCSIF, the original low-resolution TROSIF could not accurately capture the spatial variations in 0.005° GLAGPP. Specifically, in the enlarged maps in Fig. 7, the low-resolution TROSIF exhibited an obvious patch effect and lost many spatial details of GPP near the boundaries among CRO, DBF, and GRA, especially in the white ellipse region where the highly productive GRA structures had disappeared in the TROSIF, whereas the 0.005° BCSIF could resolve the spatial problems in TROSIF and captured the spatial distributions of different biomes.

Furthermore, to compare the potential of all-sky TROSIF, BCSIF, and NIRvP for monitoring GPP, we analyzed the correlations between them with FLUXGPP both at different biomes and at the pixel level, as shown in Figs. 8 and 9, respectively. The R^2 of both linear and nonlinear (a hyperbolic function based on [46]) regressions was calculated, abbreviated as $R^2{}_L$ and $R^2{}_{\text{NL}}$. Only the higher R^2 for each correlation was displayed in the figure. Due to the low percentage of pixels for SHR (0.09%), NF (0.98%), and EBF (3.74%) in China, we ignored the SHR and merged NF and EBF types into DBF as one forest type in the analysis.

In Fig. 8, all the correlations between BCSIF and FLUXGPP for four biomes were high with linear or nonlinear R^2, respectively, larger than 0.842, which implied that BCSIF had potential for monitoring GPP. Note that for forest, SAV, and CRO types, the R^2 for nonlinear regression was higher than that for linear regression, but for GRA types, the R^2 for linear regression was higher. In addition, when using BCSIF instead of TROSIF, the R^2 values were greatly improved with the averaged R^2 of four biomes, which increased from 0.472 to 0.877, probably because the downscaled SIF retained the key information of the original SIF while greatly reducing the noises. Moreover, compared with NIRvP, the scatters of BCSIF were obviously more concentrated. BCSIF produced higher correlations for four biomes: the averaged R^2 increased from 0.848 to 0.877, especially for the SAV type.

We further evaluated the spatial distributions of pixel-level temporal correlations between three products with FLUXGPP at 0.05° scales in 2019, as shown in Fig. 9. In general, the R^2 of

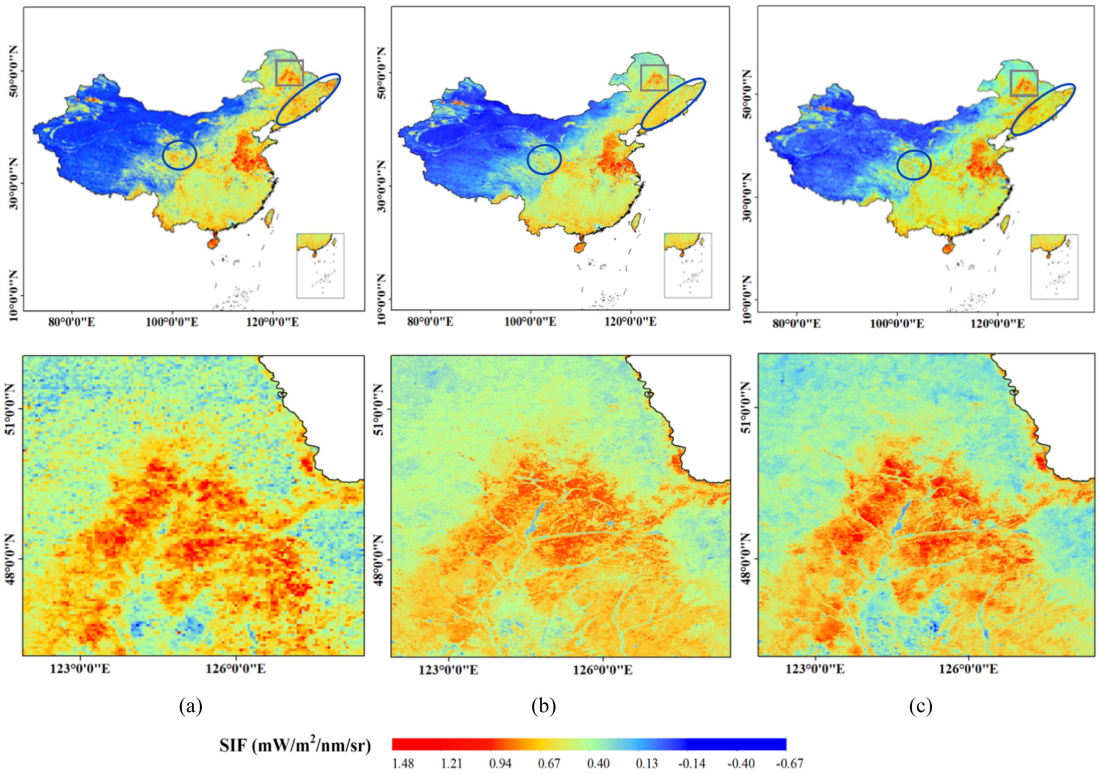


Fig. 5. Spatial patterns of (a) TROSIF, (b) $\mathrm{PSIF}_{\mathrm{HR}}$, and (c) BCSIF in August 2019. Enlarged maps show three SIF products in the gray square region.

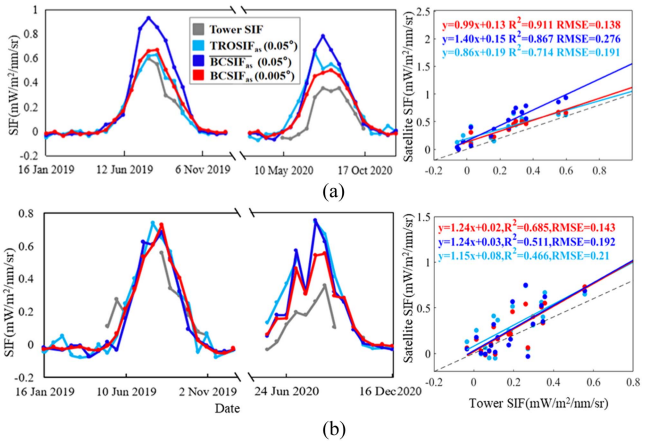


Fig. 6. Comparison between the time series of the tower-based SIF with the 0.005° BCSIF, the 0.05° reaggregated BCSIF, and the 0.05° TROSIF under all-sky conditions for (a) DM and (b) AR sites. All regressions in the right panel achieved the significance level of 0.05.

