



An integrated method for validating long-term leaf area index products using global networks of site-based measurements



Baodong Xu^{a,b,d}, Jing Li^{a,c,*}, Taejin Park^{b,**}, Qinhuo Liu^{a,c,*}, Yelu Zeng^{a,e}, Gaofei Yin^f, Jing Zhao^a, Weiliang Fan^g, Le Yang^a, Yuri Knyazikhin^b, Ranga B. Myneni^b

^a State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

^b Department of Earth and Environment, Boston University, Boston, MA 02215, USA

^c Joint Center for Global Change Studies (JCGCS), Beijing 100875, China

^d University of Chinese Academy of Sciences, Beijing 100049, China

^e Department of Global Ecology, Carnegie Institution for Science, Stanford, CA 94305 USA

^f Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, China

^g School of Environmental and Resources Science, Zhejiang A & F University, Lin'an 311300, China

ARTICLE INFO

Keywords:

Spatial representativeness grading
Spatial upscaling
Time-series ground measurements
Global networks
Leaf Area Index (LAI)
Validation

ABSTRACT

Long-term ground LAI measurements from the global networks of sites (e.g. FLUXNET) have emerged as a promising data source to validate remotely sensed global LAI product time-series. However, the spatial scale-mismatch issue between site and satellite observations hampers the use of such invaluable ground measurements in validation practice. Here, we propose an approach (Grading and Upscaling of Ground Measurements, GUGM) that integrates a spatial representativeness grading criterion and a spatial upscaling strategy to resolve this scale-mismatch issue and maximize the utility of time-series of site-based LAI measurements. The performance of GUGM was carefully evaluated by comparing this method to both benchmark LAI and other widely used conventional approaches. The uncertainty of three global LAI products (i.e. MODIS, GLASS and GEOV1) was also assessed based on the LAI time-series validation dataset derived from GUGM. Considering all the evaluation results together, this study suggests that the proposed GUGM approach can significantly reduce the uncertainty from spatial scale mismatch and increase the size of the available validation dataset. In particular, the proposed approach outperformed other widely used approaches in these two respects. Furthermore, GUGM was successfully implemented to validate global LAI products in various ways with advantaging frequent time-series validation dataset. The validation results of the global LAI products show that GLASS has the lowest uncertainty, followed by GEOV1 and MODIS for the overall biome types. However, MODIS provides more consistent uncertainties across different years than GLASS and GEOV1. We believe that GUGM enables us to better understand the structure of LAI product uncertainties and their evolution across seasonal or annual contexts. In turn, this method can provide fundamental information for further LAI algorithm improvements and the broad application of LAI product time-series.

1. Introduction

Leaf Area Index (LAI), which is defined as one half of the total green leaf area per unit ground surface area (Chen and Black, 1992), has been widely used to characterize the structure and function of vegetation (Garrigues et al., 2008). As the leaf is the primary interface for the exchange of fluxes of energy, mass (e.g. water, nutrients and CO₂) and momentum between the surface and the planetary boundary layer, LAI is identified as a key parameter in most terrestrial ecosystem models (Bonan, 1995; Liu et al., 1997; Richardson et al., 2012; Sellers et al., 1997). Thus, generating accurate, consistent and continuous long-term

global LAI datasets from remote sensing observations has drawn the attention of scientific communities (Myneni et al., 2002; Zhu et al., 2013). Several LAI products based on different combinations of sensors (e.g. MODIS, VEGETATION, MERIS, VIIRS etc.) and algorithms (e.g. using the look-up table generated from radiative transfer models, machine learning etc.) have been developed (Baret et al., 2007; Knyazikhin et al., 1998; Yan et al., 2018) and widely used in a broad range of user communities (e.g. Bi et al., 2015; Samanta et al., 2012; Zhu et al., 2016). Assessing the uncertainties associated with these LAI products through comparisons with independent ground-truth measurements (i.e. direct validation) is pivotal for their proper use in land surface

* Corresponding authors at: State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China.

** Corresponding author at: Department of Earth and Environment, Boston University, Boston, MA 02215, US.

E-mail addresses: lijing01@radi.ac.cn (J. Li), parktj@bu.edu (T. Park), liuqh@radi.ac.cn (Q. Liu).

models and other applications (Fang et al., 2012; Morisette et al., 2006; Yan et al., 2016).

