

## Comprehensive evaluation of global CI, FVC, and LAI products and their relationships using high-resolution reference data

Sijia Li<sup>a,b,\*</sup>, Hongliang Fang<sup>a,b</sup>, Yinghui Zhang<sup>a,b</sup>, Yao Wang<sup>a,b</sup>

<sup>a</sup> LREIS, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China

<sup>b</sup> College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100049, China



### ARTICLE INFO

#### Keywords:

Canopy clumping index (CI)  
Fractional vegetation cover (FVC)  
Leaf area index (LAI)  
Leaf projection function at nadir (G(0))  
Vertical structural characteristics

### ABSTRACT

In recent decades, several global clumping index (CI), fractional vegetation cover (FVC), and leaf area index (LAI) products have been generated using a range of optical satellite sensors. It is essential for the application community to understand the accuracy and uncertainty of these products. However, current validation studies have limited spatial coverage and temporal continuity, and the relationships between different products are not clearly understood. This study aims to validate the most recent global CAS-CI, GEOV2 FVC, and MODIS LAI products using the Ground-Based Observations for Validation (GBOV) and DIRECT 2.1 datasets. The relationships between these products, the leaf projection function in the nadir direction (G(0)), and the vertical structural characteristics of the forest were also analyzed. The GEOV2 FVC and MODIS LAI show good performance, whereas CAS-CI shows a slight underestimation (bias = -0.10) and strong seasonal variations. FVC and LAI are strongly correlated at different spatial scales [field, high-resolution (30 m and upscaled 3 km), and moderate-resolution (3 km)], particularly for forests [mean Pearson's correlation coefficients ( $r$ ) = 0.90]. The CI-FVC/LAI relationship is more consistent across spatial scales for grass and shrubs than for forests. The derived G(0) overestimates the reference value (bias = 0.28) mainly due to inconsistencies between the CI, FVC, and LAI products. The overstory CI is very close to the overall forest CI, indicating that the understory may be neglected when estimating the overall CI. The overstory G(0) is similar to the overall G(0) for most forest types, except for the evergreen needleleaf forest. The overall CI-FVC/LAI relationship is mainly influenced by the overstory, while the FVC-LAI relationship is affected by both the overstory and understory. Overall, the CI-FVC/LAI relationship should be further evaluated and applied cautiously across different scales. The forest vertical structural characteristics revealed in this study are important for field measurement, remote sensing estimation, and modeling studies.

### 1. Introduction

The vegetation clumping index (CI), fractional vegetation cover (FVC), and leaf area index (LAI) are critical canopy structural parameters. CI describes the degree of foliage clumping compared to that of a random distribution (Chen and Black 1992; Fang 2021a; Nilson 1971). FVC represents the proportion of the vertical projection area covered by green vegetation (Gitelson et al., 2002), and LAI is defined as one-half of the total green leaf area per unit ground area (GCOS 2016). These variables determine radiative transfer, hydrological processes, and vegetation photosynthesis (Al-Kaisi et al., 1989; Gitelson et al., 2014; Steltzer and Welker 2006) and have been widely used for the study of terrestrial carbon and water cycles (Chen et al., 2019; Zeng et al., 2018).

CI, FVC, and LAI are typically measured using indirect optical methods, such as digital hemispheric photography (DHP), LAI-2200 plant canopy analyzer, and terrestrial LiDAR (Fang et al., 2014b; Wang and Fang 2020; Zhao et al., 2012). These parameters have been included in a series of continuous field measurement networks, such as the National Ecological Observatory Network (NEON) (Kao et al., 2012), Terrestrial Ecosystem Research Network (TERN) (Karan et al., 2016), and Integrated Carbon Observation System (ICOS) (Gielen et al., 2018). To facilitate the use of measurements from operational ground-based monitoring networks and their comparison to remote sensing products, the Ground-Based Observations for Validation (GBOV) service (<https://land.copernicus.eu/global/gbov>) was initiated by the European Commission as a part of the Copernicus Global Land Service

\* Corresponding author. LREIS, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China.  
E-mail address: lisj.19b@igsnrr.ac.cn (S. Li).

(CGLS). At present, the GBOV vegetation section has incorporated most of the NEON sites and provides consistent, continuous, and quality-controlled ground measurements and upscaled high-resolution reference data.

In recent decades, numerous global CI, FVC, and LAI products have been derived from a range of optical satellite sensors, such as AVHRR, MODIS, POLDER, VEGETATION, VIIRS, and PROBA-V (Baret et al., 2007, 2013; Wei et al., 2019; Yan et al. 2016a, 2018). For CI, the biome-specific empirical relationships with the normalized difference between the hotspot and dark spot (NDHD) index are extensively used in the product generation (Chen et al., 2005; Wei and Fang 2016). For FVC and LAI, the underlying algorithms used for product generation mainly include the lookup table (LUT) method based on the radiative transfer model (RTM), machine learning, and hybrid methods (Baret et al., 2007; Jia et al., 2015; Yan et al., 2018). These products have been utilized as important inputs in land surface models and global change research (Fang et al., 2019a; Zhu et al., 2016).

The uncertainty and accuracy associated with these products are critical for their proper use in various applications (Lawrence et al., 2019; Morisette et al., 2006). The uncertainty information can be obtained through the validation of global products with upscaled high-resolution reference data (Fang et al., 2019a; Fernandes et al., 2014; Morisette et al., 2006). The best practices and protocols for product validation have been proposed by the Land Product Validation (LPV) Subgroup of the Committee on Earth Observation Satellites (CEOS) (<http://lpvs.gsfc.nasa.gov/>). Under the LPV guidelines, several validation campaigns, such as BigFoot (Cohen and Justice 1999), Validation of Land European Remote Sensing Instruments (VALERI) (Weiss 1991), and Implementation of Multiscale Agricultural Indicators Exploiting Sentinels (ImagineS) (<http://fp7-imagine.eu/>) have been conducted to establish valuable reference datasets via scaled-up field measurements with high-resolution imagery. Most validation studies are based on these high-resolution datasets and provide product quality information for different biomes on a global scale (Camacho et al., 2013; He et al., 2012; Yan et al., 2016b). Several time-series validations have been conducted on a regional scale to obtain the temporal performance of the remote sensing products (Fang et al., 2019b; Song et al., 2021).

Nevertheless, there are two deficiencies in current validation exercises. The first is the limitation of spatial coverage and temporal continuity resulting from the shortage of global continuous reference datasets. The evaluations of products on a global scale are based on historical datasets that are typically conducted at the peak of the growing season and are thus limited in their ability to capture the temporal performance of global products (Baret et al., 2006; Weiss 1991). Most time-series validations have been conducted at extremely limited sites by individual field campaigns (Fang et al., 2019b; Song et al., 2021). Second, the relationships between different biophysical products have not been clearly evaluated. Some empirical relationships between CI, FVC, and LAI have been established on the ground (Chen 1996; Fang et al., 2014a; Ryu et al., 2010) and at remote sensing product scales (Fang et al., 2021; Wei et al., 2019) and have been applied in remote sensing studies (Nikolov and Zeller 2006; Zhao et al., 2020). However, current validation studies have only evaluated different biophysical parameters separately (Camacho et al., 2013; Fuster et al., 2020; Yan et al., 2016b) and the relationships between the parameters have not been comprehensively evaluated.

The leaf projection function ( $G(\theta)$ ), defined as the average projection of a unit leaf area in the direction  $\theta$ , is another crucial canopy structural variable for radiation transfer modeling and remote sensing retrieval (Ross 1981; Sellers 1985). In particular, the leaf projection function in the nadir direction ( $G(0)$ ) is important because it may provide a solution for the derivation of the nadir CI (Fang 2021a). Recently, global monthly  $G(0)$  maps were generated for the first time based on the Beer-Lambert law using the global CI, FVC, and LAI products (Fang et al., 2021). However, the validity of the method and the output  $G(0)$  map have not been adequately evaluated using reference data.

Forests are typically characterized by vertical structure and are comprised of the overstory and understory (Wang and Fang 2020). The vertical distribution of forest structural variables greatly influences forest carbon and water cycles, as well as radiative transfer processes (Fang 2020). Some studies have reported the importance of the overstory and understory for the overall LAI estimation in field measurements (Bond-Lamberty et al., 2002; Chianucci et al., 2014a) and remote sensing studies (Brown et al., 2020; Liu et al., 2017). However, the vertical characteristics of FVC, CI, and  $G(0)$  and the relationships between these variables are still poorly understood.

The objective of this study is to comprehensively evaluate the global CI, FVC, and LAI products and their relationships. Global high-resolution reference and field data were used in the validation to enhance our understanding of the vegetation structural variables. Specifically, this study aims to:

- (1) validate the global CI, FVC, and LAI products and evaluate their relationships;
- (2) examine the validity of the  $G(0)$  estimation method and the output  $G(0)$  map;
- (3) analyze the vertical characteristics of the structural variables and their relationships.

## 2. Materials and methods

### 2.1. Global moderate-resolution products

#### 2.1.1. CAS-CI

CAS-CI (V1.1) (Table 1) is derived from the MODIS bidirectional reflectance distribution function (BRDF) product (MCD43A1, V6) with a spatial resolution of 500 m and a temporal resolution of 8 days based on an improved NDHD method (Wei et al., 2019). In the algorithm, the background influence is minimized through the adoption of different solar zenith angles according to FVC, i.e. SZA = 60° and the observational SZA were used to calculate the NDHD index for the sparsely vegetated (FVC < 25%) and the moderate to dense vegetation (FVC ≥ 25%), respectively (Wei and Fang 2016). A look-up table (LUT) based on the improved NDHD method was developed to estimate the daily CI, which was then aggregated into an 8-day product. A quality flag layer is attached to indicate whether the input BRDF is derived from the full or magnitude inversion (Wei et al., 2019).

#### 2.1.2. GEOV2 FVC

The GEOV2 FVC (V2.0) product (Table 1) is generated from SPOT/VEGETATION (SPOT/VGT) and PROBA-V observations on a 10-day basis and at 1 km using a neural network framework over the global BELMANIP sites (Baret et al., 2006). The neural network takes the corrected CYCLOPES FVC product as training data, while the daily synthesis of the top of canopy reflectance and viewing geometry from SPOT/VGT (1998–May 2014) and PROBA-V (May 2014–present) are set as input data. To ensure consistency, the PROBA-V FVC was rescaled to the SPOT/VGT FVC using a third polynomial conversion. Compared to the V1 product, multi-step filtering, temporal smoothing, gap filling, and compositing techniques have been applied to ensure continuity (Verger et al., 2014). Quality indicators are associated with the product to provide information about the number of valid daily estimates in the composition period, the length in days of composition, and its uncertainty.

#### 2.1.3. MODIS LAI

The MODIS LAI C6 product (MCD15A2H) (Table 1) is acquired from the MODIS sensors onboard the Terra and Aqua platforms every 8 days at a spatial resolution of 500 m (Myneni et al., 2015). The main algorithm is based on a LUT generated from a 3-D RTM (Knyazikhin et al., 1998a) where the clumping effect of the canopy is partly considered by assigning a priori CI value for each vegetation type (Knyazikhin et al.,

**Table 1**

Global moderate-resolution remote sensing products evaluated in this study. LUT: look-up table, NDHD: normalized difference hotspot and darkspot index, NIR: near infrared, MIR: middle infrared, NN: neural network.

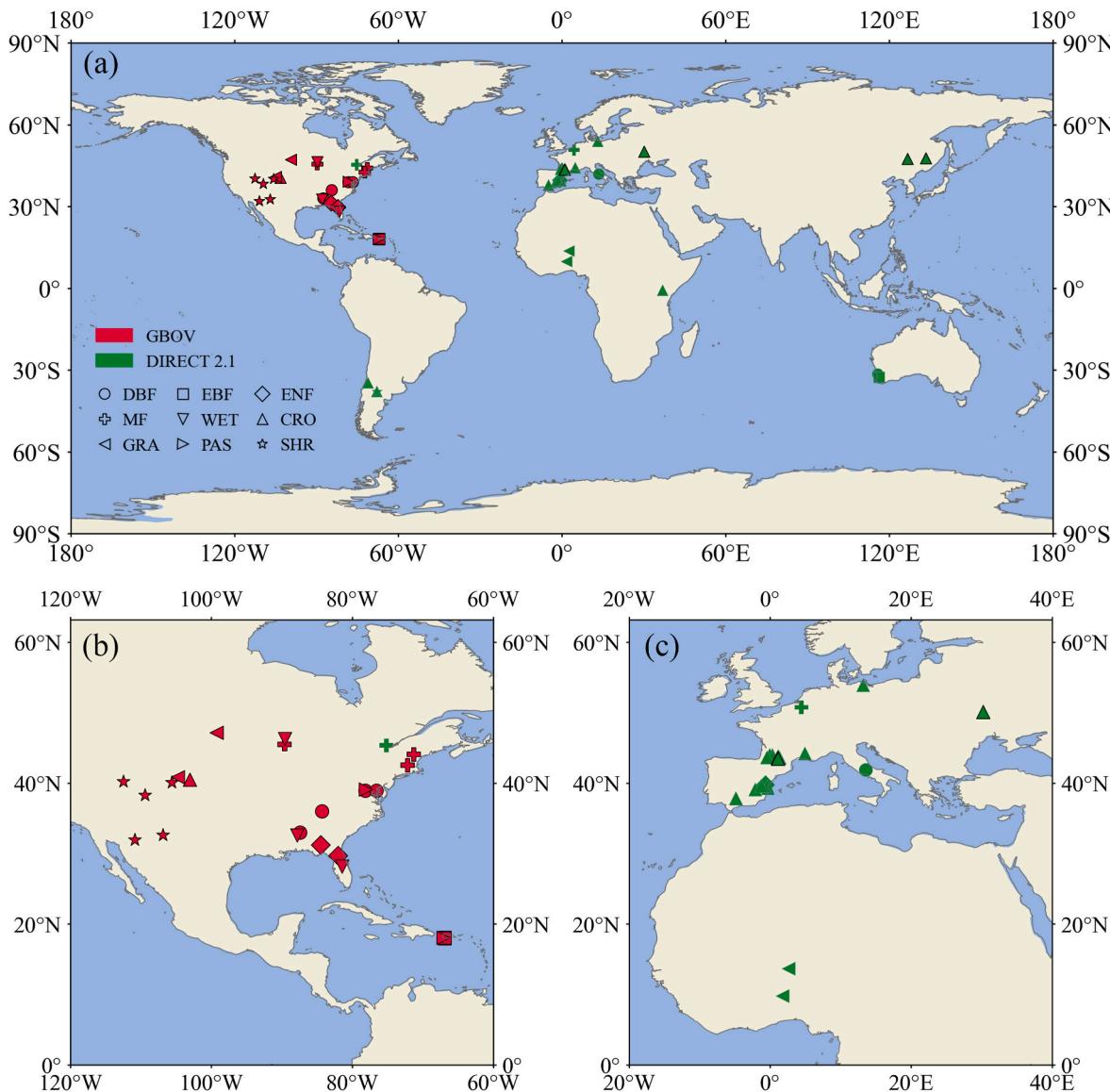
Variable	Products	Version	Spatial resolution	Temporal resolution	Algorithms	References
CI	CAS-CI	V1.1	500 m	8-day	NDHD-CI (red)	(Wei and Fang 2016; Wei et al., 2019)
FVC	GEOV2	V2.0	1 km	10-day	NN (red, NIR, MIR)	Verger et al. (2014)
LAI	MODIS	C6	500 m	8-day	LUT (red, NIR)	(Yan et al. 2016a, 2016b)

1998b). Therefore, the product is closer to the true LAI than the effective LAI (Yan et al., 2016b). The mean value of all candidate LAI estimates from the main algorithm was taken as the final solution, and the standard deviation of all acceptable LAI values served as the theoretical uncertainty. When the main algorithm is invalid, a backup algorithm employs the empirical relationship between LAI and the normalized difference vegetation index (NDVI) for different biome types to estimate LAI (Myndeni 1997). A quality indicator layer is attached, indicating whether the LAI value is derived from the main or backup algorithm.