BCSIF distributed similarly as that of the TROSIF [see Fig. 9(a) and (b)], but BCSIF showed higher R^2 values than TROSIF in most parts of China (around 83% pixels) [see Fig. 9(d)] due to reduced noises. The R^2 of BCSIF was higher than 0.4 for most pixels but relatively low in some alpine meadows of Inner Mongolia and Tibetan Plateau (blue ellipses in the figure). While in these regions, NIRvP performed better [see (Fig. 9(c)], resulting in negative values in the difference map of Fig. 9(e). While for NF type in southern Tibetan Plateau and SAV type of Yunnan province, the R^2 of NIRvP was low and BCSIF greatly outperformed NIRvP (gray rectangles in the figure), causing positive values in the difference map of Fig. 9(e). Except for these typical regions that were marked, the R^2 of BCSIF was comparable with that of NIRvP in more than half of China (around 65% of pixels exhibited the differences of R^2 lower than $\pm$ 0.05), such as central, eastern, and northeast China [see Fig. 9(e)]. A similar phenomenon for these typical regions was found for the differences of R^2 between TROSIF and NIRvP [see blue ellipses and gray rectangles in Fig. 9(f)].

V. Discussion

A. Necessity of Bias Correction for ML-Based Downscaling Approaches

In this study, we have established a downscaling strategy to correct the bias in SIF predictions based on RF. The reason for this predicted bias is that statistics-based ML models, such as RF, tend to learn the relationship (between SIF and driven variables) that occurs more frequently in most samples, and the relationship may change in some cases that cannot be well represented by one universal model in China. Therefore, once a universal model based on statistical ML is used for SIF downscaling at large spatial scales, the bias will inevitably exist. This bias has also been mentioned in other downscaling studies [16], [47].

To verify whether there is bias using other ML models, we took XGBoost and ANN models as examples. Similar to Figs. 3 and 4, we have also validated the accuracies of model testing

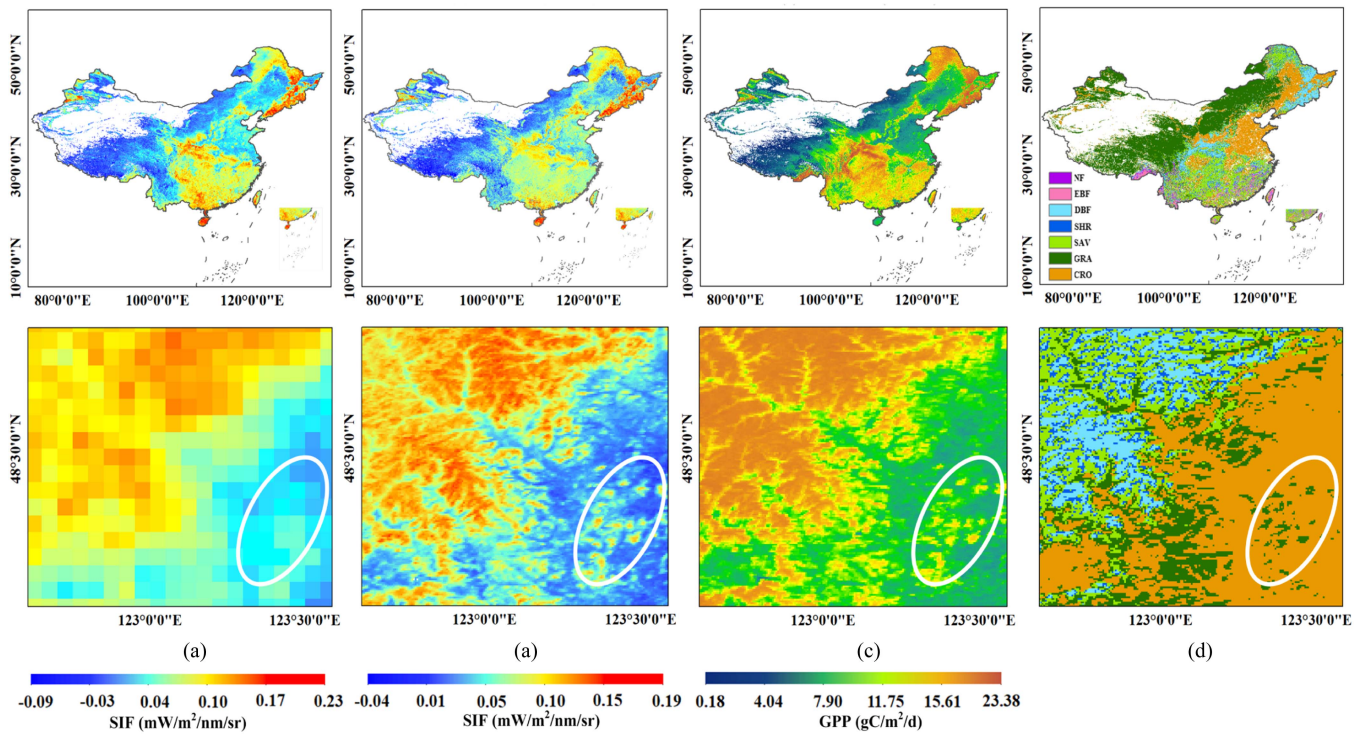


Fig. 7. Spatial patterns of annual mean of (a) all-sky TROSIF, (b) all-sky BCSIF, (c) GLAGPP, and (d) landcover in 2019. Enlarged maps show data in northeastern China.