Direct validation is the most common approach to evaluate products to understand the uncertainties associated with input, pre- or post-processing and inversion algorithms. Many regional field campaigns (e.g. VALERI, BigFoot, SAFARI 2000, etc.) have collected and provided invaluable ground LAI measurements covering a wide range of biome types and spatial variabilities (Morisette et al., 2006). In the Committee on Earth Observation Satellites (CEOS) hierarchical four-stage validation approach, current global medium-resolution LAI products (250 m–1 km) are considered to be validated at Stage 2 (“Product accuracy is estimated over a significant set of locations and time periods by comparison with reference in situ or other suitable reference data.”) after a tremendous effort from scientific communities (Camacho et al., 2013; Yan et al., 2016). However, most previous studies were limited to evaluate the temporal performance of the LAI products due to limited resources for collecting time-series of ground LAI via recursive field campaigns (Claverie et al., 2013). This restriction is critical because assessing the temporal performance of these products enables us to better understand the structure of the uncertainties and their evolution across seasonal or annual contexts (Barr et al., 2004; Weiss et al., 2007) and consequently reach the upper validation stage (Stage 3). To improve the temporal assessment capability, the global network of sites (e.g. FLUXNET), which enables to obtain continuous time-series of ground LAI measurements (hereafter, LAI_{site-TS}) using onboard instruments or recursive data collection (Baldocchi et al., 2001), emerges as a promising data source for validation (Xu et al., 2016). However, the spatial scale mismatch restricts the utilization of LAI measurements from these networks as the LAI is conventionally measured within an area of tens of meters around a site. The scale issue usually introduces undesired errors in the validation of remote sensing products because the spatially heterogeneous land surface results in incomparability between observations from sites and satellites (Yang et al., 2006b). Therefore, using LAI_{site-TS} to validate time-series of LAI products remains a challenge that should be addressed for broader application of long-term global LAI products.

Currently, two approaches are available to remedy the scale issue in utilizing LAI_{site-TS}: (1) the “bottom-up” method and (2) the spatial representativeness evaluation method (Fensholt et al., 2004; Morisette et al., 2006). The “bottom-up” method proposed by the CEOS Land Product Validation (LPV) subgroup is designed to link the ground measured LAI to the remotely sensed LAI through a rigorous upscaling procedure (Tan et al., 2005). This method first employs a two-stage sampling strategy: (a) capture the variability across the extent of a site based on multiple elementary sampling units (ESUs), and (b) capture the variability within pixels of the high spatial resolution image (HSI) by repeating measurements within each ESU. Then, a transfer function is established based on the ground LAI and the spectral measurements from the HSI to generate an LAI reference map (hereafter, LAI_{HSI}). Finally, the generated LAI_{HSI} map is spatially aggregated to retrieve scale-matched LAI (hereafter, LAI_{HSI-AGG}) for direct comparison (Morisette et al., 2006). The “bottom-up” approach is effective for both homogeneous and heterogeneous landscapes because it employs a sufficient number (20–100) of ESUs to represent regions with different spatial heterogeneities and obtain a robust transfer function between LAI and spectral characteristics. However, this method is unsuitable for most of sites where LAI measurements were repeated throughout years (e.g. FLUXNET sites) because of poor spatial sampling. Note that the “bottom-up” approach may also be unsuccessful if the HSI is unavailable because of unexpected conditions (e.g. the cloud effect and the temporal mismatch due to the satellite local passing time). The second approach is based on the evaluation of the spatial representativeness of LAI_{site-TS} (Xu et al., 2016). This method determines whether the LAI measurements are spatially representative for the product pixel by quantitatively considering the point-to-pixel comparability and within-pixel heterogeneity. Quantifying the reliability of ground observations

is advantageous because this approach can reduce potential errors from point-to-pixel inconsistency. In particular, time-series validation practice can greatly benefit from this approach because it can consider changing spatial heterogeneity within the product pixel grid over time due to variation in vegetation growth at different growth stages (Ding et al., 2014). However, the stand-alone implementation of this approach yields only a few valid (i.e. high representativeness) LAI_{site-TS} dataset from networks that can accurately represent the product pixels. Consequently, this limit in implementing the stand-alone second approach hinders the ability to assess the temporal performance of the product. Therefore, an additional processing (i.e. upscaling) is required to use less representative measurements to derive the time-series validation dataset.