## 2.2. Ground measurements and high-resolution reference data

### 2.2.1. GBOV dataset

The V2.0 version of the GBOV ground measurements and high-resolution reference maps (i.e. reference measurement (RM) and land product (LP) dataset in GBOV) were used in this study (<https://land.copernicus.eu/global/gbov/dataaccessLP/>). The dataset has been exploited in other validation exercises to validate the PROBA-V and MODIS biophysical products (EOLAB 2019). Current GBOV vegetation measurements are mainly based on the NEON sites (Fig. 1 and Table A1). For each site, multiple elemental sampling units (ESUs) (generally three



**Fig. 1.** The locations of global GBOV and DIRECT 2.1 field measurement sites used in this study (a). (b) and (c) show the sites in North America and Europe, respectively. CRO: Cultivated crops, DBF: Deciduous broadleaf forest, EBF: Evergreen broadleaf forest, ENF: Evergreen needleleaf forest, GRA: Grassland/herbaceous, MF: Mixed forest, PAS: Pasture/hay, SHR: Shrub/scrub, WET: Woody wetlands. The bold circles indicate those sites with >5 continuous measurements.

ESUs) with a nominal extent of 20 m × 20 m or 40 m × 40 m are routinely measured using DHP every 2 weeks during the growing season (Meier et al., 2018). Twelve DHP images are collected in a cross-pattern within a given ESU using a Nikon digital single-lens reflex camera (D750, D800, or D810) equipped with a full-frame fisheye lens (AF Fisheye-Nikkor 16 mm f/2.8D). Only downward-facing DHP images are collected for short canopies or understory, whereas both upward-facing and downward-facing images are acquired for forest plots with the overstory and understory.

LAI and effective LAI ( $\text{LAI}_e$ ) are estimated from the DHP (Miller 1967; Warren Wilson 1963) and the CI is derived as follows (Chen and Cihlar 1995; Fang 2021a):

$$CI = \frac{\text{LAI}_e}{\text{LAI}} \quad (1)$$

FVC is calculated as follows:

$$FVC = 1 - \overline{p(0^\circ)} \quad (2)$$

where  $\overline{p(0^\circ)}$  is the mean gap fraction in a zenith ring centered near the nadir ( $\pm 10^\circ$ ). Only the ground measurements with the best quality indicator ( $QA = 1$ ) were used for further analysis.

For the forest sites, the understory and overstory  $\text{LAI}_e$ , LAI, and FVC were combined to obtain the overall value for each ESU (Eqs. (3)–(5)), following the strategy used in the VALERI project (Baret et al., 2005):

$$\text{LAI}_e = \text{LAI}_{e_{\text{over}}} + \text{LAI}_{e_{\text{under}}} \quad (3)$$

$$\text{LAI} = \text{LAI}_{\text{over}} + \text{LAI}_{\text{under}} \quad (4)$$

$$FVC = FVC_{\text{over}} + (1 - FVC_{\text{over}}) \times FVC_{\text{under}} \quad (5)$$

where the subscripts *over* and *under* represent the structural variables of the overstory and understory, respectively. Subsequently, the overall CI was derived as the ratio of the  $\text{LAI}_e$  to the LAI (Eq. (1)).

The high-resolution LAI,  $\text{LAI}_e$ , and FVC over a 3 km × 3 km area centered on each site are derived from Landsat 8 OLI (30 m) and Sentinel-2 MSI (20 m) data, based on empirical transfer functions between vegetation indices and ground measurements. Quality control layers are attached to indicate whether the data is generated by extrapolating beyond the input and output ranges of the dataset used to establish the transfer function. The minimal and maximal days of year (DOYs) associated with each transfer function are also provided.

## 2.2.2. DIRECT 2.1 dataset

The DIRECT 2.1 ground database provided by CEOS LPV compiles LAI, FVC, and fraction of absorbed photosynthetically active radiation (FAPAR) averaged values over a 3 km × 3 km area (<https://calvalportal.ceos.org/lpv-direct-v2.1>) (Fig. 1). It contains 176 sites and 280 LAI samples, 122 FVC and 128 FAPAR samples (2000–2021) collected from a range of field campaigns conducted on relatively flat and homogenous sites, such as BigFoot (Cohen and Justice 1999), VALERI (Weiss 1991), ImagineS (<http://fp7-imagines.eu/>), the Earth Observation Laboratory (EOLAB), and ESA. The sampling scheme, ground measurements, and *in situ* data upscaling with high-resolution imagery were performed under the framework recommended by the CEOS LPV subgroup (<http://lpvs.gsfc.nasa.gov/>). For the forest sites, both overstory and understory information are considered in the DIRECT 2.1.

The vegetation structural field measurement data for northeastern China crops (NECC) was obtained from seasonal field campaigns carried out to obtain continuous measurements of crop structural variables in Honghe (2012, 2013, and 2019) and Hailun (2016) in NE China (Fang 2021b). The Honghe site ( $47.65^\circ\text{N}$ ,  $133.51^\circ\text{E}$ ) comprises homogeneous paddy rice, while the Hailun site ( $47.41^\circ\text{N}$ ,  $126.82^\circ\text{E}$ ) is made up of maize, sorghum, and soybean. Five plots were selected at each site. Typically, four elemental sampling units (approximately 15 m × 15 m) were sampled from each plot to reduce random sampling errors. The CI,

$\text{LAI}_e$ , LAI, FVC, and average leaf inclination angle (ALA) were extracted from the DHP data by post-processing using the CAN-EYE software (Weiss and Baret, 2017). Further information on the sampling protocol and data processing are available in the literature (Fang et al. 2014a, 2018; Zhang et al., 2021). This dataset is now open to access in PANGAEA (<https://doi.pangaea.de/10.1594/PANGAEA.939444>) and has been integrated into DIRECT 2.1.

## 2.2.3. Generation of NECC high-resolution reference data

To bridge the scale gap between the NECC dataset and global products, NECC high-resolution reference data was generated with cloud-free HJ-1, harmonized Landsat, and Sentinel-2 (HLS) (Claverie et al., 2018) imagery based on random forest models for 2012, 2016, and 2019, respectively, with the exception of 2013 due to relatively few ground measurements. The HLS V1.4 dataset with a spatial resolution of 30 m (L30 and S30) starting from 2013 was used for upscaling (<https://hls.gsfc.nasa.gov>). HJ-1 images accessed from the China Centre for Resources Satellite Data and Application (CRESDA, <http://218.247.138.119:7777/DSSPlatform/index.html>) were used as a complement to HLS to enhance cloudless frequency and expand high-resolution reference data to 2012. The HJ-1 images were first processed with geometrical and atmospheric corrections. Then, the HJ-1 reflectance was normalized to the view and illumination geometry of HLS using the c-factor technique (Roy et al., 2016) which was also used in the HLS generation (Claverie et al., 2018), and its spectral bands were linearly adjusted to the corresponding HLS bands to maintain consistency.

In the random forest model, the reflectances of the blue, green, red, and near-infrared bands were used as explanatory variables, whereas  $\text{LAI}_e$ , LAI, and FVC were used as the explained variables. The model was trained with yearly ground measurements and assessed by 10-fold cross-validation, for which nine-tenths of the samples were used for training and the remainder of the samples were used to evaluate the model. This step was run 10 times until all samples were looped. The optimal hyperparameters of the model (e.g. the number of trees and the maximum depth of the tree) were determined by selecting parameters corresponding to the best performance after iterative running (model performance is shown in Fig. S1). See Table S2 to obtain the NECC high-resolution reference data.

In this study, only sites including all  $\text{LAI}_e$ , LAI, and FVC high-resolution reference maps from 2003 to 2019 (Fig. 1, Table A1) were used to match the global remote sensing products.

## 2.3. Evaluation of global products and their relationships

### 2.3.1. Validation of the global products

The remote sensing products for the period 2003–2019 were validated following the protocol recommended by CEOS LPV (<http://lpvs.gsfc.nasa.gov/>). For consistency, both remote sensing products and high-resolution reference data within the 3 km × 3 km area around each site were calculated and compared.

The remote sensing LAI and FVC for the area were calculated with a simple averaging algorithm from the original data (Table 1). The remote sensing  $\text{LAI}_e$  (500 m) was first calculated from the LAI and CI ( $\text{LAI}_e = \text{LAI} \times \text{CI}$ ) before spatial averaging to 3 km. Subsequently, the remote sensing CI was calculated using Eq. (1) for the entire area. For CAS-CI and MODIS LAI, only good quality pixels from the main retrieval algorithms and only those from the direct retrieval method for GEOF2 FVC were considered in the validation. A similar calculation method was applied to the high-resolution reference data. For the GBOV high-resolution reference data, only those acquired within the minimal and maximal DOYs associated with the transfer function were used in the validation. During the calculation, we have only considered the data pairs whose good retrievals were >50% within the area around each site. Remote sensing products were validated with coherent high-resolution reference data within the compositing periods. All validations were performed under the WGS-84 geographic coordinate system.

A new structural variable, specific vegetation cover (SVC), defined as the ratio of FVC to LAI, was proposed to characterize the ability of vegetation to cover the ground per unit leaf area (Fang et al., 2021):

$$SVC = \frac{FVC}{LAI} \quad (6)$$

SVC was also validated, accompanied by CI, FVC, and LAI. The coefficient of determination ( $R^2$ ), bias, root mean square error (RMSE), and the relative RMSE (RRMSE), calculated by dividing the RMSE by the average reference value, were selected as the evaluation metrics. The uncertainty agreement ratio (UAR), indicating the percentage of samples that satisfy the uncertainty requirements (Brown et al., 2020), was calculated to provide a benchmark for the practical use of the product:

$$UAR(\%) = \frac{N_{good}}{N_{all}} \cdot 100\% \quad (7)$$

where  $N_{good}$  is the number of samples that satisfy the uncertainty requirements and  $N_{all}$  represents the number of samples used for validation. The uncertainty requirements for UAR calculation are 0.05 unit or 10% for FVC, CI, SVC, and G(0) and 15% for LAI, as defined by GCOS (2016). Time-series validation was conducted to assess the performance of these products over different biome types and seasonal phases.

### 2.3.2. Evaluation of relationships between CI, FVC, and LAI

The relationships between CI, FVC, and LAI were evaluated at different scales. At the field scale, the relationships were computed for sites with continuous measurements (Fig. 1, Table A1). The relationships from high-resolution reference data (30 m) were evaluated with those from field measurements at the ESU level. Next, the high-resolution data were upscaled to 3 km, and the resulting relationships were compared with those obtained from moderate-resolution remote sensing products. The relationship between any two variables was expressed by the Pearson correlation coefficients ( $r$ ) at a confidence level of 95%.

### 2.4. Estimation of $G(0)$

Under the turbid medium assumption, the relationship between the CI, FVC, and LAI can be expressed by the Beer-Lambert law (Nilson 1971):

$$p(\theta) = \exp^{-\frac{G(\theta)*LAI*\cos(\theta)}{\cos(\theta)}} \quad (8)$$

where  $p(\theta)$ ,  $CI(\theta)$ , and  $G(\theta)$  denote the gap fraction, CI, and leaf projection function in the zenith direction  $\theta$ , respectively. Specifically, the gap fraction in the nadir direction can be expressed as:

$$p(0^\circ) = 1 - FVC \quad (9)$$

Afterward, the leaf projection function at the nadir direction can be estimated from:

$$G(0) \cdot CI(0) = \frac{\ln(1 - FVC)}{CI(0) * LAI} \quad (10)$$

In Eq. (10), the CI at the nadir direction ( $CI(0)$ ) can be approximated with the whole CI obtained from remote sensing (Fang et al., 2021):

$$G(0) \cdot CI \approx \frac{\ln(1 - FVC)}{CI * LAI} \quad (11)$$

Through field measurements of ALA, the reference  $G(0)$  can be obtained by assuming that the leaf angle distribution can be expressed by an ellipsoidal function (Campbell 1990):

$$G(\theta) \cdot ALA = \frac{\sqrt{(\chi^2 + \tan^2 \theta)} \cos \theta}{\chi + 1.774(\chi + 1.182)^{-0.73}} \quad (12)$$

where  $\chi$  represents the length ratio of the horizontal to the vertical semi-axis and can be obtained from the ALA:

$$\chi = -3 + \left( \frac{ALA}{9.65} \right)^{-0.6061} \quad (13)$$

This study evaluated the influence of different CI schemes on the  $G(0)$  estimation with NECC field data by comparing the  $G(0)$  obtained through Eqs. 10–12.

### 2.5. Assessment of the vertical structural characteristics in forest

The structural variables of the overall canopy, overstory, and understory were compared for the typical sites with the GBOV ground data. The impact of excluding the understory on the overall structural variables was evaluated by the relative difference (RD) between the overstory and the overall values:

$$RD(\%) = \frac{Var_{over} - Var_{all}}{Var_{all}} \cdot 100\% \quad (14)$$

where  $Var_{over}$  and  $Var_{all}$  represent the structural variables of the overstory and the overall canopy, respectively.

The relationships between CI, FVC, and LAI were analyzed separately for the overall canopy, overstory, and understory using field measurement data from the GBOV. The analysis was performed at the site level where the CI, FVC, and LAI were aggregated from ESUs within the sites.