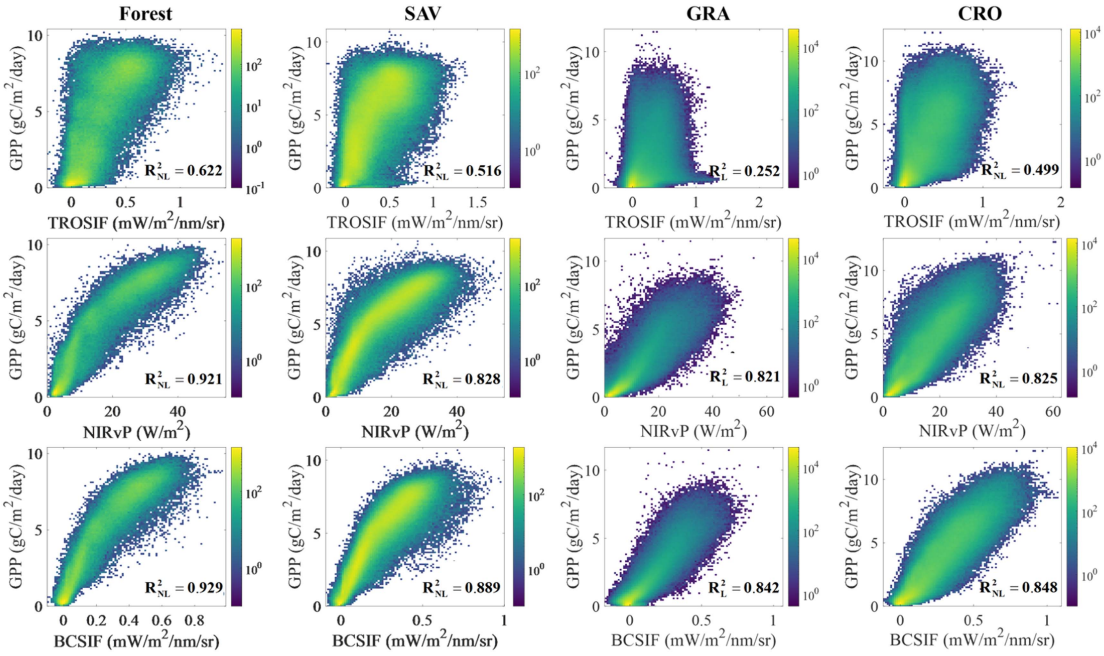


Fig. 8. Correlations between the 16-day all-sky TROSIF, BCSIF, and NIRvP with the FLUXGPP at 0.05° scales for different biomes in 2019. $R^2{}_L$ and $R^2{}_{\mathrm{NL}}$ values denote the averaged coefficients of determinations of the linear and hyperbolic nonlinear regressions.

and investigated the performance of SIF predicting before and after the bias correction. We have found that both XGBoost and ANN models produce comparable accuracies with the RF model for testing samples (see Fig. 12). In addition, similar biases, as shown in Fig. 4(a), exist in the SIF predictions by XGBoost and ANN, which caused that some scatters (light yellow dots in the figure) were obviously off the linear trend (see Fig. 13). Through the bias correction for XGBoost and ANN, the scatters are more concentrated with higher R^2 and lower RMSE values. This result illustrates the necessity of bias correction when