Here, we propose an integrated approach, namely, the Grading and Upscaling of Ground Measurements (GUGM), which reconciles the *pros* and *cons* of the above two approaches. The GUGM is expected to be more suitable for LAI_{site-TS} than the current methods, i.e. the “bottom-up” method and the spatial representativeness evaluation method, in two respects: (1) reducing the uncertainty of the upscaled LAI dataset, and (2) increasing the size of the available LAI dataset. This paper aims to (a) provide a full description of the GUGM, (b) evaluate the performance of the GUGM compared to that of conventional approaches, and (c) implement the proposed approach on three global LAI products: MODIS, GLASS and GEOV1. The paper is organized as follows. Section 2 describes the framework of the GUGM method. Section 3 introduces the data and detailed methods in this study. Section 4 provides the results and discussion for the evaluation of GUGM and the application of GUGM for the three global LAI products. Finally, Section 5 provides concluding remarks on this study.

2. Framework of the GUGM method

The proposed GUGM method mainly includes two sequential processes, i.e. spatial representativeness grading and spatial upscaling. GUGM first ingests LAI_{site-TS}, the reflectance of the HSI and a land cover map as inputs, and then generates a LAI validation dataset (hereafter, LAI_{site-HSI}), which is directly comparable to LAI products with minimizing potential scale effects. Note that LAI_{site-HSI} is generated by using the combination of LAI_{site-TS} and LAI_{HSI-AGG} at a given spatial resolution. The framework of GUGM is shown in Fig. 1 and a detailed description of each step is provided below.

2.1. Spatial representativeness grading

The method presented by Xu et al. (2016) uses three indicators to evaluate the spatial representativeness of LAI_{site-TS}: Dominant Vegetation Type Percent (*DVTP*), Relative Absolute Error (*RAE*) and Coefficient of Sill (*CS*). These indicators are calculated based on the HSI. The *DVTP* is defined as the percent of the area covered by the vegetation type which was observed at the LAI field site. It indicates whether the site-observed vegetation type is the same as the dominant vegetation type in the product pixel grid. The *RAE* quantifies the point-to-pixel consistency by calculating the absolute difference between the LAI_{site-TS} and LAI_{HSI-AGG}, and then dividing by LAI_{HSI-AGG} in the product pixel grid. The *CS*, defined as the ratio of the square root of the sill value from a fitted variogram function to LAI_{HSI-AGG}, describes the spatial heterogeneity caused by different vegetation densities within a pixel grid. To adequately compute a variogram, we secured sufficient pair samples (e.g. 4160 and 860 for minimum and maximum lag distance within 1-km area, respectively). To grade the spatial representativeness of the measurements, the proper selection of the thresholds for *DVTP*, *RAE*, and *CS* is critical. For the sake of brevity, a detailed description of the threshold selection for the three indicators is not provided here (see Section 2 of Xu et al. (2016) for the details). Based on the established rules, the spatial representativeness of LAI_{site-TS} in the product pixel grid is divided into five levels (Levels 0–4), as shown in Table 1. Level 0