## 3. Results

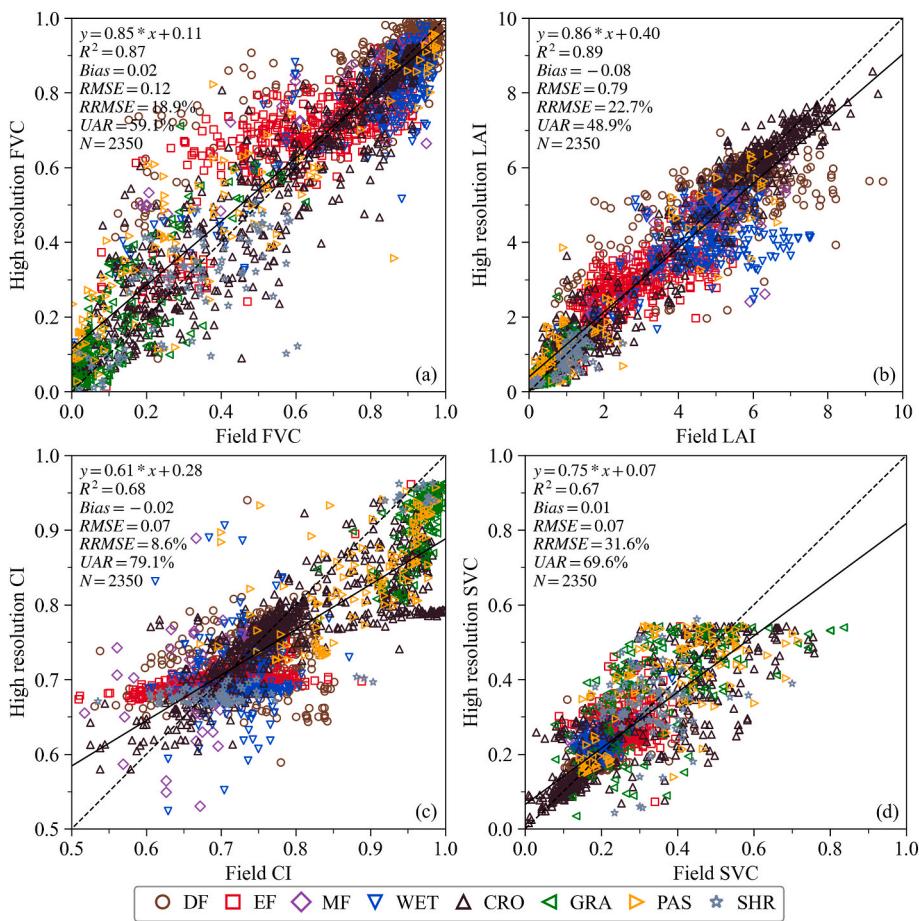
### 3.1. Evaluation of CI, FVC, and LAI products and their relationships

#### 3.1.1. Product validation

Fig. 2 shows a comparison of high-resolution CI, FVC, LAI, and SVC with field measurements at the ESU level. The high-resolution FVC and LAI agree well with field measurements ( $R^2 > 0.87$ ) and show better performance than the high-resolution CI and SVC. The high-resolution CI slightly underestimates the field CI, especially for high values ( $>0.85$ ), which mainly occurs for GRA. An upper limit of 0.56 (Fig. 2(d)) is observed for the high-resolution SVC, which mainly occurs for grassland and crops when LAI and FVC are very low, accounting for a small minority proportion and does not affect the main results. This upper limit is related to the empirical transfer function used to generate the GBOV high-resolution reference data. GBOV uses an exponential formula to fit the vegetation index and LAI but a linear relationship to fit the vegetation index and FVC. In general, DF, EF, MF, and WET have higher LAI and FVC and are characterized by more clustered foliages ( $CI < 0.8$ ) and a relatively weak leaf coverage capacity ( $SVC < 0.4$ ). In contrast, grassland, PAS, and SHR have lower LAI and FVC but higher CI and SVC than the other types. The crop parameters vary larger seasonally than those of the other types due to the strong seasonal variation of crops.

Fig. 3 presents a comparison between moderate-resolution products and high-resolution reference data. Overall, the FVC and LAI products show a high degree of correspondence with the reference data ( $R^2 = 0.83$  and 0.72, respectively). In contrast, CI and SVC compare poorly with the reference data (bias = -0.10 and 0.07, respectively). Although FVC and LAI show high  $R^2$ , their RRMSE are larger than 30%. The CI product underestimates the reference value (bias = -0.10). SVC show a scattered distribution (RRMSE = 83.0%). Nevertheless, CI and SVC are characterized by a higher UAR (31.6% and 35.6%) than FVC and LAI.

Specifically, the moderate-resolution FVC product slightly overestimates high-resolution reference data for forest types with strong seasonal variation (i.e. DBF, MF, and WET) when  $FVC > 0.8$ , but shows underestimations when  $FVC < 0.8$ . While the EBF FVC overestimates the reference FVC (bias = 0.07). EBF shows good performance among all biomes ( $R^2 = 0.80$ , RMSE = 0.09, and UAR = 56%; Table 2). For ENF, FVC presents both overestimation and underestimation and is characterized by a poor accuracy ( $R^2 = 0.19$  and UAR = 15%; Table 2). For



**Fig. 2.** Comparison of the high-resolution CI, FVC, LAI, and SVC with field measurements at the ESU level. The solid and dash lines represent the regression and 1:1 lines, respectively. DF: deciduous forest; EF: evergreen forest. See Fig. 1 for other acronyms.

non-forest types, different degrees of overestimation are observed for CRO, GRA, PAS, and SHR (bias = 0.14, 0.14, 0.13, and 0.07, respectively) (Table 2), especially when FVC < 0.4. Nevertheless, the moderate-resolution FVC is highly correlated with the high-resolution FVC for non-forest types (all  $R^2 > 0.75$ ) (Table 2).

As shown in Fig. 3, the LAI product shows underestimation for MF (bias = -0.68) but overestimates the reference values of EBF (bias = 1.76) (Table 2). The points of DBF and WET are scattered with both overestimation and underestimation, particularly when LAI < 5. For ENF, the LAI is highly consistent with the reference LAI (RMSE = 0.56) (Table 2). For non-forest types, when LAI is low, the estimation accuracy of the LAI product is good. However, the LAI product underestimates when LAI > 3, particularly for PAS and CRO. Generally, forests have higher RMSE and lower  $R^2$  than non-forests, whereas greater RRMSE is found for non-forests, especially for grass (RRMSE = 110%) (Table 2).

The CI product shows the best consistency for CRO (RRMSE = 12% and UAR = 63%; Table 2). For GRA, PAS, and forests, the CI product significantly underestimates the high-resolution reference CI with biases up to -0.16 (Table 2). The CI product exhibits a larger variation for SHR, whereas the high-resolution CI is stabler ( $\sim 0.7$ ). As the variation of field and reference CI is relatively low for SHR (Fig. 2(c)), this implies that the CI product may overestimate its seasonal amplitude.

SVC shows poor correspondence with the reference values for all types (Fig. 3), indicative of issues in the original FVC and LAI products. The forest values are typically gathered in clusters with a lower bias and  $R^2$  (Table 2), whereas the non-forests values are highly scattered. SVC overestimates the reference values for CRO, SHR, PAS, and GRA with biases of up to 0.16 but underestimates EBF (bias = -0.11) (Table 2). In addition, for SHR, a distinct difference exists when high-resolution SVC

$$= 0.$$

Fig. 4 compares the time series of the moderate-resolution products with the reference data for several typical sites (Fig. B1). In general, these moderate-resolution products well capture the seasonal variation for almost all sites, except for the EBF site. The FVC product overestimates the reference value in the mid-growing season, especially for non-forest sites. However, FVC underestimates the reference value for forest sites in the early and late growing seasons. For evergreen forest sites, FVC shows stronger seasonal variations than the high-resolution reference.

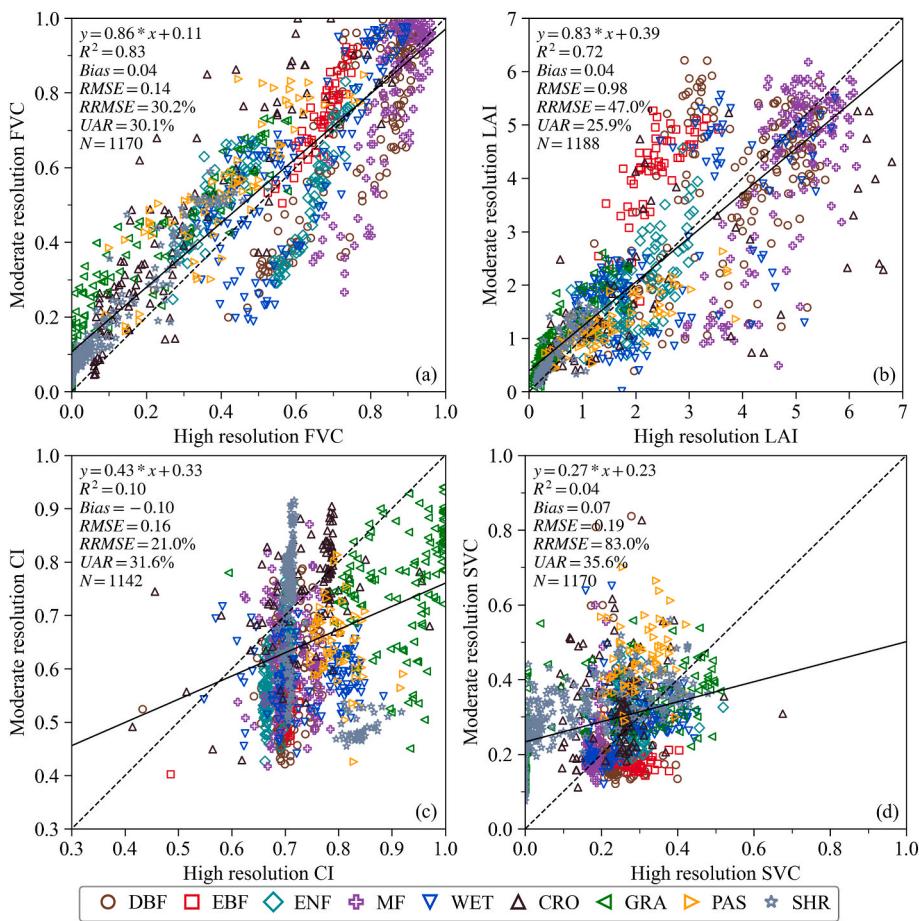
The LAI product presents high fluctuations for forest sites in the mid-phenological stage, which partly leads to the scattering of forest values, as shown in Fig. 3(b). In addition, LAI underestimates the high-resolution LAI for DBF, ENF, and MF during the green-up and senescence periods.

The CI product exhibits significant seasonal fluctuations up to 0.5, especially for non-forests, whereas smooth seasonal variations are observed for the high-resolution reference CI (< 0.2). This indicates that the CI product may overestimate the seasonal amplitude.

SVC shows shallow bowl-shaped seasonal variations, especially for biomes with high seasonality, such as DBF, MF, WET, and PAS. The PAS site shows a large discrepancy between the moderate-resolution SVC and the reference value, mainly during the green-up and senescence stages.

### 3.1.2. Evaluation of relationships

Fig. 5 shows the relationships between CI, FVC, and LAI at different spatial resolutions for each site. Generally, most CI-FVC/LAI relationships are negative, while FVC is positively correlated with LAI. The CI-



**Fig. 3.** Comparison of the moderate-resolution CI, FVC, LAI, and SVC with the high-resolution reference data at site level. The solid line and dash line indicate the fitting line and 1:1 line, respectively. See Fig. 1 for the biome type acronyms.

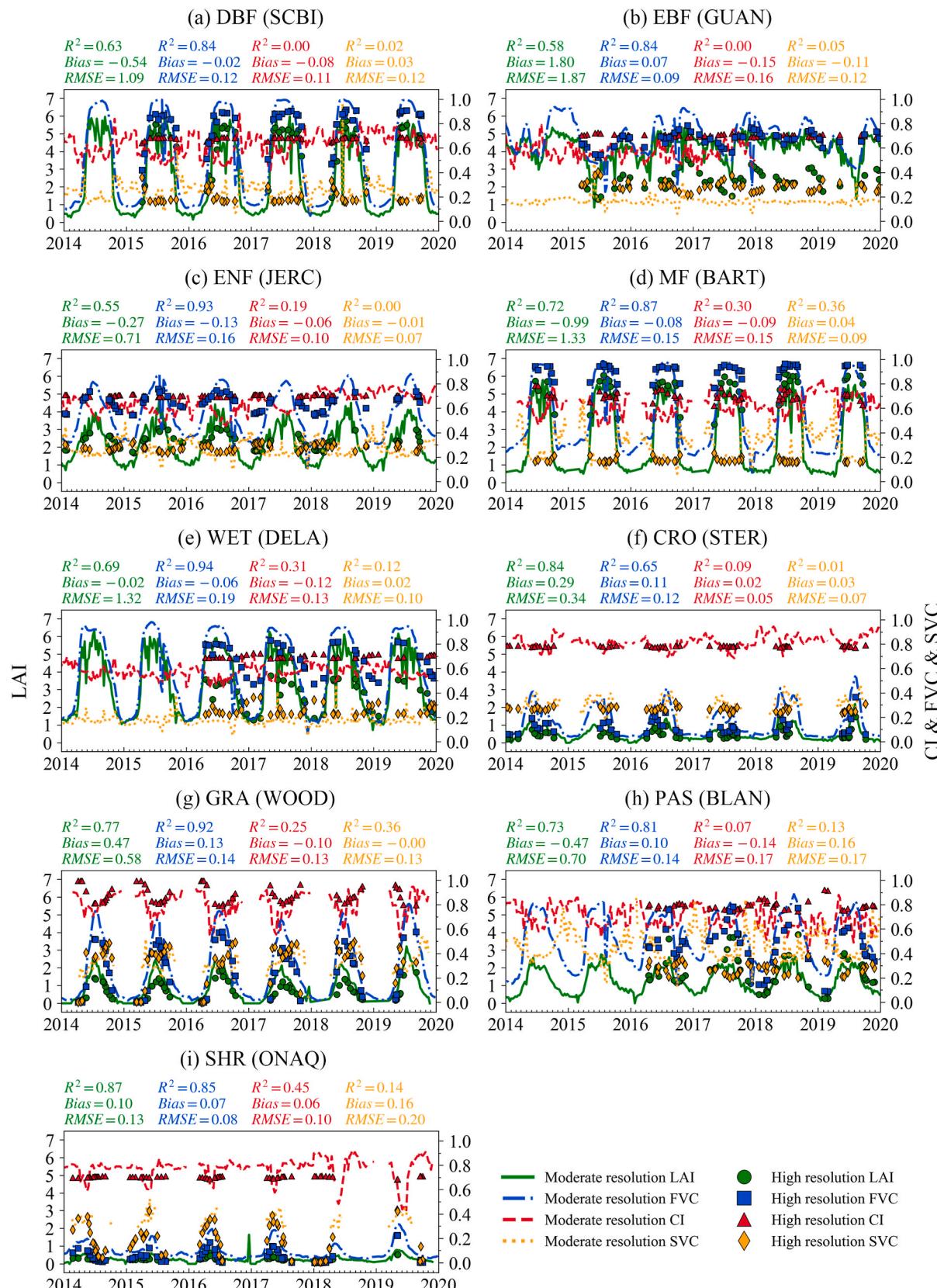
**Table 2**

Statistics from validation of the moderate-resolution products using high-resolution reference data for different biome types.