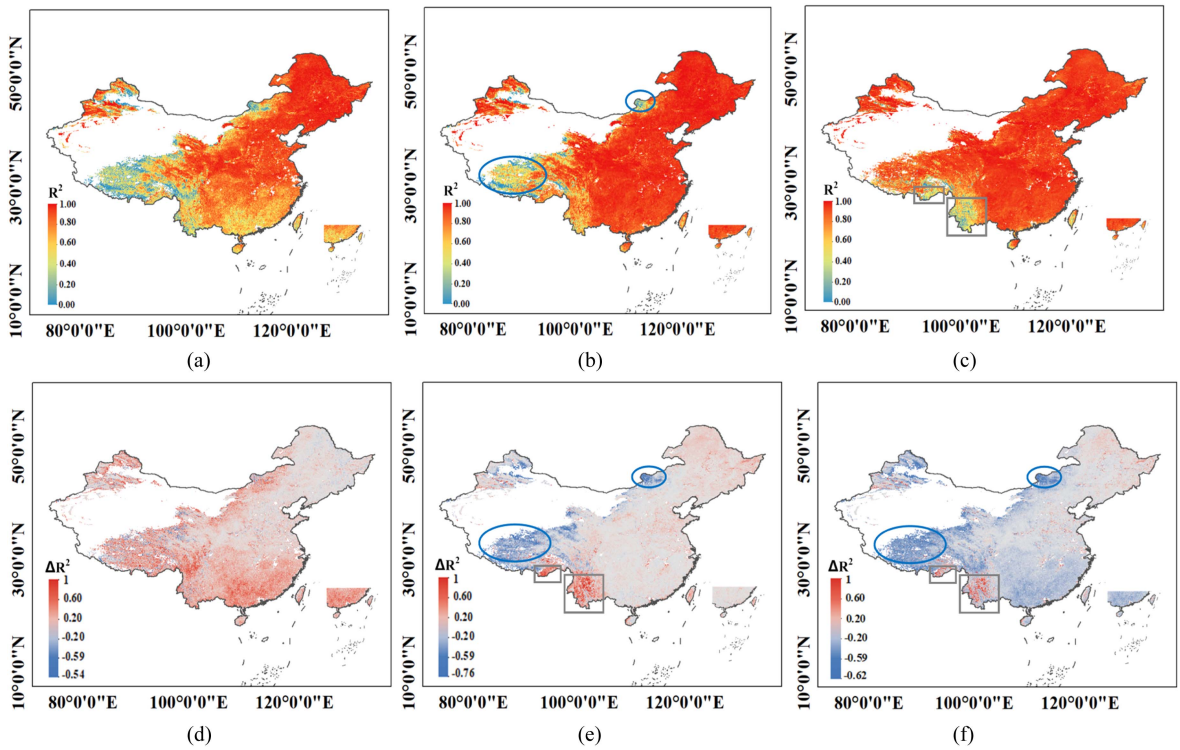


Fig. 9. Spatial distribution of the temporal correlations among the 16-day all-sky TROSIF, BCSIF, and NIRvP with the FLUXGPP, and the differences of corresponding correlations between them at 0.05° scales in 2019. All regressions achieved the significance level of 0.05. (a) R^2 ($TROSIF_{as}$ vs. FLUXGPP). (b) R^2 ($BCSIF_{as}$ vs. FLUXGPP). (c) R^2 (NIRvP vs. FLUXGPP). (d) R^2 ($BCSIF_{as}$ vs. FLUXGPP)-R^2 ($TROSIF_{as}$ vs. FLUXGPP). (e) R^2 ($BCSIF_{as}$ vs. FLUXGPP)-R^2 (NIRvP vs. FLUXGPP). (f) R^2 ($TROSIF_{as}$ vs. FLUXGPP)-R^2 (NIRvP vs. FLUXGPP).

using other statistics-based ML models and further strengthens the effectiveness of our downscaling strategy with bias correction.

B. *Prospects of Downscaled SIF for GPP Monitoring*

The emitted SIF signal contains photosynthetic physiological information about the plant, while NIRvP represents the vegetation reflected radiation at the near-infrared band, which is mainly governed by leaf biochemical optical parameters, canopy structure, and incident and observed geometries. Several studies [48], [49] demonstrated that SIF can better track the variation of photosynthesis than NIRvP, especially under conditions of drought or temperature stresses. According to Fig. 9, compared with TROSIF, BCSIF has a similar spatial pattern in temporal correlations with FLUXGPP, and the typical regions where SIF obviously outperformed or underperformed NIRvP for R^2 also distribute similarly. Thus, it can be inferred that our downscaled SIF still preserves the physiological information of original TROPOMI SIF, which can play an important role in monitoring vegetation GPP.

In China, both the BCSIF and NIRvP have high temporal consistencies with FLUXGPP [see Fig. 9(b) and (c)], which indicates that both two products have good potential for GPP monitoring. For the spatial–temporal correlations of different biomes, BCSIF is superior to NIRvP, and the advantage is more noticeable for other types than the forest type, which agrees with the results in [2]. But for temporal correlations, BCSIF shows relatively low R^2 values in some alpine meadows of Inner Mongolia and Tibetan Plateau, which may be due to the low satellite SIF values and little seasonal dynamic fluctuation caused by sparse grassland coverage in these regions. On the contrary, in the savanna region of Yunnan province, BCSIF has obviously higher R^2 values than NIRvP, which is probably caused by the physiological information in downscaled SIF. Specifically, the reason may be that the less precipitation in winter-to-spring transition of Yunnan due to the monsoon climate results in a suppression of photosynthesis. This dynamic pattern of the plant's physiological response to stress can be tracked by SIF but has little influence on the reflectance-based NIRvP. Similarly, the different dynamic responses between SIF/GPP and VI or APAR to the stress have also been reported for evergreen forests in Colorado, USA, by the research [3].

Moreover, compared with the original TROPOMI SIF, BCSIF has significant advantages in spatial details and reducing noises, which provides a better product for the monitoring of vegetation photosynthesis or physiological state in China. The finer SIF products with more spatial details at 0.005° resolution can better capture the spatial variations in GPP. There are also several studies emphasizing the importance of fine-resolution SIF in crop yield prediction [50], [51]. On the other hand, BCSIF largely eliminates the noises in TROSIF, which can be concluded from Figs. 6, 8, and 9. From this perspective, BCSIF has the same benefits as the enhanced SIF products that combine

NIRvP and satellite SIF physiological information generated by Liu et al. [2].

C. Limitations

First, although the data sources have passed the quality control before modeling, the uncertainties in them will introduce errors in the downscaled SIF. On the one hand, the original TROPOMI SIF retrievals still have errors from the sensors, retrieval models, and cloud contamination, which will be involved in the downscaling modeling and validation process, even though we have reduced the errors by averaging three soundings and by refining cloudy fractions. Besides, the uncertainties come from the errors inherent in explanatory variables, such as MODIS reflectance and ERA5 meteorological parameters, as well as the interpolation process of those variables.

Second, the uncertainties involved in the resampling of the bias map will also cause errors in the downscaled SIF. This study used the nearest neighbor interpolation and Gaussian filter of the coarse-resolution bias to calculate the high-resolution bias. We assumed that the 0.05° coarse bias is evenly distributed to each 0.005° grid within it and ignored the different weights in the bias of different 0.005° grids. This strategy is acceptable because most of the pixel's differences inside the 0.05° coarse-resolution grid are already represented by the SIF predicted by ML, and the remaining bias is not governed by the driving variables (e.g., vegetation reflectance, PAR, or Ta) used in SIF. Besides, within the scope of existing knowledge, it is difficult to give the law and influencing factors of the bias in the present results. For instance, we did not find an apparent relationship between the spatial pattern of bias in Fig. 11 with that of SIF in Fig. 5 and the landcover map in Fig. 7. Although this strategy will introduce some errors and does not produce a completely accurate bias map, it is effective in constraining the 0.005° ML-based SIF predictions to be closer to the original SIF.

Third, more robust and accurate deep learning networks or coupling model-driven and data-driven algorithms need to be presented for the reconstruction of SIF data in future work. In addition, how different ways of using methods work should be quantitatively evaluated further.