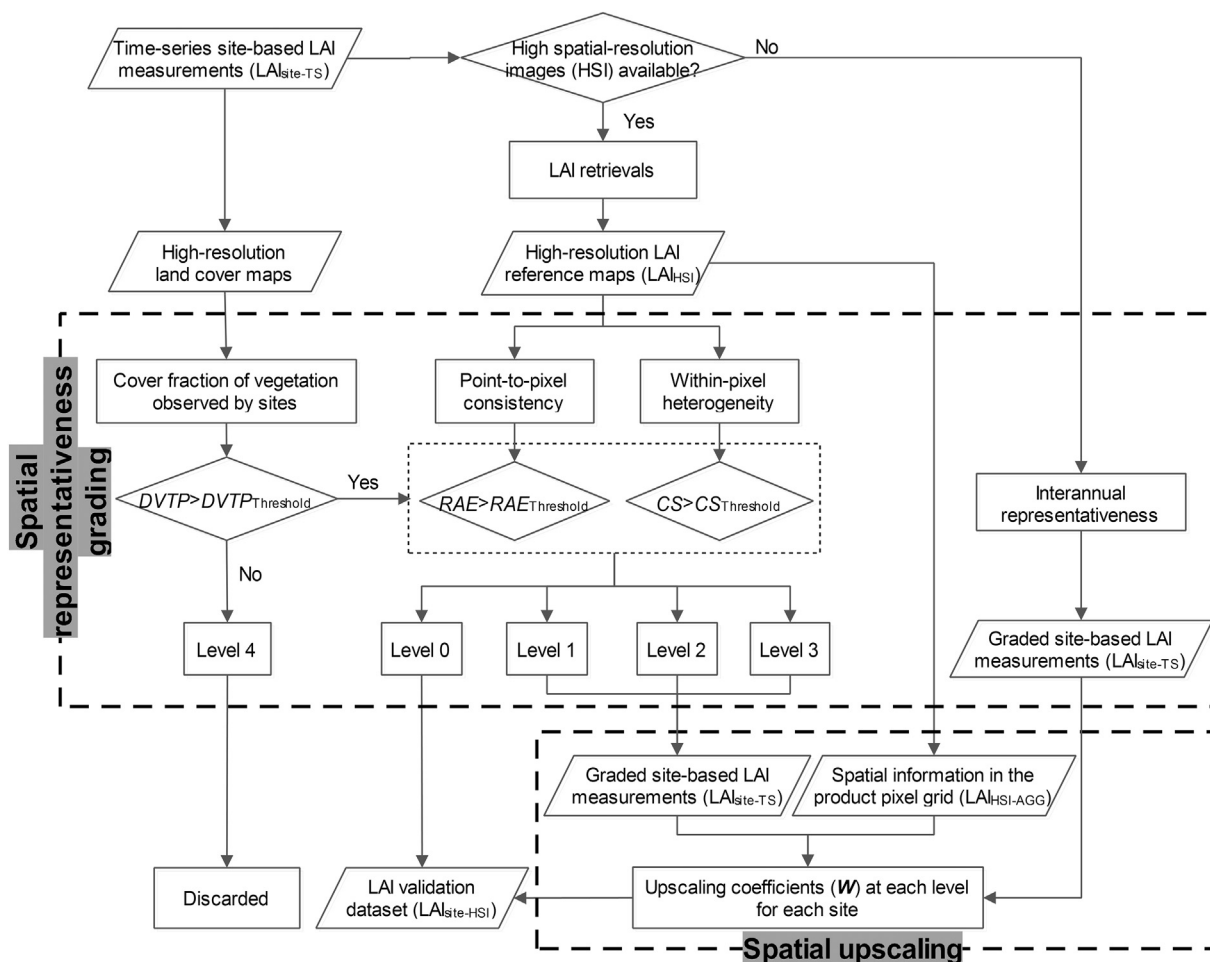


Fig. 1. The framework of the GUGM method.

Table 1

Grading of spatial representativeness for LAI_{site-TS}. The number in the bracket indicates the threshold of each indicator in this study. The threshold determination will be introduced in Section 3.3.

Level	Indicators		
	DVTP > DVTP _{Threshold} (60%)	RAE < RAE _{Threshold} (7%)	CS < CS _{Threshold} (6%)
0	✓	✓	✓
1	✓	✓	×
2	✓	×	✓
3	✓	×	×
4	×	–	–

measurements have good representativeness in the targeted pixel grid and can be used directly to validate products without spatial upscaling. Level 1, Level 2 and Level 3 measurements in the product pixel grid are not sufficiently representative due to their low point-to-pixel consistency or high spatial heterogeneity, which may introduce unexpected errors into the product validation results. Thus, following a rigorous spatial upscaling process is required for these three levels. Note that the stratification of those ground measurements through spatial upscaling into three levels (i.e. Levels 1–3) is beneficial to reduce the uncertainty occurring during the upscaling process (see an example in Supplementary material, Fig. S1) because mixing measurements with different degrees of representativeness reduces the quality of the LAI that are upscaled from more representative measurements. In detail, the uncertainty of upscaled LAI_{site-TS} would be increased by the over-corrected effect for high representativeness level of LAI_{site-TS} and under-corrected

effect for low representativeness level of LAI_{site-TS} if we used only one upscaling coefficient for all the LAI_{site-TS} without the additional grading process. Measurements classified as Level 4 cannot be used for product validation because of the inconsistency between the vegetation type observed by the site and the product pixel, indicating that LAI_{site-TS} is unable to represent the LAI value of the dominant vegetation type within the product pixel grid.