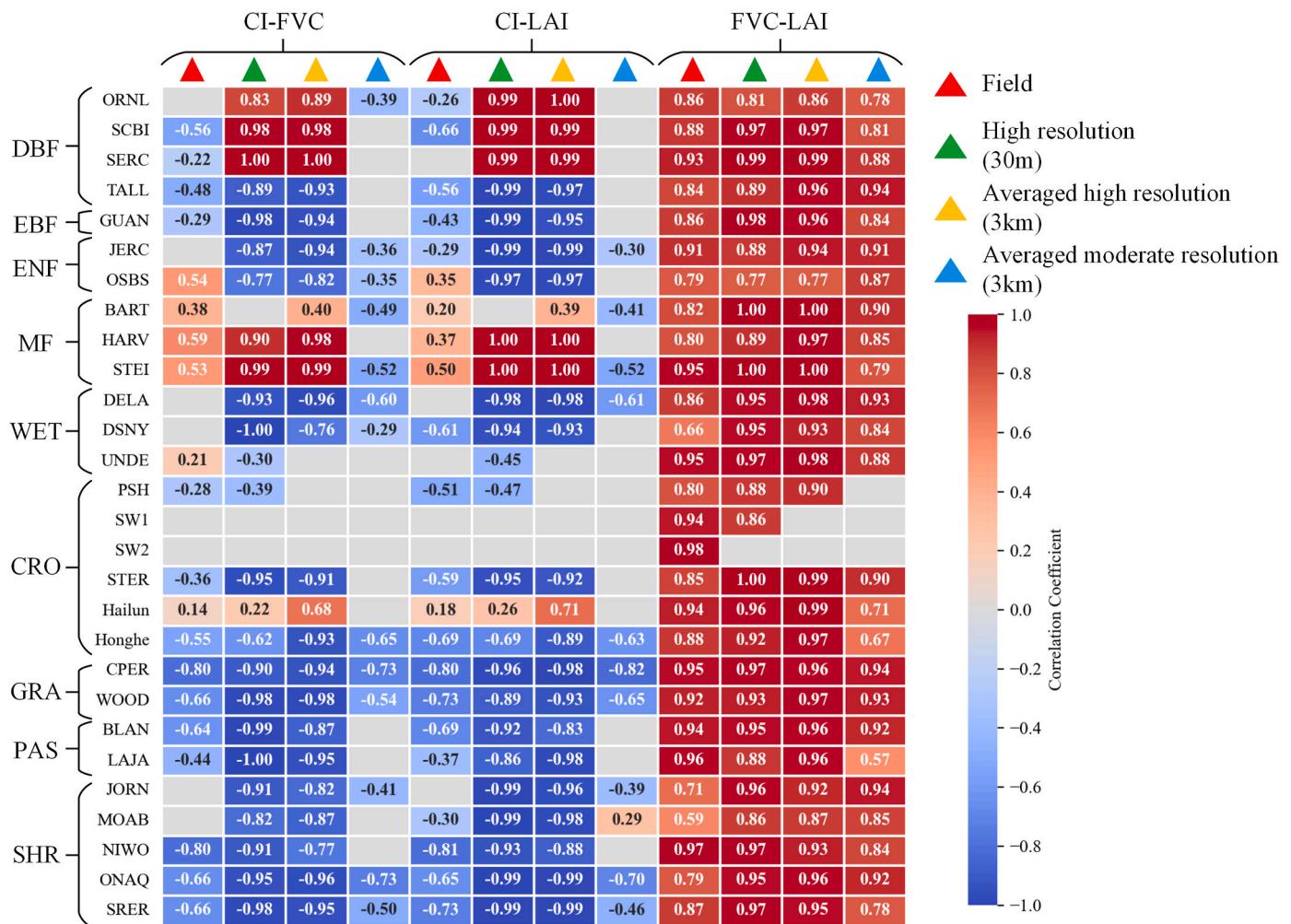
		DBF	EBF	ENF	MF	WET	CRO	GRA	PAS	SHR
FVC	$R^2$	0.72	0.80	<b>0.19</b>	0.71	0.65	0.78	0.86	0.75	0.93
	Bias	-0.04	<b>0.07</b>	0.01	-0.04	0.01	<b>0.14</b>	<b>0.14</b>	<b>0.13</b>	0.07
	RMSE	0.14	0.09	0.15	0.15	0.14	0.18	0.16	0.16	0.09
	RRMSE	18%	14%	29%	17%	23%	62%	104%	39%	109%
	UAR	42%	<b>56%</b>	<b>15%</b>	57%	34%	16%	16%	21%	23%
LAI	$R^2$	0.33	0.53	0.43	0.56	0.52	0.55	0.80	0.68	0.82
	Bias	0.01	<b>1.76</b>	-0.05	<b>-0.68</b>	-0.01	-0.15	0.41	-0.26	0.12
	RMSE	<b>1.26</b>	<b>1.85</b>	<b>0.56</b>	<b>1.29</b>	1.10	1.22	0.50	0.61	0.21
	RRMSE	33%	77%	30%	26%	44%	76%	<b>110%</b>	41%	65%
CI	UAR	45%	0%	30%	55%	22%	16%	4%	34%	15%
	$R^2$	0.03	0.15	0.40	0.00	0.00	0.23	0.09	0.07	0.54
	Bias	<b>-0.13</b>	<b>-0.15</b>	<b>-0.09</b>	<b>-0.10</b>	<b>-0.16</b>	-0.01	<b>-0.16</b>	<b>-0.15</b>	-0.03
	RMSE	0.16	0.16	0.11	0.14	0.18	<b>0.09</b>	0.19	0.17	0.17
SVC	RRMSE	22%	23%	16%	20%	25%	12%	20%	22%	23%
	UAR	25%	12%	38%	30%	16%	<b>63%</b>	25%	21%	37%
	$R^2$	0.00	0.00	0.03	0.04	0.08	0.01	0.01	0.08	0.36
	Bias	-0.01	-0.11	0.03	0.05	0.00	<b>0.09</b>	<b>0.13</b>	<b>0.15</b>	<b>0.16</b>
G(0)	RMSE	0.11	0.12	0.08	0.17	0.10	0.16	0.39	0.17	0.20
	RRMSE	49%	41%	30%	92%	36%	67%	<b>176%</b>	58%	<b>126%</b>
	UAR	<b>54%</b>	7%	48%	<b>73%</b>	47%	27%	19%	7%	16%
	$R^2$	0.00	0.36	0.22	0.01	0.04	0.01	0.10	0.04	0.61
G(0)	Bias	0.22	<b>-0.03</b>	0.16	<b>0.41</b>	0.26	0.32	0.28	<b>0.57</b>	0.31
	RMSE	0.56	<b>0.22</b>	0.31	0.81	0.42	0.69	0.51	0.65	0.36
	RRMSE	90%	<b>32%</b>	54%	129%	72%	175%	170%	133%	161%
	UAR	15%	18%	8%	13%	15%	12%	25%	0%	10%

FVC correlation is similar to that between the CI and LAI. The CI-FVC/LAI relationships show relatively weak consistency across the increasing spatial scale. Specifically, at the field scale, the CI shows

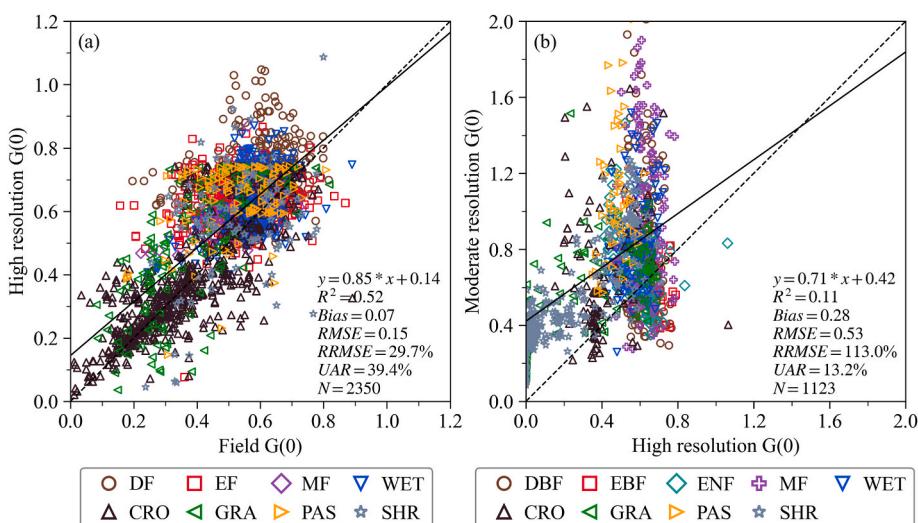
negative or non-significant correlations with FVC and LAI for most sites, whereas for MF sites, the CI-FVC/LAI relationships are positive. The CI-FVC/LAI relationships for the high-resolution reference data coincide



**Fig. 4.** Comparisons of time-series moderate-resolution LAI, FVC, CI, and SVC with high-resolution reference data at typical sites for different biome types. One most representative site for each biome type is selected (Table S1). The statistical metrics for LAI, FVC, CI, and SVC are displayed in different colors. See Table A1 and Fig. 1 for the site and biome type acronyms, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 5.** The correlations between CI, FVC, and LAI obtained from field measurements, high-resolution reference data (30 m), spatially averaged high-resolution data (3 km), and the moderate-resolution remote sensing products (3 km), respectively. The gray cell represents there is no significant correlation between two variables. See Table A1 and Fig. 1 for the site and biome type acronyms, respectively.



**Fig. 6.** Comparisons of G(0) derived from high-resolution data and field measurements (a), and between remote sensing products and high-resolution data (b). The solid line and dash line indicate the fitting line and 1:1 line, respectively. See Figs. 1 and 2 for the biome type acronyms.

with those of the field measurements for EBF, MF, and non-forests. However, for other biomes, the high-resolution reference data are unable to fully capture the CI-FVC/LAI relationships at the field measurement level. The CI-FVC/LAI relationships of the remote sensing products are consistent with those of the high-resolution reference data for non-forests, especially for GRA, characterized by correlations with identical signs. Positive FVC-LAI correlations are well maintained for most sites at different spatial scales. However, PSH, SW1, and SW2 show no significant correlations.

### 3.2. Validation of $G(0)$

The high-resolution reference  $G(0)$  generally agrees well with the  $G(0)$  from the field measurements ( $R^2 = 0.52$  and  $RMSE = 0.15$ ), but a slight overestimation (bias = 0.07) is observed, especially for  $G(0) > 0.4$  (Fig. 6(a)). The moderate-resolution  $G(0)$  significantly overestimates the high-resolution reference data (bias = 0.28,  $RMSE = 0.53$ ) (Fig. 6(b)). Among all biome types, EBF showed the best correspondence between the moderate-resolution and high-resolution reference  $G(0)$  (bias = -0.03) (Table 2), whereas the moderate-resolution  $G(0)$  for PAS and MF products largely overestimate their reference values (bias = 0.57 and 0.41, respectively) (Table 2). The outlying  $G(0)$  values ( $G(0) > 1$ ) shown in Fig. 6(b) are attributed to the inconsistent CI, FVC, and LAI values.

Fig. 7 shows the influence of the two CI schemes on the  $G(0)$  estimations. The  $G(0)$  estimated with the nadir CI agrees well with the reference  $G(0)$  from the average leaf inclination angle (Fig. 7(a)). In contrast, the  $G(0)$  from whole CI underestimates reference  $G(0)$  up to 0.26 and 0.30 for maize and soybean, respectively (Fig. 7(b)). This difference is caused by the difference between the whole CI and  $CI(0)$ . The whole CI coincides with  $CI(0)$  for rice and sorghum but is larger than  $CI(0)$  for maize and soybean (Fig. 7(c)).

### 3.3. Vertical structural characteristics in forests

Fig. 8 shows the seasonal variations in canopy structural variables and the influence of excluding the understory for different forest types at typical sites. The overstory is characterized by a higher FVC and LAI than the understory, but for the ENF site, FVC and LAI of the understory and overstory are comparable. The CI and SVC of the overstory are less than those of the understory, and the understory is characterized by a stronger seasonal variation in the CI. There is no clear difference between the  $G(0)$  of the overstory and understory for the DBF, MF, and WET sites. However, for the ENF site, the  $G(0)$  of the overstory is less than 0.5, whereas the  $G(0)$  of the overstory is  $> 0.5$ .

Regarding the effects of excluding the understory, the negative RD

for FVC, LAI, and CI indicates that ignoring the understory will lead to an underestimation of up to 45.43%, 50.30%, and 13.34%, respectively (Fig. 8). Compared to FVC and LAI, the impact of excluding the understory on the overall CI is relatively lower ( $|RD| < 13.34\%$ ). The RD of SVC is positive (up to 30.91%), which indicates that neglecting the understory results in an increase in the overall SVC. Additionally, the contribution of excluding the understory to the overall  $G(0)$  is weak (<10%) for most biomes. However, for the ENF site, a negative RD up to -36.14% is observed (Fig. 8(f)), indicating that neglecting the understory will cause a large underestimation of the overall  $G(0)$  for this biome. The LAI RD is the largest during the mid-growing period, whereas this phenomenon is not common to other structural variables.

Fig. 9 shows the relationships between CI and FVC for the overall canopy, overstory, and understory. The overall CI-FVC relationship is dominated by the overstory for most woody biomes, except for WET, whose overall canopy relationship is mainly influenced by the understory, especially for  $FVC < 0.6$ . For the DBF overstory, CI is negatively correlated with FVC, while for the overstory of the other woody biomes, positive relationships are observed. In contrast, CI is negatively correlated with FVC for the understory of all woody biomes. The relationship between CI and LAI is highly similar to that between CI and FVC, as shown in Fig. S2.

Fig. 10 shows the relationships between FVC and LAI for the overall canopy, overstory, and understory, respectively. The FVC-LAI relationships for the overall canopy are similar to those of the overstory for DBF, MF, and WET. The slight relationship difference between the overall canopy and the overstory is because of the understory influence. For ENF, the overall relationship is similar to that of the understory.