Finally, the validation methods of downscaled SIF should be more adequate and diversified in the future. Since the "truth" of the downscaled SIF at the space scale is absent, this kind of downscaling work is needed. A wider range of SIF observation by aerial or unmanned aerial vehicle platforms is more ideal for the authenticity verification of SIF predictions.

VI. Conclusion

In this study, we presented an improved SIF downscaling strategy, which corrected the prediction bias when using RF model, and then generated a 0.005° downscaled SIF product in China (named BCSIF). Compared with the SIF directly predicted by RF, BCSIF was more consistent with the original TROSIF. The validation showed that our 0.005° BCSIF had a better consistency with tower-based SIF than the 0.05° TROPOMI SIF with an averaged R^2 increased from 0.590 to 0.798 at two sites due to the reduced noises and the more comparable spatial scale between them. In China, the BCSIF had high consistencies with FLUXGPP both for biome-specific spatial–temporal correlations and pixel-level temporal correlations, which verified its good potential for GPP monitoring. For the spatial–temporal correlations with FLUXGPP in different biomes, BCSIF outperformed TROSIF for linear/nonlinear R^2 due to its lower noises, and also outperformed NIRvP, especially for the SAV type. For temporal correlation, the performance between BCSIF and NIRvP was comparable in more than half of China. The regions where SIF obviously outperformed or underperformed NIRvP in temporal correlations distributed similarly between BCSIF and TROSIF (i.e., the savanna region in Yunnan province and some alpine meadows in Inner Mongolia or Tibetan Plateau, respectively). It implied that BCSIF preserved the physiological information in the original SIF that might not contained in the NIRvP.

Our downscaled BCSIF has advantages in spatial details, reducing noises, and preserving physiological information of the original TROPOMI SIF, which provided data support for fine-scale ecosystem monitoring.

Data Availability: BCSIF products at 0.005°, 16-day resolutions in China during 2019 and 2020 are now available for free.[2]

Appendix

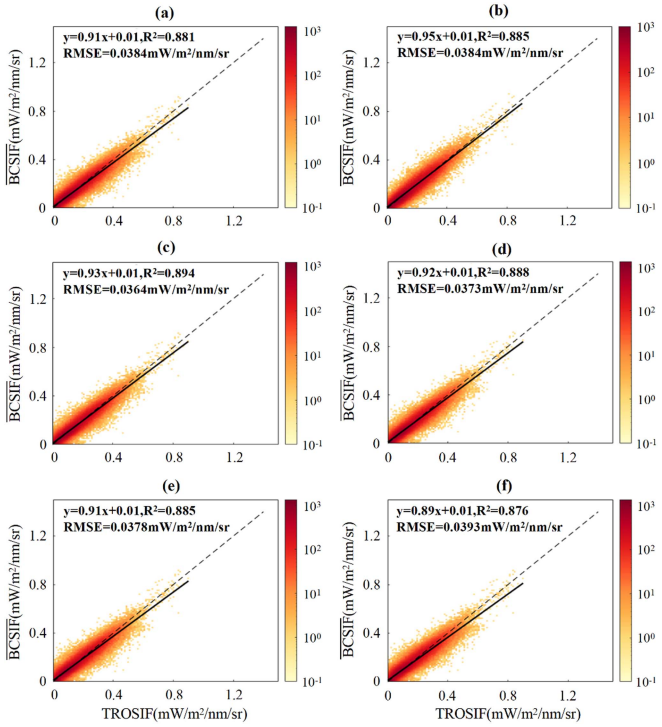


Fig. 10. Scatter diagrams between TROSIF and the reaggregated downscaled $\overline{\text{BCSIF}}$ at the 0.05° scale in January 2019 among different settings of standard deviation of the Gaussian function. (a) $\sigma = 1$. (b) $\sigma = 3$. (c) $\sigma = 5$. (d) $\sigma = 7$. (e) $\sigma = 9$. (f) $\sigma = 11$.

[2][Online]. Available: https://doi.org/10.5281/zenodo.10371968

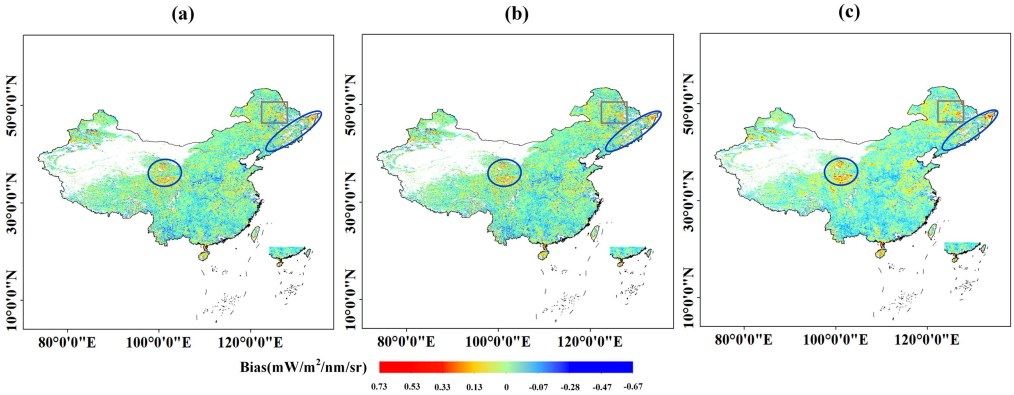


Fig. 11. Spatial patterns of the (a) 0.05° low-resolution prediction bias β_{LR}, (b) 0.005° high-resolution prediction bias β_{HR}, and (c) filtered 0.005° high-resolution bias β'_{HR} in August 2019.