The grading method inevitably requires HSI that covers each LAI_{site-TS} to enable the calculation of the three indicators. However, some LAI measurements lack the corresponding clear-sky HSI for adjacent dates because of the long revisit cycles of HSI satellites or other negative effects (e.g. clouds). For these cases, the proposed GUGM framework also includes a back-up rule for spatial representativeness grading when the adjacent HSI is unavailable. First, it should be noted that the ground sites are designed for long-term observations, thus the vegetation types are generally consistent on the same observational dates in different years to preserve the comparability of measurements. Here, we assumed that the relative spatial variability on the surface surrounding a site does not change among adjacent observational dates (< 4 days) in consecutive years (Chen et al., 2012; Kim et al., 2006). Then, the level of spatial representativeness also remains stable if the LAI_{site-TS} was measured for the same vegetation type (Xu et al., 2016). Therefore, even when HSI are unavailable, the LAI_{site-TS} can be graded based on similar spatial representativeness on adjacent dates in different years. For example, suppose a site that has monthly LAI_{site-TS} from 2005 to 2006. Unfortunately, the HSI for LAI_{site-TS} on 9 July 2006 is unavailable due to the cloudy condition. In this case, the back-up rule in GUGM applies the representativeness level that is extracted from 11 July 2005 (< 4 days in consecutive years) to this LAI_{site-TS} based on

forementioned our presumption. Note that potential effect of irregular crop rotation on the back-up rule is minimal in our case because 94.7% of observations based on the back-up rule across cropland sites was assured as regular cropping practices during our observation period. For the spatial representativeness level of $LAI_{site-TS}$ from the back-up rule, Section 2.2 shows the details regarding the use of the spatial upscaling process in this case.

2.2. Spatial upscaling

The strategy to spatially upscale $LAI_{site-TS}$ from networks was proposed by Qin et al. (2013). The theoretical formulation is shown as Eq. (1)

$$LAI_{site-HSI}(d) = W^T LAI_m(d) \quad \text{and} \quad LAI_m(d) = [1, LAI_{site-TS}(d)]^T \quad (1)$$

where $LAI_{site-TS}(d)$ and $LAI_{site-HSI}(d)$ represent the time-series LAI ground measurement and the upscaled LAI in the product pixel grid on date d , respectively. W is the vector of combination coefficients, which determines the upscaling results of the LAI measurements ($LAI_{site-HSI}(d)$) on all the dates. To calculate W , a cost function is established in Eq. (2) by combining all the dates (M) at one site.

$$J = \text{Minimize} \left\{ \sum_{d=1}^M [LAI_{site-HSI}(d) - W^T LAI_m(d)]^2 \right\} \quad (2)$$

Then, W can be derived by minimizing the cost function using the ordinary least-squares (OLS) algorithm. By setting the derivative of J with respect to W equal to zero, W is calculated using Eq. (3).

$$W = [LAI_m(d)^T LAI_m(d)]^{-1} LAI_m(d)^T LAI_{site-HSI}(d) \quad (3)$$

where $LAI_{site-HSI}(d)$ is the solution of the upscaling method that cannot be directly obtained. To solve for W , the representative LAI value ($LAI_{HSI-AGG}(d)$) is estimated from LAI_{HSI} in the product pixel grid to replace $LAI_{site-HSI}(d)$, as shown in Eq. (4).

$$LAI_{HSI-AGG}(d) = \frac{\sum_{i=1}^n LAI_{HSI}(d)}{N} \quad (4)$$

where $LAI_{HSI}(d)$ denotes the LAI value of the i th vegetated pixel in the HSI, N is the total number of high-resolution pixels in the product pixel grid, and n is the number of vegetated pixels in the high-resolution land cover map. According to Eq. (2) and (3), if $LAI_{HSI-AGG}(d)$ is correct or includes only random errors and no biases, the value of W can be calculated accurately. However, some noises, such as errors from atmospheric correction or LAI retrieval algorithms, are hard to avoid in the LAI estimation for HSI. Thus, the value of W calculated using Eq. (2) and the OLS method will easily be overfitted due to those noise-influenced LAI. To reduce these overfitting effects, a regularization term is added to the cost function (Qin et al., 2013; Tarantola, 2005). Then, the cost function can be re-written as follows:

$$J = \text{Minimize} \left\{ \sum_{d=1}^M [LAI_{HSI-AGG}(d) - W^T LAI_m(d)] \sigma^{-2} [LAI_{HSI-AGG}(d) - W^T LAI_m(d)] + \alpha W^T W \right\} \quad (5)$$

where α and σ are the regularization parameter and the standard deviation of $LAI_{HSI-AGG}(d)$, respectively, and both are correlated with W and $LAI_m(t)$. Then, the optimal value of W can be determined by setting the derivative of J equal to zero:

$$W = \sigma^{-2} S [LAI_m(d)]^T LAI_{HSI-AGG}(d) \quad \text{and} \quad S = [\alpha I + \sigma^{-2} (LAI_m(d))^T LAI_m(d)]^{-1} \quad (6)$$

where I is the identity matrix. If α is close or equal to zero, W in Eq. (6) reduces to the OLS result given by Eq. (3). Therefore, the derived value of W in this study also considers the errors from the LAI retrievals. A

Bayesian linear regression approach was implemented to compute α and σ in Eq. (6) based on the iteration strategy proposed by Chen and Martin (2009) and Qin et al. (2013). Finally, the upscaling coefficient W calculated from Eq. (6) for each level at each site was applied to Eq. (1) to upscale the corresponding $LAI_{site-TS}$. The spatial upscaling method in this study is advantageous because its upscaling coefficient W for each level at each site was already derived based on $LAI_{site-TS}$ and $LAI_{HSI-AGG}$, which can also be used to upscale $LAI_{site-TS}$ at this site without the corresponding HSI at this level to increase the size of the available validation dataset ($LAI_{site-HSI}$).

3. Data and methods

3.1. Global site-based LAI measurements

FLUXNET (<http://fluxnet.ornl.gov/>, July 2014) and CERN (Chinese Ecosystem Research Network, <http://www.cerndata.ac.cn/>, September 2014) are the main observation networks of $LAI_{site-TS}$. FLUXNET includes seven regional networks maintaining > 500 observation sites worldwide, and CERN has collected biophysical variables from 40 sites across inland China. Since these networks are not particularly designed for collecting LAI, most sites have only a single or few LAI measurements, which is not sufficient to validate the time-series of products. Only the 38 sites collecting LAI time-series measurements over multiple years were further selected for the analysis in this study. The selected sites marked by green dots in Fig. 2 are mainly located in the USA, Canada and China, among which are 17 forest sites, 18 cropland sites, and 3 grassland sites (Supplementary material, Table S1). A total of 1693 ground measurements from 2001 to 2011 were acquired across all the sites, and the number of measurements during a year ranged from 90 to 235, which suggests that the temporal trajectory of products can be adequately evaluated with these observation networks.

At each network-site, a tens of meters quadrat was first selected, and each quadrat was divided into sub-quadrats based on the network protocol (Barr et al., 2004; Fu et al., 2010; Gonsamo and Chen, 2014). Then, several randomly selected sub-quadrats were used to obtain LAI. There are two conventional ways to measure ground LAI quantity at network sites: direct (or semi-direct) and indirect methods. The direct method involves the harvesting or collection of sampled foliar elements and scales to quantify the true LAI of a given area. The indirect method infers LAI from measurements of the transmission of radiation through the canopy using the radiative transfer theory (Ross, 1989) and employs optical instruments (e.g. LAI 2200, AccuPAR, digital hemispherical photography (DHP), etc.) to measure the transmittance but usually provides effective LAI, which are not taking account of clumping. For crops and grasses, we did not consider the clumping effect in these biomes because their typical clumping index is higher than 0.9 (Tang et al., 2007). For forest biome case, by ignoring the clumping of forest foliage, the effective LAI often underestimates the true LAI. To retrieve and upscale true LAI, the clumping effect in forest biomes was corrected by considering the clumping index in the LAI retrieval model (see details in Section 3.3).