## 4. Discussion

### 4.1. Product evaluation

As shown in Fig. 3, only 30.1% of the GEOV2 FVC are within the uncertainty requirement of 0.05 unit or 10%. This level of uncertainty is similar to previous results based on the DIRECT sites only ( $UAR = 33.0\%$ ) (Verger et al., 2014). However, a substantially increased number of reference samples (1170 vs. 24) covering different growing phases were utilized in the present study. GEOV2 FVC underestimates the reference values for MF, DBF, and WET in the early and late seasons, whereas it overestimates the crops, grasslands, and PAS (Fig. 3(a)). Overestimations have also been reported by other researchers for similar biome types (Liu et al. (2019)). The discrepancies between the FVC product and the reference data may be attributed to differences in their algorithms. The FVC product is derived from the red, NIR, and MIR

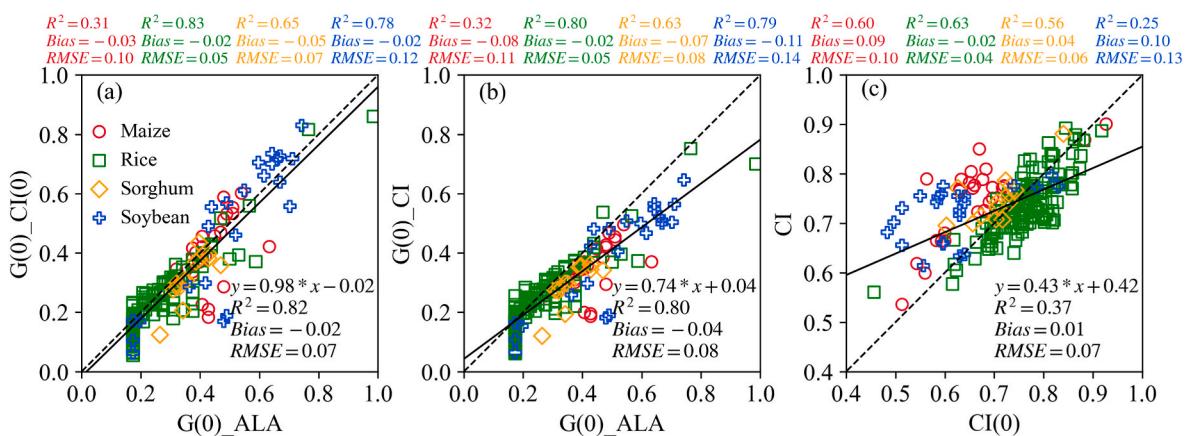
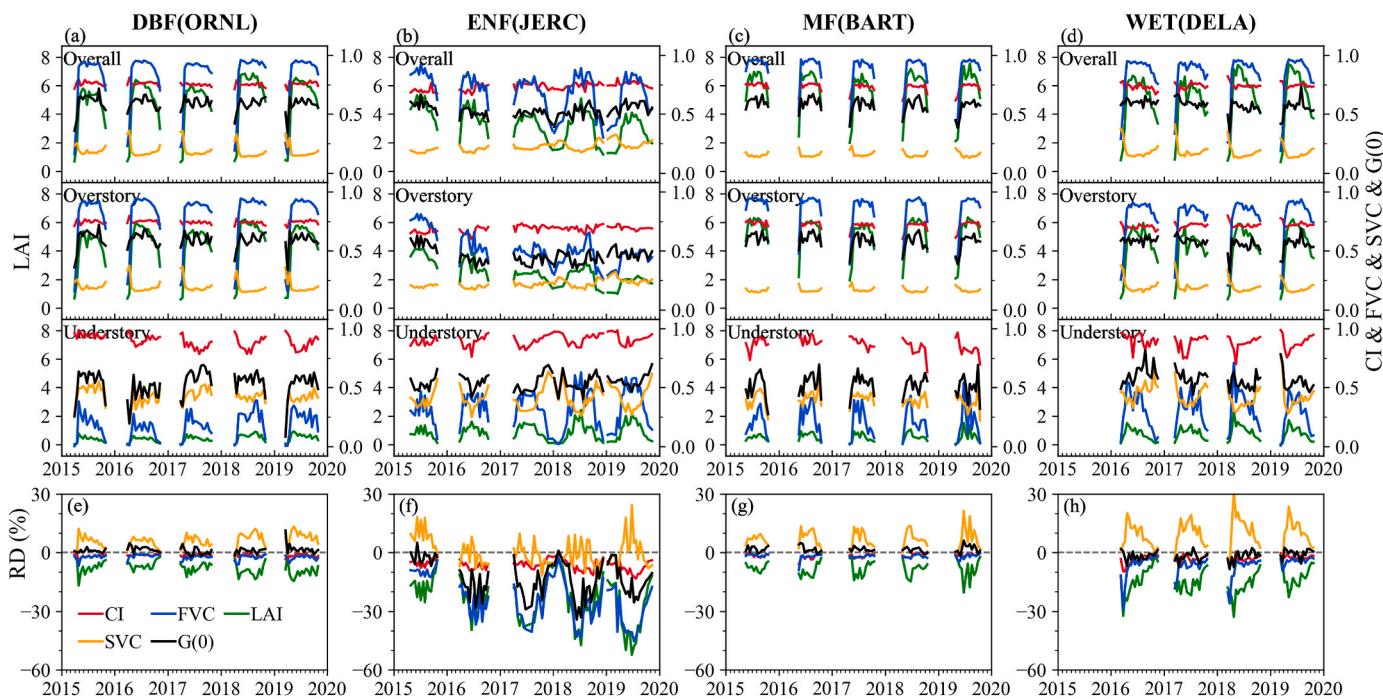
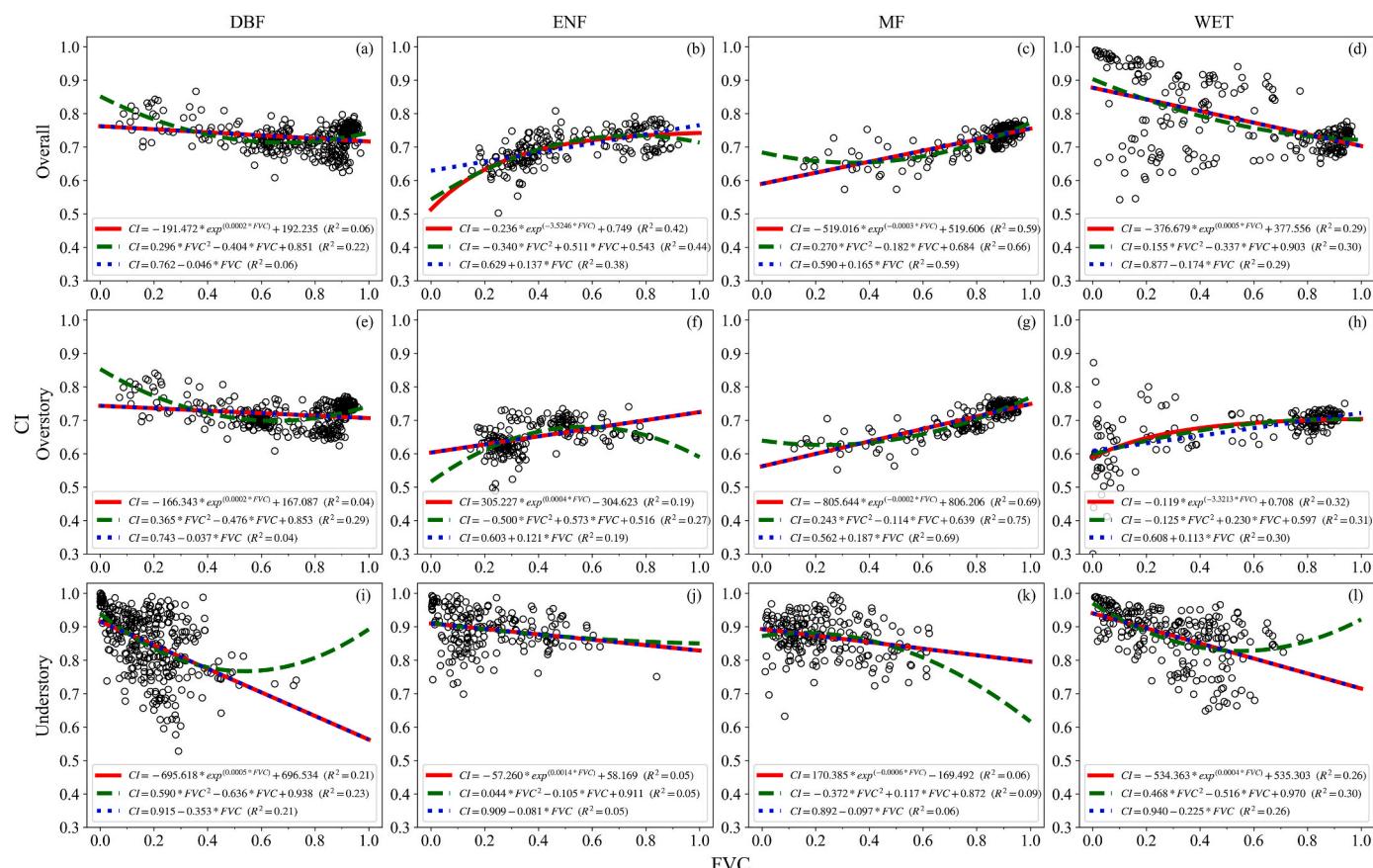


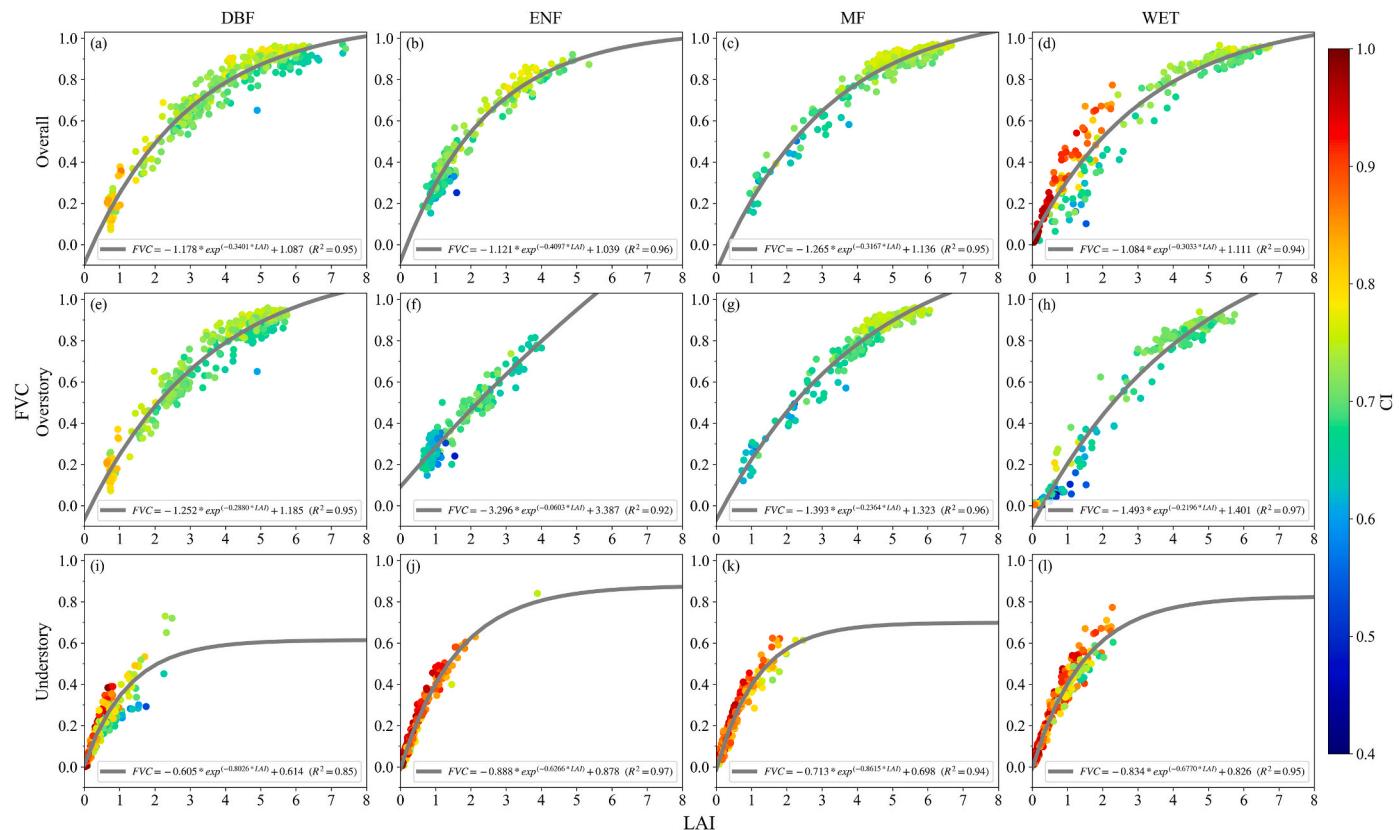
Fig. 7. Evaluation of two CI schemes on the  $G(0)$  estimation using field data from the NECC sites.  $G(0)_ALA$ : the  $G(0)$  from the average leaf angle,  $G(0)_CI(0)$ : the  $G(0)$  estimated from the nadir CI,  $G(0)_CI$ : the  $G(0)$  estimated from the whole CI. The statistical metrics for different crop types are displayed in different colors. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 8.** The seasonal variations of canopy structural variables and the influence of excluding the understory for different forest types. The upper panels show seasonal variations of CI, FVC, LAI, SVC, and G(0) for overall canopy, overstory, and understory, respectively. The lower panels show the relative differences (RD) of CI, FVC, LAI, SVC, and G(0) between the overall and overstory canopies. Only one most representative site for each biome type (Table S1) is shown. See Table A1 and Fig. 1 for the site and biome type acronyms, respectively.



**Fig. 9.** The relationships between CI and FVC at the site level for the overall canopy (upper), overstory (middle), and understory (lower), respectively. The red, green, and blue lines represent exponential, polynomial, and linear fittings, respectively. See Fig. 1 for the biome type acronyms. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 10.** The relationships between FVC and LAI at the site level for the overall canopy (upper), overstory (middle), and understory (lower), respectively. See Fig. 1 for the biome type acronyms.

reflectances using the neural network method (Table 1) (Verger et al., 2014), whereas the high-resolution reference data are generated by upscaling the field measurements by empirical transfer functions. The spectral reflectances used for the generation of the product are mainly influenced by green leaves with a higher chlorophyll content. However, for forests in the early and late growing seasons, the upward DHP used in the ground measurement is equally sensitive to woody tissue and green and senescent leaves because of the difficulty in distinguishing them (Fang et al., 2014b; Woodgate et al., 2015). This can lead to the larger high-resolution FVC than the FVC product. For non-forests, brown leaves still contribute slightly to the spectral reflectances used for the product generation, whereas the classification algorithm used for processing the downward DHP is highly sensitive to green leaves only (Meyer and Neto 2008). This results in a high-resolution FVC that may be lower than the FVC product. Currently, the global FVC products mainly generated with the neural network method (Jia et al., 2015; Verger et al., 2014), and the product accuracy is restricted by the accuracy and representativeness of the training samples. The increasing abundance of high-resolution reference data can be regarded as a training base for improving the accuracy of the FVC product.