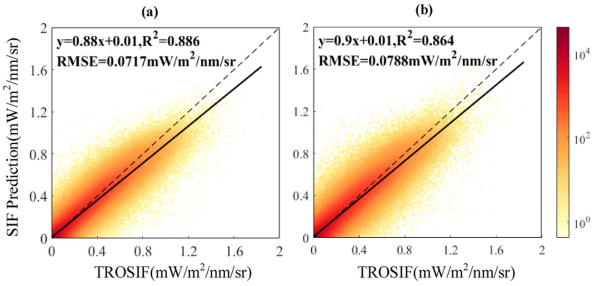


Fig. 12. Scatter diagrams between the 0.05° TROSIF and predicted SIF by (a) XGBoost model and (b) ANN model for the testing samples in 2019.

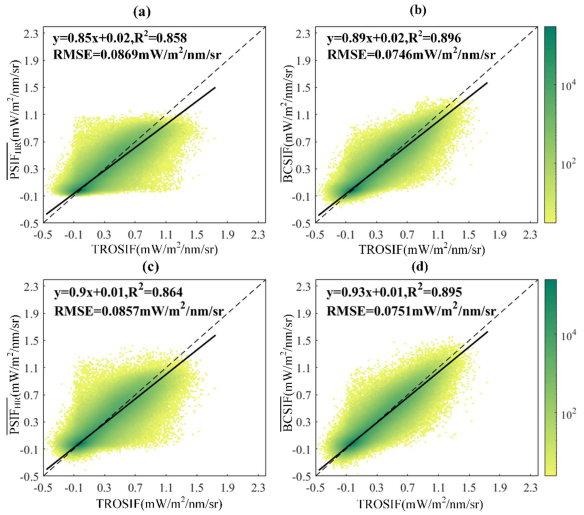


Fig. 13. Scatter diagrams between TROSIF with PSIF and BCSIF based on XGBoost and ANN models at the 0.05° scale in 2019. (a) $\overline{\mathrm{PSIF_{HR}}}$ based on XGBoost model. (b) $\overline{\mathrm{BCSIF}}$ based on XGBoost model. (c) $\overline{\mathrm{PSIF_{HR}}}$ based on ANN model. (d) $\overline{\mathrm{BCSIF}}$ based on ANN model.

ACKNOWLEDGMENT

The authors would like to thank Prof. L. Liu and the related group from the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, for generously providing the long-term data from flux tower sites, and Dr. M. Jung for providing the FLUXCOM GPP products.

REFERENCES

[1] A. Porcar-Castell et al., "Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges," *J. Exp. Botany*, vol. 65, no. 15, pp. 4065–4095, Aug. 2014.

[2] X. Liu et al., "A simple approach to enhance the TROPOMI solar-induced chlorophyll fluorescence product by combining with canopy reflected radiation at near-infrared band," *Remote Sens. Environ.*, vol. 284, Jan. 2023, Art. no. 113341.

[3] T. S. Magney et al., "Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence," *PNAS*, vol. 116, no. 24, pp. 11640–11645, Jun. 2019.

[4] Y. Zhang and J. Peñuelas, "Combining solar-induced chlorophyll fluorescence and optical vegetation indices to better understand plant phenological responses to global change," *J. Remote Sens.*, vol. 3, Sep. 2023, Art. no. 0085.

[5] P. Köhler, L. Guanter, and J. Joiner, "A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data," *Atmos. Meas. Techn.*, vol. 8, pp. 2589–2608, Jun. 2015.

[6] J. Joiner, Y. Yoshida, A. P. Vasilkov, Y. Yoshida, L. A. Corp, and E. M. Middleton, "First observations of global and seasonal terrestrial chlorophyll fluorescence from space," *Biogeosciences*, vol. 8, pp. 637–651, Mar. 2011.

[7] J. Joiner et al., "Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: Methodology, simulations, and application to GOME-2," *Atmos. Meas. Techn.*, vol. 6, pp. 2803–2823, Oct. 2013.

[8] C. Frankenberg et al., "Prospects for chlorophyll fluorescence remote sensing from the orbiting carbon observatory-2," *Remote Sens. Environ.*, vol. 147, pp. 1–12, May 2014.

[9] L. Guanter et al., "The TROPOSIF global sun-induced fluorescence dataset from the Sentinel-5P TROPOMI mission," *Earth Syst. Sci. Data*, vol. 13, no. 11, pp. 5423–5440, Nov. 2021.

[10] P. Köhler, C. Frankenberg, T. S. Magney, L. Guanter, J. Joiner, and J. Landgraf, "Global retrievals of solar-induced chlorophyll fluorescence with TROPOMI: First results and intersensor comparison to OCO-2," *Geophys. Res. Lett.*, vol. 45, no. 19, pp. 10 456–10 463, Oct. 2018.

[11] S. Du et al., "Retrieval of global terrestrial solar-induced chlorophyll fluorescence from TanSat satellite," *Sci. Bull.*, vol. 63, pp. 1502–1512, Nov. 2018.

[12] S. Du, X. Liu, J. Chen, and L. Liu, "Prospects for solar-induced chlorophyll fluorescence remote sensing from the SIFIS payload onboard the TECIS-1 satellite," *J. Remote Sens.*, vol. 2022, Sep. 2022, Art. no. 9845432.

[13] G. Duveiller and A. Cescatti, "Spatially downscaling sun-induced chlorophyll fluorescence leads to an improved temporal correlation with gross primary productivity," *Remote Sens. Environ.*, vol. 182, pp. 72–89, Sep. 2016.