3.2. High-resolution land cover and reflectance

The 30-m spatial resolution global land cover map (GlobeLand30, <http://www.globallandcover.com/>) produced by the National Geomatics Center of China (Chen et al., 2014) was used to compute the DVTP indicator as part of the spatial representativeness grading. The classification approach was first based on the integration of pixel- and object-based methods with knowledge (POK-based), and then a knowledge-based interactive verification procedure was developed to improve the classification quality (Chen et al., 2015). GlobeLand30 identifies the 10 land cover types, including cropland, forest and grassland, which are comparable with site-based classes. Since two 30-m land cover datasets for the years 2000 and 2010 have been produced,

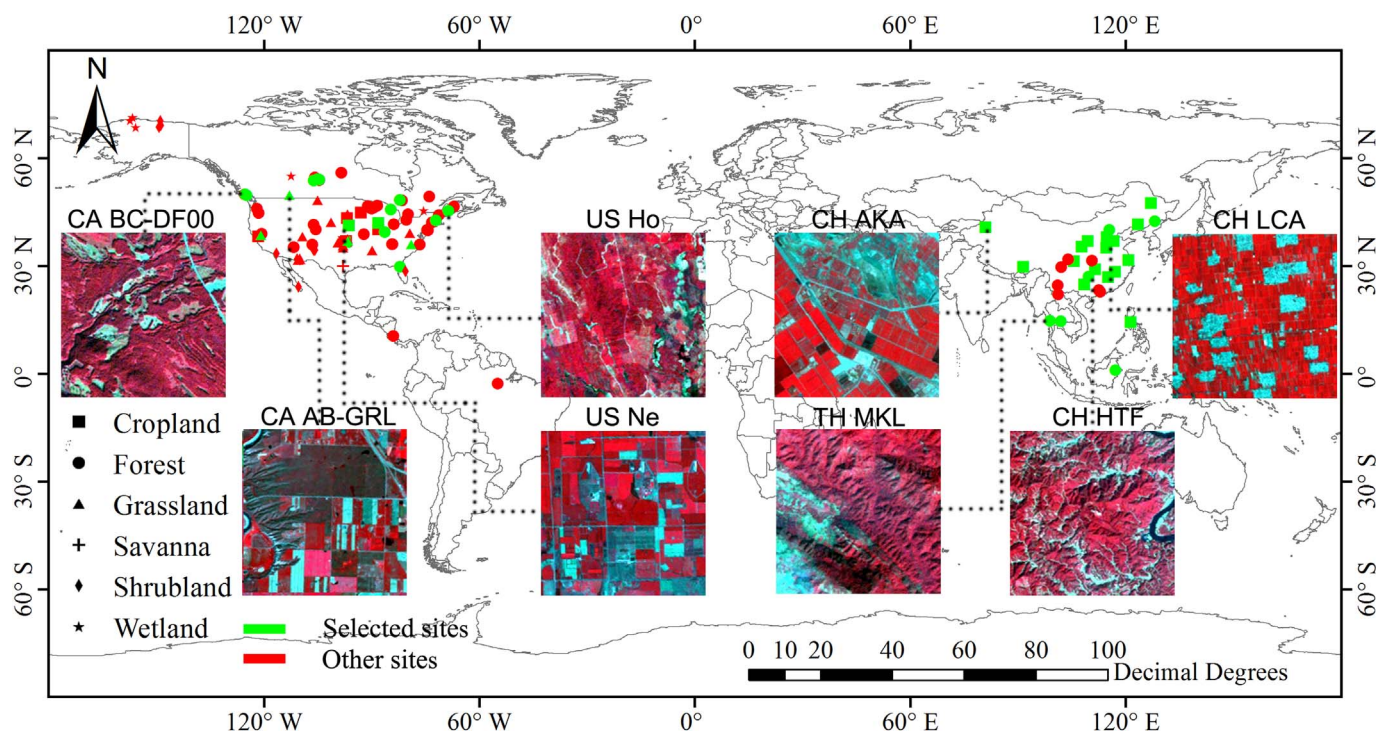


Fig. 2. Spatial distribution of the 124 global sites for which LAI measurements are available. Green dots ($n = 38$) represent the sites with LAI time-series measurements selected for analysis in this study. The color images represent the $5\text{ km} \times 5\text{ km}$ area TM pseudocolor (RGB: near-infrared, red and green bands, respectively.) images with the sites at their center points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

we selected one land cover dataset that is close to the observation date of the $LAI_{\text{site-TS}}$ to calculate the $DVTP$. Previous studies reported that the overall accuracy of Globeland30 products is $> 80\%$ based on a two-rank sampling strategy to represent global conditions, and the accuracies of cropland, forest and grassland are 83%, 84% and 72%, respectively (Chen et al., 2015). In particular, the evaluation result of Globeland30 based on the ground vegetation type from the 38 sites showed that its accuracy (89.5%) was highly reasonable. This reported accuracy imbues our confidence in the high-resolution land cover map and suggests a good reliability of the $DVTP$ indicator.