The MODIS LAI shows a medium degree of correspondence with the high-resolution data ( $R^2 = 0.72$ , RMSE = 0.98) (Fig. 3(b)). This accuracy is comparable to that of recent validation studies conducted over the GBOV site ( $R^2 = 0.61\text{--}0.74$  and RMSE = 0.89–1.19) (Brown et al., 2020; EOLAB 2019), whereas it is relatively lower than that of previous validation studies over the BELMANIP2.1 sites ( $R^2 = 0.77$  and RMSE = 0.66) (Yan et al., 2016b). The UAR calculated here for MODIS LAI is 25.9% (Fig. 3(b)), indicating that further efforts should be made to improve existing products for meeting the requirement of LAI accuracy proposed by GCOS (2016). The systematic overestimation of MODIS LAI for EBF, also reported in another study (Yan et al., 2016b), may be attributed to the overestimation of the reflectance data because of cloud

contamination. The discrepancy for DBF and WET (Fig. 3(b)) can also be explained by the algorithm difference between the product and the reference data. These two biomes are characterized by a complex structure, which may lead to difficulties in the retrieval algorithm. Therefore, regional tuning of global biophysical generation algorithms is crucial.

Fig. 3 shows that 31.6% of CAS-CI are within the uncertainty requirement of 0.05 unit or 10%, which is higher than those of the GEOV2 FVC and MODIS LAI. Nevertheless, CAS-CI displays a slight underestimation for almost all biomes (bias = −0.09). Several factors may have contributed to this underestimation. First, the CI product typically captures structural information near the top of the canopy with a lower CI value, whereas the field CI responds to the lower layer of the canopy with more random foliage (Pisek et al., 2015). Second, the field CI in the study is estimated from DHP with the LX method (Lang and Yueqin 1986) and the clumping effect below the shoot level is not considered for needleleaf forests (Chen and Cihlar 1995; Ryu et al., 2010). In addition, the LX method cannot fully correct the clumping effect compared to the CLX method (Leblanc et al., 2005), resulting in an overestimation of the field CI. Fig. 4 shows that the seasonal amplitude of the CAS-CI was overestimated, especially for non-forests, which may be caused by the residual background impact and the input BRDF quality (section 2.1.1). This also contributes to the low consistency between CAS-CI and the high-resolution reference data ( $R^2 = 0.10$ , RMSE = 0.16, Fig. 3).

SVC was proposed as a new structural index to indicate the capability of an unit area foliage to cover the ground surface (Fang et al., 2021). A bowl shape is found for the SVC seasonal profile, especially for biomes with high seasonality, such as DBF, MF, WET, and PAS (Figs. 4 and 8). The SVC seasonal profile is off phase with the LAI profile, which is consistent with that found from global remote sensing data (Fang et al., 2021). GRA and CRO typically have a relatively higher SVC than the

other types for light interception (Fig. 2 and S3), which was also reported in a previous study (Fang et al., 2021).

This study validated SVC derived from the products and found better performance for forests than for non-forests (Fig. 3). SVC is a derived variable from FVC and LAI; therefore, its performance is determined by the consistency between FVC and LAI. For PAS, the SVC overestimation is caused by the FVC overestimation and the LAI underestimation. However, for MF, although simultaneous underestimations of FVC and LAI exist, SVC shows a smaller bias than that of PAS. The distinct difference between the moderate-resolution SVC and the reference value for SHR is caused by the inconsistency between the moderate-resolution and reference FVCs when FVC = 0 (Fig. 3(a)).

Different relationships between CI and FVC/LAI were found in the field and at high- and moderate-resolution scales for most biome types, particularly for DBF and MF (Fig. 5). For GRA, the negative CI-FVC/LAI relationships remain consistent across different spatial scales. Previous studies have established empirical relationships between CI, FVC, and LAI at different scales (Fang et al., 2014a; Ryu et al., 2010; Wei et al., 2019), and the relationships have been applied for remote sensing estimation (Nikolov and Zeller 2006; Zhao et al., 2020). The findings presented in this study suggest that caution should be exercised when applying empirical relationships across different spatial scales.

#### 4.2. G(0) estimation

In this study, G(0) was estimated based on the Beer-Lambert law (Eq. (10)) and shows relatively poor correspondence with the reference data ( $R^2 = 0.11$ , bias = 0.28). This is mainly attributed to the inconsistency among existing products. Indeed, the G(0) estimation can be enhanced when new, improved CI, FVC, and LAI products are available. The estimation of G(0) assumes that the whole CI is equal to the nadir CI (CI(0)). This assumption has been used in other studies (Chen et al., 1997; Pinty et al., 2006; Ryu et al., 2012) and can result in a bias of <0.3 in the estimated G(0) (Fig. 7(b)). CI(0) can be estimated through indirect optical measurement (Ryu et al., 2012) and empirical functions with the zenith angle (Fang et al., 2014b). Since the satellite CI(0) product is currently lacking, it is therefore necessary to generate the global CI(0) first in order to estimate the global G(0).

In addition, new FVC and LAI products derived from space-borne LiDAR, such as the Geoscience Laser Altimetry System (GLAS) (Tang et al., 2019) and Global Ecosystem Dynamics Investigation (GEDI) (Dubayah et al., 2020) may help improve the G(0) estimation. On the other hand, the global G(0) may also be derived using global field measured data, such as the TRY plant trait database (<https://www.try-db.org>) and other published datasets (Hinojo-Hinojo and Goulden 2020; Pisek and Adamson 2020).

#### 4.3. The understory impact

The total canopy is composed of both understory and overstory parts. Excluding the understory leads to a negative bias for the overall FVC (−45.43%) and LAI (−50.30%) (Fig. 8). Similar underestimations have been reported by previous studies on the ground (Bond-Lamberty et al., 2002; Chianucci et al., 2014b) and at remote sensing levels (Brown et al., 2020; Liu et al., 2017). The impact of excluding the understory on the overall LAI is found to be largest during the mid-growing period. However, for the other structural variables, the influence of excluding the understory is mixed (Fig. 8) because of the different combination methods for the overall values.

This study demonstrates that the understory has a higher CI than the overstory (Figs. 8–10). Similar findings have been reported in previous studies on ENF, EBF, DBF, and MF (Pisek et al., 2015). Neglecting the understory will lead to an RD not larger than −13.34% for CI (Fig. 8), suggesting that the overstory CI may be regarded as an approximation of the overall CI.

The understory has a higher SVC than the overstory, and neglecting

the understory will result in an increase in the overall SVC (up to 30.91%) (Fig. 8). This is because SVC is defined as the ratio of FVC to LAI, and neglecting the understory has a relatively smaller influence on FVC than on LAI. Therefore, caution should be exercised in interpreting the understory and overstory SVC.

For ENF, ignoring the understory underestimates the overall G(0). ENF is typically composed of a broadleaf understory and needleleaf overstory, and the broadleaf understory G(0) is generally larger than that of the needleleaf overstory (Arthur Sampson and Smith 1993; Wright et al., 2006). On the contrary, neglecting the understory has little impact on the overall G(0) for DBF, MF, and WET sites ( $|RD| < 10\%$ ) (Fig. 8), which suggests that the overstory G(0) can be regarded as the overall value for these biome types. The understory has little impact on the CI and FVC/LAI relationship for nearly all biome types except for WET (Fig. 9). This indicates that the CI-FVC/LAI relationships for the overstory and the total canopy are similar.

#### 4.4. Limitations and future perspectives

The global CI, FVC, and LAI products and their relationships were evaluated using high-resolution reference data. The high-resolution reference dataset used in this study is mostly located in the Northern Hemisphere (Fig. 1), and some sites are mixed with multiple biome types (Table S1). Although the new GBOV V3.0 datasets are available, they were not used in the present study due to a lack of high-resolution LAI<sub>e</sub> data for CI calculation and validation.

The GBOV will be developed to incorporate additional NEON sites and other environmental monitoring networks (TERN and ICOS) to improve its global representativeness (Brown et al., 2020). Meanwhile, individual validation campaigns will also expand the spatial extent of the validation dataset in particular ecosystems (Fang et al., 2019b; Song et al., 2021). With an increasing number of sites, it is preferable to select homogeneous sites for further validation to obtain high-quality reference data. For heterogeneous sites, it is necessary to evaluate the spatial representativeness of field measurements using grading criteria (Xu et al., 2018) and to increase the number of samples to adequately represent the spatial variability of different structural variables.

Evaluating the relationships between multiple variables provide more information about the product consistency for the application community, and thus would help promote the product usage. Continuous and consistent time series reference data provide a good opportunity to evaluate the relationships between the biophysical products at representative sites.

SVC is a special case of the concept of *unit leaf area* applied to FVC. The concept of unit leaf area indicates the leaf coverage efficiency and can be extended to the FAPAR, gross primary production, evapotranspiration, and rainfall interception. The seasonal SVC variation and the differences between the different biomes are associated with biological meanings (Fig. 8 and Fig. S3). This demonstrates its potential usage in ecological and phenological researches. In addition, SVC may be used as a product consistency indicator on a global scale because SVC is always  $\leq 1$  (Fang et al., 2021).

#### 5. Conclusion

In this study, a comprehensive evaluation of the global CI, FVC, and LAI products and their relationships was conducted using high-resolution reference data. GEOV2 FVC and MODIS LAI show good performance ( $R^2 = 0.83$  and 0.72, RMSE = 0.14 and 0.98, respectively). CAS-CI slightly underestimates the reference value ( $R^2 = 0.10$ , bias = −0.10, RMSE = 0.16) and thus further improvement is needed. FVC and LAI show a strong positive correlation across different scales, whereas the relationship between CI and FVC/LAI is different across different spatial scales, indicating that the CI-FVC/LAI relationships should not be directly applied across different scales. The derived G(0) overestimates the reference value (bias = 0.28) due to the inconsistency between the

CI, FVC, and LAI products. The G(0) estimation could be improved by using the global FVC, LAI, and CI(0) products, e.g., from space-borne LiDAR data.

The overstory CI is close to the overall forest CI, and the overstory G(0) is similar to the overall value for most forest types, except for ENF. The overall CI-FVC/LAI relationship is mainly influenced by the overstory, while the FVC-LAI relationship is affected by both the overstory and understory. The vertical structural characteristics found in this study are significant for field measurement, remote sensing estimation, and modeling studies. Future research can be conducted for other remote sensing products in more representative areas.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Appendix C. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.srs.2022.100066>.

#### Appendix A. A list of selected field measurement sites used in the study

Both GBOV and DIRECT 2.1 field measurement data were used in this study. The biome type for GBOV was determined through the percentages of each land cover type within a 3 km × 3 km area centered on the site (see Table S1). The biome type for DIRECT 2.1 is from the dataset.

**Table A1**

Global GBOV and DIRECT 2.1 field measurement sites used in the study. For the Continuity column, “Y” (“N”) indicates >5 (otherwise) continuous measurements. See Fig. 1 for the biome type acronyms.

Site	Code	Latitude	Longitude	Biome	Year	Project	Continuity
25de Mayo_Alfalfa	25de Mayo_Alfalfa	-37.91	-67.75	CRO	2014	DIRECT 2.1	N
25de Mayo_Shurb	25de Mayo_Shurb	-37.94	-67.79	SHR	2014	DIRECT 2.1	N
AHSPECT-CON	AHSPECT-CON	43.97	0.34	CRO	2015	DIRECT 2.1	N
AHSPECT-CRE	AHSPECT-CRE	43.99	-0.05	CRO	2015	DIRECT 2.1	N
AHSPECT-MTO	AHSPECT-MTO	43.57	1.37	CRO	2015	DIRECT 2.1	N
AHSPECT-PEY	AHSPECT-PEY	43.67	0.22	CRO	2015	DIRECT 2.1	N
AHSPECT-SAV	AHSPECT-SAV	43.82	1.17	CRO	2015	DIRECT 2.1	N
AHSPECT-URG	AHSPECT-URG	43.64	-0.43	CRO	2015	DIRECT 2.1	N
Albufera	Albufera	39.27	-0.32	CRO	2014	DIRECT 2.1	N
Barrax	Barrax	39.07	-2.10	CRO	2004, 2005, 2009, 2010, 2014, 2015	DIRECT 2.1	N
Camerons	Camerons	-32.60	116.25	EBF	2004	DIRECT 2.1	N
Collelongo	Collelongo	41.85	13.59	DBF	2015	DIRECT 2.1	N
Demmin	Demmin	53.89	13.21	CRO	2004	DIRECT 2.1	N
Donga	Donga	9.77	1.78	GRA	2005	DIRECT 2.1	N
Gnangara	Gnangara	-31.53	115.88	DBF	2004	DIRECT 2.1	N
Hailun	Hailun	47.41	126.82	CRO	2016	DIRECT 2.1	Y
Honghe	Honghe	47.65	133.52	CRO	2012, 2019	DIRECT 2.1	Y
LaReina_Cordoba_1	LaReina_Cordoba_1	37.82	-4.86	CRO	2014	DIRECT 2.1	N
LaReina_Cordoba_2	LaReina_Cordoba_2	37.79	-4.83	CRO	2014	DIRECT 2.1	N
Larose	Larose	45.38	-75.22	MF	2003	DIRECT 2.1	N
Liria	Liria	39.75	-0.70	ENF	2017	DIRECT 2.1	N
Moncada	Moncada	39.52	-0.39	CRO	2014, 2017	DIRECT 2.1	N
Muragua-Upper-Tana	Muragua-Upper-Tana	-0.77	36.97	CRO	2016	DIRECT 2.1	N
Plan_De_Dieu	Plan_De_Dieu	44.20	4.95	CRO	2004	DIRECT 2.1	N
Pshenichne	PSH	50.08	30.23	CRO	2013, 2014, 2015	DIRECT 2.1	Y
SanFernando	SanFernando	-34.72	-71.00	CRO	2015	DIRECT 2.1	N
Sonian	Sonian	50.77	4.41	MF	2004	DIRECT 2.1	N
SouthWest_1	SW1	43.55	1.09	CRO	2013	DIRECT 2.1	Y
SouthWest_2	SW2	43.45	1.15	CRO	2013	DIRECT 2.1	Y
Utiel	Utiel	39.58	-1.26	CRO	2006	DIRECT 2.1	N
Wankama	Wankama	13.65	2.64	GRA	2005	DIRECT 2.1	N
Bartlett Experimental Forest	BART	44.06	-71.29	MF	2014-2019	GBOV	Y
Blandy Experimental Farm	BLAN	39.06	-78.07	PAS	2015-2019	GBOV	Y
Central Plains Experimental Range	CPER	40.82	-104.75	GRA	2014-2019	GBOV	Y
Disney Wilderness Preserve	DELA	32.54	-87.80	WET	2016-2019	GBOV	Y
Disney Wilderness Preserve	DSNY	28.13	-81.44	WET	2013-2019	GBOV	Y
Guanica Forest	GUAN	17.97	-66.87	EBF	2015-2019	GBOV	Y
Harvard Forest	HARV	42.54	-72.17	MF	2014-2019	GBOV	Y