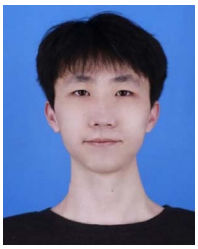
[14] G. Duveiller et al., "A spatially downscaled sun-induced fluorescence global product for enhanced monitoring of vegetation productivity," *Earth Syst. Sci. Data*, vol. 12, no. 2, pp. 1101–1116, May 2020.
[15] J. Wen et al., "A framework for harmonizing multiple satellite instruments to generate a long-term global high spatial-resolution solar-induced chlorophyll fluorescence (SIF)," *Remote Sens. Environ.*, vol. 239, Mar. 2020, Art. no. 111644.
[16] Y. Ma, L. Liu, X. Liu, and J. Chen, "An improved downscaled sun-induced chlorophyll fluorescence (DSIF) product of GOME-2 dataset," *Eur. J. Remote. Sens.*, vol. 55, no. 1, pp. 168–180, Dec. 2022.
[17] Y. Zhang, J. Joiner, S. H. Alemohammad, S. Zhou, and P. Gentine, "A global spatially contiguous solar-induced fluorescence (CSIF) dataset using neural networks," *Biogeosciences*, vol. 15, no. 19, pp. 5779–5800, Oct. 2018.
[18] P. Gentine and S. H. Alemohammad, "Reconstructed solar-induced fluorescence: A machine learning vegetation product based on MODIS surface reflectance to reproduce GOME-2 solar-induced fluorescence," *Geophys. Res. Lett.*, vol. 45, no. 7, pp. 3136–3146, Mar. 2018.
[19] J. Gensheimer, A. J. Turner, P. Köhler, C. Frankenberg, and J. Chen, "A convolutional neural network for spatial downscaling of satellite-based solar-induced chlorophyll fluorescence (SIFnet)," *Biogeosciences*, vol. 19, no. 6, pp. 1777–1793, Mar. 2022.
[20] Z. Zhang, W. Xu, Q. Qin, and Z. Long, "Downscaling solar-induced chlorophyll fluorescence based on convolutional neural network method to monitor agricultural drought," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1012–1028, Feb. 2021.
[21] Y. Ma, L. Liu, R. Chen, S. Du, and X. Liu, "Generation of a global spatially continuous TanSat solar-induced chlorophyll fluorescence product by considering the impact of the solar radiation intensity," *Remote Sens.*, vol. 12, no. 13, Jun. 2020, Art. no. 2167.
[22] X. Li, J. Xiao, and B. He, "Chlorophyll fluorescence observed by OCO-2 is strongly related to gross primary productivity estimated from flux towers in temperate forests," *Remote Sens. Environ.*, vol. 204, pp. 659–671, Jan. 2018.
[23] L. Yu, J. Wen, C. Y. Chang, C. Frankenberg, and Y. Sun, "High-resolution global contiguous SIF of OCO-2," *Geophys. Res. Lett.*, vol. 46, no. 3, pp. 1449–1458, Feb. 2019.
[24] C. Frankenberg et al., "New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity," *Geophys. Res. Lett.*, vol. 38, no. 17, Sep. 2011, Art. no. L17706.
[25] Z. Y. Zhang, J. M. Chen, L. Guanter, L. M. He, and Y. G. Zhang, "From canopy-leaving to total canopy far-red fluorescence emission for remote sensing of photosynthesis: First results from TROPOMI," *Geophys. Res. Lett.*, vol. 46, no. 21, pp. 12030–12040, Nov. 2019.
[26] C. Schaaf and Z. Wang, "MCD43A4 MODIS/Terra+Aqua nadir BRDF-adjusted reflectance daily L3 global 500m," NASA LP DAAC, 2015.
[27] R. G. Allen, L. S. Pereira, D. Raes, and M. Smith, "Crop evapotranspiration: Guidelines for computing crop water requirements," *FAO Irrigation Drainage*, vol. 56, pp. 147–151, Jan. 1998.
[28] M. Jung et al., "The FLUXCOM ensemble of global land-atmosphere energy fluxes," *Sci. Data*, vol. 6, May 2019, Art. no. 74.
[29] M. Jung et al., "Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach," *Biogeosciences*, vol. 17, pp. 1343–1365, Mar. 2020.
[30] W. Yuan et al., "Global estimates of evapotranspiration and gross primary production based on MODIS and global meteorology data," *Remote Sens. Environ.*, vol. 114, pp. 1416–1431, Jul. 2010.
[31] M. A. Friedl et al., "Global land cover mapping from MODIS: Algorithms and early results," *Remote Sens. Environ.*, vol. 83, pp. 287–302, Nov. 2002.
[32] B. Dechant et al., "Canopy structure explains the relationship between photosynthesis and sun-induced chlorophyll fluorescence in crops," *Remote Sens. Environ.*, vol. 241, May 2020, Art. no. 111733.
[33] D. D. Baldocchi et al., "Outgoing near-infrared radiation from vegetation scales with canopy photosynthesis across a spectrum of function, structure, physiological capacity, and weather," *JGR Biogeosciences*, vol. 125, no. 7, Jul. 2020, Art. no. e2019JG005534.
[34] L. Liu et al., "Estimating maize GPP using near-infrared radiance of vegetation," *Sci. Remote Sens.*, vol. 2, Dec. 2020, Art. no. 100009.
[35] J. Kong et al., "Matching high resolution satellite data and flux tower footprints improves their agreement in photosynthesis estimates," *Agricultural Forest Meteorol.*, vol. 316, Apr. 2022, Art. no. 108878.
[36] C. Jiang, K. Guan, G. Wu, B. Peng, and S. Wang, "A daily, 250 m and real-time gross primary productivity product (2000–present) covering the contiguous United States," *Earth Syst. Sci. Data*, vol. 13, no. 2, pp. 281–298, Feb. 2021.
[37] Y. Zhang et al., "ChinaSpec: A network for long-term ground-based measurements of solar-induced fluorescence in China," *JGR Biogeosci.*, vol. 126, no. 3, Mar. 2021, Art. no. e2020JG006042.
[38] S. W. Maier, K. P. Günther, and M. Stellmes, "Sun-induced fluorescence: A new tool for precision farming," in *Digital Imaging and Spectral Techniques: Applications to Precision Agriculture and Crop Physiology*, vol. 66. Madison, WI, USA: American Society of Agronomy, Jan. 2004, pp. 207–222.
[39] L. Guanter et al., "Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements," *Remote Sens. Environ.*, vol. 121, pp. 236–251, Jun. 2012.
[40] N. C. Parazoo et al., "Towards a harmonized long-term spaceborne record of far-red solar-induced fluorescence," *JGR Biogeosci.*, vol. 124, no. 8, pp. 2518–2539, Aug. 2019.
[41] J. A. Berry et al., "New methods for measurement of photosynthesis from space," *Geophys. Res. Lett.*, vol. 38, Apr. 2013, Art. no. L17706.
[42] Y. Yoshida et al., "The 2010 Russian drought impact on satellite measurements of solar-induced chlorophyll fluorescence: Insights from modeling and comparisons with parameters derived from satellite reflectances," *Remote Sens. Environ.*, vol. 166, pp. 163–177, Sep. 2015.
[43] L. Gu, J. Han, J. D. Wood, C. Y. Chang, and Y. Sun, "Sun-induced Chl fluorescence and its importance for biophysical modeling of photosynthesis based on light reactions," *New Phytologist*, vol. 223, pp. 1179–1191, Aug. 2019.
[44] J. Hu, J. Jia, Y. Ma, L. Liu, and H. Yu, "A reconstructed global daily seamless SIF product at 0.05 degree resolution based on TROPOMI, MODIS and ERA5 data," *Remote Sens.*, vol. 14, no. 6, Mar. 2022, Art. no. 1504.
[45] J. Hu, L. Liu, J. Guo, S. Du, and X. Liu, "Upscaling solar-induced chlorophyll fluorescence from an instantaneous to daily scale gives an improved estimation of the gross primary productivity," *Remote Sens.*, vol. 10, no. 10, Oct. 2018, Art. no. 1663.
[46] A. Damm et al., "Far-red sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An assessment based on observational and modeling approaches," *Remote Sens. Environ.*, vol. 166, pp. 91–105, Sep. 2015.
[47] S. M. Njuki, C. M. Mannaerts, and Z. Su, "An improved approach for downscaling coarse-resolution thermal data by minimizing the spatial averaging biases in random forest," *Remote Sens.*, vol. 12, no. 21, Oct. 2020, Art. no. 3507.
[48] Y. Zeng et al., "Combining near-infrared radiance of vegetation and fluorescence spectroscopy to detect effects of abiotic changes and stresses," *Remote Sens. Environ.*, vol. 270, Mar. 2022, Art. no. 112856.
[49] A. J. Turner, P. Kohler, T. S. Magney, C. Frankenberg, I. Fung, and R. C. Cohen, "Extreme events driving year-to-year differences in gross primary productivity across the US," *Biogeosciences*, vol. 18, no. 24, pp. 6579–6588, Dec. 2021.
[50] L. Zhang, Z. Zhang, Y. Luo, J. Cao, and F. Tao, "Combining optical, fluorescence, thermal satellite, and environmental data to predict county-level maize yield in China using machine learning approaches," *Remote Sens.*, vol. 12, no. 1, Dec. 2019, Art. no. 21.
[51] B. Peng et al., "Assessing the benefit of satellite-based solar-induced chlorophyll fluorescence in crop yield prediction," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 90, Aug. 2020, Art. no. 102126.

Jiaochan Hu received the B.S. degree in remote sensing science and technology from Wuhan University, Wuhan, China, in 2012, and the Ph.D. degree in cartography and geographic information system from the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2019.

She is currently an Associate Professor with Green Shipping and Carbon Neutrality Lab, College of Environmental Sciences and Engineering, Dalian Maritime University, Dalian, China. Her research interests include remote sensing and the application of solar-induced chlorophyll fluorescence.

Jia Jia received the B.S. degree in environmental engineering from the College of Resources and Environment, Linyi University, Linyi, China, in 2020, and the M.S. degree in environmental science and engineering in 2023 from Green Shipping and Carbon Neutrality Lab, College of Environmental Science and Engineering, Dalian Maritime University, Dalian, China, where she is currently working toward the Ph.D. degree in computer science and technology with the Center of Hyperspectral Imaging in Remote Sensing, Information Science and Technology College.

Her research interest is data reconstruction of the satellite-based chlorophyll fluorescence.



Zihan Ma received the B.S. degree in computer science and technology from the College of Information Science and Engineering, Shandong Normal University, Jinan, China, in 2023. He is currently working toward the M.S. degree in computer technology with the Center of Hyperspectral Imaging in Remote Sensing (CHIRS), Information Science and Technology College, Dalian Maritime University, Dalian, China.

His research interests include data reconstruction of the satellite-based chlorophyll fluorescence based on deep learning.

Keyu Yuan received the B.S. degree in environmental engineering from the College of Chemical Engineering, North China University of Science and Technology, Tangshan, China, in 2022. She is currently working toward the M.S. degree in environmental science and engineering with Green Shipping and Carbon Neutrality Lab, College of Environmental Science and Engineering, Dalian Maritime University, Dalian, China.

Her research interest is the retrieval of solar-induced chlorophyll fluorescence based on satellite data.

Haoyang Yu (Member, IEEE) received the B.S. degree in information and computing science from Northeastern University, Shenyang, China, in 2013, and the Ph.D. degree in cartography and geographic information system from the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2019.

He is currently an Associate Professor with the Center of Hyperspectral Imaging in Remote Sensing, Information Science and Technology College, Dalian Maritime University, Dalian, China. His research focuses on models and algorithms for hyperspectral image processing, analysis, and applications.

Liangyun Liu received the Ph.D. degree in optical engineering from the Xi'an Institute of Optical and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China, in 2000.

He is currently a Professor with Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. His research majors in quantitative remote sensing of vegetation.

Dr. Liu was a recipient of five government prizes, including two second-class national scientific and technological progress awards, and was also honored by the National Science Fund for Distinguished Young Scholars in 2018.