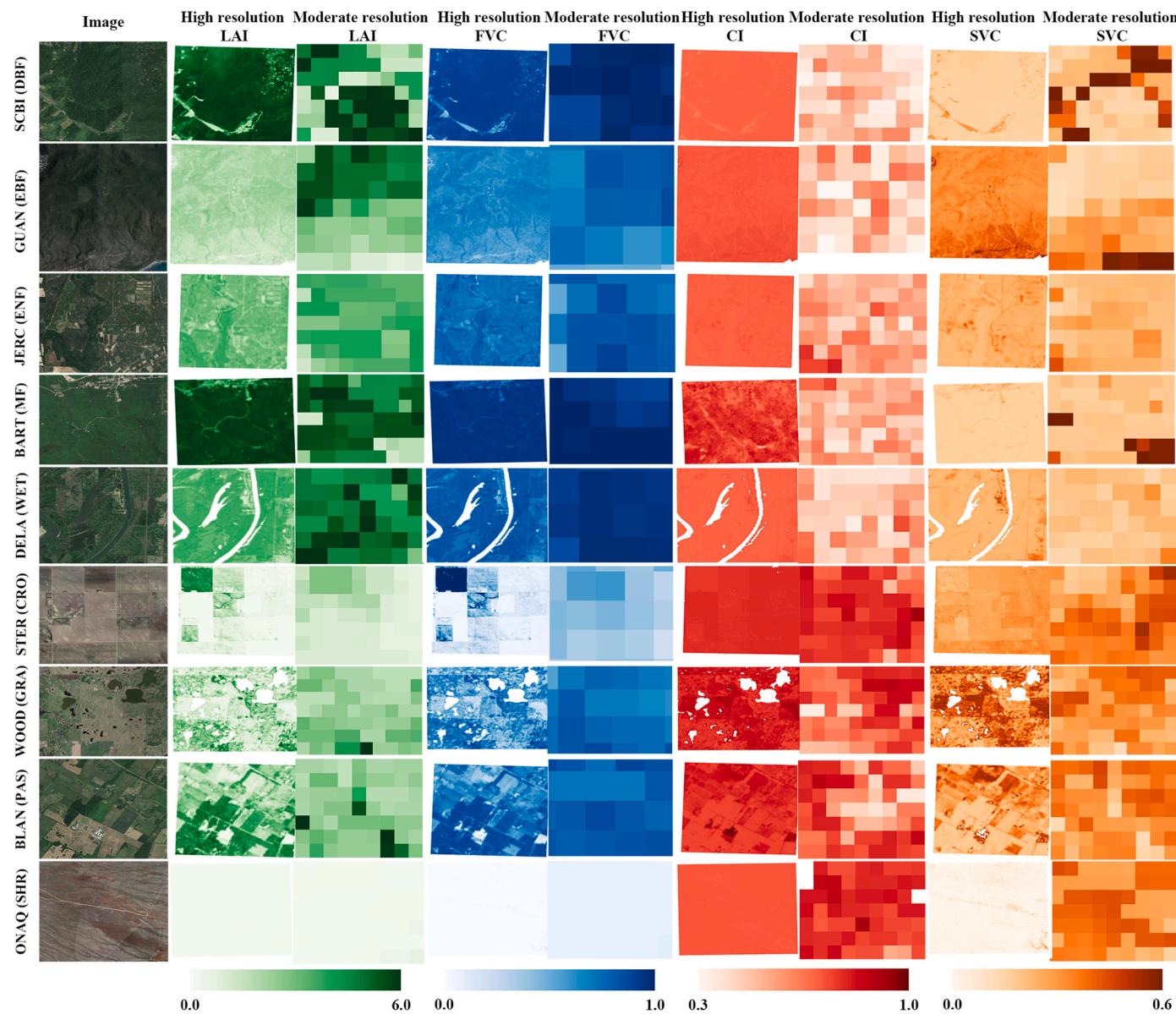
(continued on next page)

**Table A1 (continued)**

Site	Code	Latitude	Longitude	Biome	Year	Project	Continuity
Jones Ecological Research Center	JERC	31.19	-84.47	ENF	2013-2019	GBOV	Y
Jornada	JORN	32.59	-106.84	SHR	2015-2019	GBOV	Y
Lajas Experimental Station	LAJA	18.02	-67.08	PAS	2016-2019	GBOV	Y
Moab	MOAB	38.25	-109.39	SHR	2015-2019	GBOV	Y
Niwot Ridge Mountain Research Station	NIWO	40.05	-105.58	SHR	2015-2019	GBOV	Y
North Sterling	STER	40.46	-103.03	CRO	2014-2019	GBOV	Y
Oak Ridge	ORNL	35.96	-84.28	DBF	2014-2019	GBOV	Y
Onaqui	ONAQ	40.18	-112.45	SHR	2014-2019	GBOV	Y
Ordway-Swisher Biological Station	OSBS	29.68	-82.01	ENF	2013-2019	GBOV	Y
Santa Rita Experimental Range	SRER	31.91	-110.84	SHR	2016-2019	GBOV	Y
Smithsonian Conservation Biology Institute	SCBI	38.89	-78.14	DBF	2014-2019	GBOV	Y
Smithsonian Environmental Research Center	SERC	38.89	-76.56	DBF	2015-2019	GBOV	Y
Steigerwaldt Land Services	STEI	45.51	-89.59	MF	2015-2019	GBOV	Y
Talladega National Forest	TALL	32.95	-87.39	DBF	2014-2019	GBOV	Y
UNDERC	UNDE	46.23	-89.54	WET	2014-2019	GBOV	Y
Woodworth	WOOD	47.13	-99.24	GRA	2014-2019	GBOV	Y

**Appendix B. Spatial pattern of CI, FVC, LAI, and SVC for some typical sites**

The spatial patterns of CI, FVC, LAI, and SVC within a 3 km × 3 km area centered on selected typical sites are displayed in Fig. B1 in the WGS-84 projection. One most representative site for each biome type was selected (Fig. 4 and Table S1).



**Fig. B1.** High- and moderate-resolution LAI, FVC, CI, and SVC maps for selected typical sites in July and August 2016. The images are from the Google satellite base map (<http://mt1.google.com/vt/lyrs=s&x={x}&y={y}&z={z}>) in QGIS software ([www.qgis.org](http://www.qgis.org)). See Table A1 and Fig. 1 for the site and biome type acronyms, respectively.

## References

- Al-Kaisi, M., Brun, L.J., Enz, J.W., 1989. Transpiration and evapotranspiration from maize as related to leaf area index. Agric. For. Meteorol. 48, 111–116.
- Arthur Sampson, D., Smith, F.W., 1993. Influence of canopy architecture on light penetration in lodgepole pine (*Pinus contorta* var. *latifolia*) forests. Agric. For. Meteorol. 64, 63–79.
- Baret, F., Weiss, M., Allard, D., Garrigues, S., Leroy, M., Jeanjean, H., Fernandes, R., Myneni, R., Privette, J., Morissette, J., 2005. VALERI: A Network of Sites and a Methodology for the Validation of Medium Spatial Resolution land Satellite Products [EB/OL]. <https://hal.archives-ouvertes.fr/hal-03221068/>.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujeau, J.L., Leroy, M., 2007. LAI, FAPAR and fCover CYCLOPS global products derived from VEGETATION. Rem. Sens. Environ. 110, 275–286.
- Baret, F., Morissette, J.T., Fernandes, R.A., Champeaux, J.L., Myneni, R.B., Chen, J., Plummer, S., Weiss, M., Bacour, C., Garrigues, S., Nickeson, J.E., 2006. Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: proposition of the CEOS-BELMANIP. IEEE Trans. Geosci. Rem. Sens. 44, 1794–1803.
- Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholczyk, P., Smets, B., 2013. GEOV1: LAI and FAPAR essential climate variables and FCover global time series capitalizing over existing products. Part1: Principles of development and production. Rem. Sens. Environ. 137, 299–309.
- Bond-Lamberty, B., Wang, C., Gower, S., Norman, J., 2002. Leaf area dynamics of a boreal black spruce fire chronosequence. Tree Physiol. 22, 993–1001.
- Brown, L.A., Meier, C., Morris, H., Pastor-Guzman, J., Bai, G., Lerebourg, C., Gobron, N., Lanconelli, C., Clerici, M., Dash, J., 2020. Evaluation of global leaf area index and fraction of absorbed photosynthetically active radiation products over North America using Copernicus Ground Based Observations for Validation data. Rem. Sens. Environ. 247, 111935.
- Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., Weiss, M., 2013. GEOV1: LAI, FAPAR essential climate variables and FCover global time series capitalizing over existing products. Part 2: validation and intercomparison with reference products. Rem. Sens. Environ. 137, 310–329.
- Campbell, G., 1990. Derivation of an angle density function for canopies with ellipsoidal leaf angle distributions. Agric. For. Meteorol. 49, 173–176.
- Chen, J.M., 1996. Optically-based methods for measuring seasonal variation of leaf area index in boreal conifer stands. Agric. For. Meteorol. 80, 135–163.
- Chen, J.M., Black, T.A., 1992. Foliage area and architecture of plant canopies from sunfleck size distributions. Agric. For. Meteorol. 60, 249–266.
- Chen, J.M., Blanken, P.D., Black, T.A., Guilbeault, M., Chen, S., 1997. Radiation regime and canopy architecture in a boreal aspen forest. Agric. For. Meteorol. 86, 107–125.

- Chen, J.M., Cihlar, J., 1995. Plant canopy gap-size analysis theory for improving optical measurements of leaf-area index. *Appl. Opt.* 34, 6211–6222.
- Chen, J.M., Ju, W., Ciais, P., Viovy, N., Liu, R., Liu, Y., Lu, X., 2019. Vegetation structural change since 1981 significantly enhanced the terrestrial carbon sink. *Nat. Commun.* 10, 4259.
- Chen, J.M., Menges, C.H., Leblanc, S.G., 2005. Global mapping of foliage clumping index using multi-angular satellite data. *Rem. Sens. Environ.* 97, 447–457.
- Chianucci, F., Cutini, A., Corona, P., Puletti, N., 2014a. Estimation of leaf area index in understory deciduous trees using digital photography. *Agric. For. Meteorol.* 198–199, 259–264.
- Chianucci, F., Puletti, N., Venturi, E., Cutini, A., Chiavetta, U., 2014b. Photographic assessment of overstory and understory leaf area index in beech forests under different management regimes in Central Italy. *For. Stud.* 61, 27–34.
- Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., Justice, C., 2018. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Rem. Sens. Environ.* 219, 145–161.
- Cohen, W.B., Justice, C.O., 1999. Validating MODIS terrestrial ecology products: linking in situ and satellite measurements. *Rem. Sens. Environ.* 70, 1–3.
- Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurt, G., Kellner, J., Luthcke, S., Armstrong, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marsalis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the earth's forests and topography. *Science of Remote Sensing* 1, 100002.
- EO LAB, 2019. Scientific Quality Evaluation of LAI/FAPAR/FCOVER Collection 1km Version 1 and Version 2 Products. [https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1\\_SQE2019\\_LAI1km-V1%26V2\\_11.00.pdf](https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1_SQE2019_LAI1km-V1%26V2_11.00.pdf).
- Fang, H., 2020. Retrieval of forest vertical leaf area index and clumping index through field measurement and remote sensing techniques: a review. *Chin. Sci. Bull.* 66, 3141–3153. <https://doi.org/10.1360/TB-2020-1057>.
- Fang, H., 2021a. Canopy clumping index (CI): a review of methods, characteristics, and applications. *Agric. For. Meteorol.* 303, 108374.
- Fang, H., 2021b. Vegetation structural field measurement data for Northeastern China Crops (NECC). In: PANGAEA. <https://doi.org/10.1594/PANGAEA.939444>.
- Fang, H., Baret, F., Plummer, S., Schaepman-Strub, G., 2019a. An overview of global leaf area index (LAI): methods, products, validation, and applications. *Rev. Geophys.* 57 (3), 739–799.
- Fang, H., Jiang, C., Li, W., Wei, S., Baret, F., Chen, J.M., Garcia-Haro, J., Liang, S., Liu, R., Myneni, R.B., 2014a. Characterization and intercomparison of global moderate resolution leaf area index (LAI) products: analysis of climatologies and theoretical uncertainties. *J. Geophys. Res. Biogeosci.* 118, 529–548. <https://doi.org/10.1002/jgrc.20051>.
- Fang, H., Li, S., Zhang, Y., Wei, S., Wang, Y., 2021. New insights of global vegetation structural properties through an analysis of canopy clumping index, fractional vegetation cover, and leaf area index. *Science of Remote Sensing* 4, 100027. <https://doi.org/10.1016/j.srs.2021.100027>.
- Fang, H., Li, W., Wei, S., Jiang, C., 2014b. Seasonal variation of leaf area index (LAI) over paddy rice fields in NE China: intercomparison of destructive sampling, LAI-2200, digital hemispherical photography (DHP), and AccuPAR methods. *Agric. For. Meteorol.* 198–199, 126–141. <https://doi.org/10.1016/j.agrformet.2014.08.005>.
- Fang, H., Ye, Y., Liu, W., Wei, S., Li, M., 2018. Continuous estimation of canopy leaf area index (LAI) and clumping index over broadleaf crop fields: an investigation of the PASTIS-57 instrument and smartphone applications. *Agric. For. Meteorol.* 253–254, 48–61. <https://doi.org/10.1016/j.agrformet.2018.02.003>.
- Fang, H., Zhang, Y., Wei, S., Li, W., Ye, Y., Sun, T., Liu, W., 2019b. Validation of global moderate resolution leaf area index (LAI) products over croplands in northeastern China. *Rem. Sens. Environ.* 233, 111377.
- Fernandes, R., Plummer, S., Nightingale, J., Baret, F., Camacho, F., Fang, H., Garrigues, S., Gobron, N., Lang, M., Lacaze, R., LeBlanc, S., Meroni, M., Martinez, B., Nilson, T., Pinty, B., Pisek, J., Sonnentag, O., Verger, A., Welles, J., Weiss, M., Widlowski, J.L., 2014. Global leaf area index product validation good practices. In: Land Product Validation Subgroup (WGC/C/CEOS). 10.5067/doc/ceoswgc/lpv/lai.002.
- Fuster, B., Sánchez-Zapero, J., Camacho, F., García-Santos, V., Verger, A., Lacaze, R., Weiss, M., Baret, F., Smets, B., 2020. Quality assessment of PROBA-V LAI, FAPAR and fCOVER collection 300 m products of Copernicus global land service. *Rem. Sens.* 12 (6), 1017.
- GCOS, 2016. The global observing system for climate: implementation needs (GCOS-200). In: World Meteorological Organization.
- Gielen, B., Acosta, M., Altimir, N., Buchmann, N., Cescatti, A., Ceschia, E., Fleck, S., Hörtnagl, L., Klumpp, K., Kolari, P., Lohila, A., Loustau, D., Marañón-Jiménez, S., Manise, T., Matteucci, G., Merbold, L., Metzger, C., Moureaux, C., Montagnani, L., Nilsson, M.B., Osborne, B., Papale, D., Pavelka, M., Saunders, M., Simioni, G., Soudani, K., Sonnentag, O., Tallec, T., Tuittila, E.-S., Peichl, M., Pokorný, R., Vincke, C., Wohlfahrt, G., 2018. Ancillary vegetation measurements at ICOS ecosystem stations. *Int. Agrophys.* 32, 645–664.
- Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. *Rem. Sens. Environ.* 80, 76–87.
- Gitelson, A.A., Peng, Y., Arkebauer, T.J., Schepers, J., 2014. Relationships between gross primary production, green LAI, and canopy chlorophyll content in maize: implications for remote sensing of primary production. *Rem. Sens. Environ.* 144, 65–72.
- He, L., Chen, J.M., Pisek, J., Schaaf, C.B., Strahler, A.H., 2012. Global clumping index map derived from the MODIS BRDF product. *Rem. Sens. Environ.* 119, 118–130.
- Hinojo-Hinojo, C., Goulden, M., 2020. A compilation of canopy leaf inclination angle measurements across plant species and biome types. <https://doi.org/10.7280/D1T97H>.
- Jia, K., Liang, S., Liu, S., Li, Y., Xiao, Z., Yao, Y., Jiang, B., Zhao, X., Wang, X., Xu, S., Cui, J., 2015. Global land surface fractional vegetation cover estimation using general regression neural networks from MODIS surface reflectance. *IEEE Trans. Geosci. Rem. Sens.* 53, 4787–4796.
- Kao, R.H., Gibson, C.M., Gallery, R.E., Meier, C.L., Barnett, D.T., Docherty, K.M., Blevins, K.K., Travers, P.D., Azuaje, E., Springer, Y.P., Thibault, K.M., McKenzie, V. J., Keller, M., Alves, L.F., Hinckley, E.-L.S., Parnell, J., Schimel, D., 2012. NEON terrestrial field observations: designing continental-scale, standardized sampling. *Ecosphere* 3 (12), 1–17. <https://doi.org/10.1890/ES12-00196.1>.
- Karan, M., Liddell, M., Prober, S.M., Arndt, S., Beringer, J., Boer, M., Cleverly, J., Eamus, D., Grace, P., Van Gorsel, E., Hero, J.M., Hutley, L., Macfarlane, C., Metcalfe, D., Meyer, W., Pendall, E., Sebastian, A., Wardlaw, T., 2016. The Australian SuperSite Network: continental, long-term terrestrial ecosystem observatory. *Sci. Total Environ.* 568, 1263–1274.
- Knyazikhin, Y., Martonchik, J.V., Diner, D.J., Myneni, R.B., Verstraete, M., Pinty, B., Gobron, N., 1998a. Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data. *J. Geophys. Res. Atmos.* 103, 32239–32256.
- Knyazikhin, Y., Martonchik, J.V., Myneni, R.B., Diner, D.J., Running, S.W., 1998b. Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *J. Geophys. Res. Atmos.* 103, 32257–32275.
- Lang, A., Yueqin, X., 1986. Estimation of leaf area index from transmission of direct sunlight in discontinuous canopies. *Agric. For. Meteorol.* 37, 229–243.
- Lawrence, D.M., Fisher, R.A., Koven, C.D., Oleson, K.W., Swenson, S.C., Bonan, G., Collier, N., Ghimire, B., Van Kampenhout, L., Kennedy, D., 2019. The Community Land Model version 5: description of new features, benchmarking, and impact of forcing uncertainty. *J. Adv. Model. Earth Syst.* 11, 4245–4287.
- Leblanc, S.G., Chen, J.M., Fernandes, R., Deering, D.W., Conley, A., 2005. Methodology comparison for canopy structure parameters extraction from digital hemispherical photography in boreal forests. *Agric. For. Meteorol.* 129, 187–207.
- Liu, D., Jia, K., Wei, X., Xia, M., Zhang, X., Yao, Y., Zhang, X., Wang, B., 2019. Spatiotemporal comparison and validation of three global-scale fractional vegetation cover products. *Rem. Sens.* 11, 2524.
- Liu, Y., Liu, R., Pisek, J., Chen, J.M., 2017. Separating overstory and understory leaf area indices for global needleleaf and deciduous broadleaf forests by fusion of MODIS and MISR data. *Biogeosciences* 14, 1093–1110.
- Meier, C., Everhart, J., Jones, K., 2018. TOS Protocol and Procedure: Measurement of Leaf Area Index. <https://data.neonscience.org/api/v0/documents/NEON.DOC.014039vK>.
- Meyer, G.E., Neto, J.C., 2008. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* 63, 282–293.
- Miller, J., 1967. A formula for average foliage density. *Aust. J. Bot.* 15, 141–144.
- Morissette, J.T., Baret, F., Privette, J.L., Myneni, R.B., Nickeson, J.E., Garrigues, S., Shabanov, N.V., Weiss, M., Fernandes, R.A., Leblanc, S.G., Kalacska, M., Sanchez-Azofeifa, G.A., Chubey, M., Rivard, B., Stenberg, P., Rautiainen, M., Voipio, P., Manninen, T., Pilant, A.N., Lewis, T.E., Iáñez, J.S., Colombo, R., Meroni, M., Busetto, L., Cohen, W.B., Turner, D.P., Warner, E.D., Petersen, G.W., Seufert, G., Cook, R., 2006. Validation of global moderate-resolution LAI products: a framework proposed within the CEOS land product validation subgroup. *IEEE Trans. Geosci. Rem. Sens.* 44, 1804–1817.
- Myneni, R., Knyazikhin, Y., Park, T., 2015. In: DAAC, N.E.L.P. (Ed.), CD15A2H MODIS/Terra+Aquatic Leaf Area Index/FPAR 8-day L4 Global 500m SIN Grid V006. <https://lpdaac.usgs.gov/products/mcd15a2hv006/>.
- Myneni, R.B., 1997. Estimation of global leaf area index and absorbed par using radiative transfer models. *IEEE Trans. Geosci. Rem. Sens.* 35, 1380–1393.
- Nikolov, N., Zeller, K., 2006. Efficient retrieval of vegetation leaf area index and canopy clumping factor from satellite data to support pollutant deposition assessments. *Environ. Pollut.* 141, 539–549.
- Nilson, T., 1971. A theoretical analysis of the frequency of gaps in plant stands. *Agric. Meteorol.* 8, 25–38.
- Pinty, B., Lavergne, T., Dickinson, R.E., Widlowski, J.L., Gobron, N., Verstraete, M.M., 2006. Simplifying the interaction of land surfaces with radiation for relating remote sensing products to climate models. *J. Geophys. Res.* 111 (D2).
- Pisek, J., Adamson, K., 2020. Dataset of leaf inclination angles for 71 different Eucalyptus species. *Data Brief* 33, 106391.
- Pisek, J., Govind, A., Arndt, S.K., Hocking, D., Wardlaw, T.J., Fang, H., Matteucci, G., Longdoz, B., 2015. Intercomparison of clumping index estimates from POLDER, MODIS, and MISR satellite data over reference sites. *ISPRS J. Photogrammetry Remote Sens.* 101, 47–56.
- Ross, J., 1981. The Radiation Regime and Architecture of Plant Stands. Springer Science & Business Media.
- Roy, D.P., Zhang, H.K., Ju, J., Gomez-Dans, J.L., Lewis, P.E., Schaaf, C.B., Sun, Q., Li, J., Huang, H., Kovalsky, V., 2016. A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Rem. Sens. Environ.* 176, 255–271.
- Ryu, Y., Nilson, T., Kobayashi, H., Sonnentag, O., Law, B.E., Baldocchi, D.D., 2010. On the correct estimation of effective leaf area index: does it reveal information on clumping effects? *Agric. For. Meteorol.* 150, 463–472.
- Ryu, Y., Verfaillie, J., Macfarlane, C., Kobayashi, H., Sonnentag, O., Vargas, R., Ma, S., Baldocchi, D.D., 2012. Continuous observation of tree leaf area index at ecosystem scale using upward-pointing digital cameras. *Rem. Sens. Environ.* 126, 116–125.
- Sellers, P.J., 1985. Canopy reflectance, photosynthesis and transpiration. *Int. J. Rem. Sens.* 6, 1335–1372.
- Song, B., Liu, L., Zhao, J., Chen, X., Gao, Y., Zhang, H., Zhang, X., 2021. Validation of four coarse-resolution leaf area index products over croplands in China using field measurements. *IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens.* 14, 9372–9382.

- Steltzer, H., Welker, J.M., 2006. Modeling the effect of photosynthetic vegetation properties on the NDVI-LAI relationship. *Ecology* 87, 2765–2772.
- Tang, H., Armston, J., Hancock, S., Marselis, S., Goetz, S., Dubayah, R., 2019. Characterizing global forest canopy cover distribution using spaceborne lidar. *Rem. Sens. Environ.* 231, 111262.
- Verger, A., Baret, F., Weiss, M., 2014. Near real-time vegetation monitoring at global scale. *IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens.* 7, 3473–3481.
- Wang, Y., Fang, H., 2020. Estimation of LAI with the LiDAR technology: a review. *Rem. Sens.* 12 (20), 3457.
- Warren Wilson, J., 1963. Estimation of foliage denseness and foliage angle by inclined point quadrats. *Aust. J. Bot.* 11, 95–105.
- Wei, S., Fang, H., 2016. Estimation of canopy clumping index from MISR and MODIS sensors using the normalized difference hotspot and darkspot (NDHD) method: the influence of BRDF models and solar zenith angle. *Rem. Sens. Environ.* 187, 476–491.
- Wei, S., Fang, H., Schaaf, C.B., He, L., Chen, J.M., 2019. Global 500 m clumping index product derived from MODIS BRDF data (2001–2017). *Rem. Sens. Environ.* 232, 111296.
- Weiss, M., 1991. VALERI: Deriving Leaf Area Index and Average Leaf Inclination Angle from LAI2000 Measurements, pp. 1–4.
- Weiss, M., Baret, F., 2017. CAN-EYE V6.4.91 User Manual. <https://www6.paca.inrae.fr/can-eye/Documentation/Documentation>.
- Woodgate, W., Disney, M., Armston, J.D., Jones, S.D., Suarez, L., Hill, M.J., Wilkes, P., Soto-Berelov, M., Haywood, A., Mellor, A., 2015. An improved theoretical model of canopy gap probability for Leaf Area Index estimation in woody ecosystems. *For. Ecol. Manag.* 358, 303–320.
- Wright, I.J., Leishman, M.R., Read, C., Westoby, M., 2006. Gradients of light availability and leaf traits with leaf age and canopy position in 28 Australian shrubs and trees. *Funct. Plant Biol.* 33, 407–419.
- Xu, B., Li, J., Park, T., Liu, Q., Zeng, Y., Yin, G., Zhao, J., Fan, W., Yang, L., Knyazikhin, Y., Myneni, R.B., 2018. An integrated method for validating long-term leaf area index products using global networks of site-based measurements. *Rem. Sens. Environ.* 209, 134–151.
- Yan, K., Park, T., Chen, C., Xu, B., Song, W., Yang, B., Zeng, Y., Liu, Z., Yan, G., Knyazikhin, Y., Myneni, R.B., 2018. Generating global products of LAI and FPAR from SNPP-VIIRS data: theoretical background and implementation. *IEEE Trans. Geosci. Rem. Sens.* 56, 2119–2137.
- Yan, K., Park, T., Yan, G., Chen, C., Yang, B., Liu, Z., Nemani, R., Knyazikhin, Y., Myneni, R., 2016a. Evaluation of MODIS LAI/FPAR product collection 6. Part 1: consistency and improvements. *Rem. Sens.* 8 (5), 359.
- Yan, K., Park, T., Yan, G., Liu, Z., Yang, B., Chen, C., Nemani, R., Knyazikhin, Y., Myneni, R., 2016b. Evaluation of MODIS LAI/FPAR product collection 6. Part 2: validation and intercomparison. *Rem. Sens.* 8.
- Zeng, Z., Piao, S., Li, L.Z.X., Wang, T., Ciais, P., Lian, X., Yang, Y., Mao, J., Shi, X., Myneni, R.B., 2018. Impact of Earth greening on the terrestrial water cycle. *J. Clim.* 31, 2633–2650.
- Zhang, Y., Fang, H., Wang, Y., Li, S., 2021. Variation of intra-daily instantaneous FAPAR estimated from the geostationary Himawari-8 AHI data. *Agric. For. Meteorol.* 307, 108535.
- Zhao, F., Strahler, A.H., Schaaf, C.L., Yao, T., Yang, X., Wang, Z., Schull, M.A., Román, M.O., Woodcock, C.E., Olofsson, P., Ni-Meister, W., Jupp, D.L.B., Lovell, J.L., Culvenor, D.S., Newham, G.J., 2012. Measuring gap fraction, element clumping index and LAI in Sierra Forest stands using a full-waveform ground-based lidar. *Rem. Sens. Environ.* 125, 73–79.
- Zhao, J., Li, J., Liu, Q., Xu, B., Yu, W., Lin, S., Hu, Z., 2020. Estimating fractional vegetation cover from leaf area index and clumping index based on the gap probability theory. *Int. J. Appl. Earth Obs. Geoinf.* 90, 102–112.
- Zhu, Z., Piao, S., Myneni, R.B., Huang, M., Zeng, Z., Canadell, J.G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., Cao, C., Cheng, L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y., Liu, R., Mao, J., Pan, Y., Peng, S., Peñuelas, J., Poulter, B., Pugh, T.A.M., Stocker, B.D., Viovy, N., Wang, X., Wang, Y., Xiao, Z., Yang, H., Zaehle, S., Zeng, N., 2016. Greening of the Earth and its drivers. *Nat. Clim. Change* 6, 791–795.