In GUGM, RAE and CS used to evaluate the spatial representativeness of $LAI_{\text{site-TS}}$ are computed based on the high-resolution LAI maps (i.e. LAI_{HSI}). The LAI_{HSI} maps are generated by a neural network trained using radiative transfer models (RTMs). To generate the LAI_{HSI} maps, all available Landsat TM/ETM+ surface reflectance data acquired on dates close to those of $LAI_{\text{site-TS}}$ were collected from the US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA) On Demand Interface (<https://espa.cr.usgs.gov/>). The Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) applies the atmospheric correction routines of the 6S radiative transfer model to Level-1 Landsat TM/ETM+ data, with inputs of water vapor, ozone, geopotential height, aerosol optical thickness, and digital elevation to generate surface reflectance (Masek et al., 2006). Given the long revisit cycle (16 days) of the Landsat satellites and the cloud effects, only 568 clear TM/ETM+ images, which covered approximately 43% of the available $LAI_{\text{site-TS}}$, were carefully selected in this study. Based on the temporal distribution of the acquired images, in $> 89\%$ of the cases, differences of < 8 days existed between the acquisition dates of the $LAI_{\text{site-TS}}$ and the TM/ETM+ images, and approximately only 10% of the temporal differences were longer than 8 days (≥ 14 days: 1%) (Supplementary material, Fig. S2). Network sites with long temporal differences in the Landsat data are mainly situated in evergreen forest sites. The higher temporal availability (< 14 days: 99%) of clear Landsat TM/ETM+ indicates only a minor difference in the acquisition dates, and we believe the impact of this difference is minimal in this study.

3.3. Generation of the global LAI time-series validation dataset

Based on the Landsat TM/ETM+ reflectance data, a neural network was trained using RTMs and then used to retrieve the LAI (i.e. LAI_{HSI}). This method includes three steps: the generation of training dataset using RTMs, the neural network training and the LAI retrievals using the trained neural network (Claverie et al., 2013). In this study, we used the Scattering from Arbitrarily Inclined Leaves (SAIL) model (Verhoef, 1984) to simulate the canopy reflectance of cropland and grassland, and the Four-Scale model (Chen and Leblanc, 1997) to simulate the canopy reflectance of forest (Yin et al., 2015). As mentioned in Section 3.1, the clumping index for cropland and grassland is relatively large, so the clumping effect was ignored in the SAIL model when simulating the canopy reflectance for cropland and grassland. For forest, the clumping index is an important input parameter in the Four-Scale model to simulate the canopy reflectance, and thus the clumping effect was considered in the LAI retrievals. The input variables of the SAIL and Four-Scale models were determined (Supplementary material, Table S2-S4) according to Yin et al. (2015). Note that the spectral response function (SRF) of the TM/ETM+ sensor in the red and near infra-red (NIR) bands was used to generate the reflectance for each sensor in the training dataset. A back-propagation neural network, with demonstrated efficiency in retrieving LAI (Baret et al., 2013; Verger et al., 2011), was trained using the generated training dataset. The neural network contains three layers: the input layer, the hidden layer and the output layer. The input layer has 4 neurons, representing four input variables: the reflectances of the red and NIR bands, the solar zenith angle and the relative azimuth angle between the sun and the sensor. In this study, based on the recommendation of Mather and Koch (2010), we set 8 neurons in the hidden layer, i.e. twice the number of neurons in the input layer ($= 4$), because this empirically driven combination maximizes the performance of the neural network. The unique output variable was the LAI, which was the single linear neuron of the output layer. The neural network was calibrated based on the methodologies proposed by Baret et al. (2013) and Verger et al. (2011). Note that the

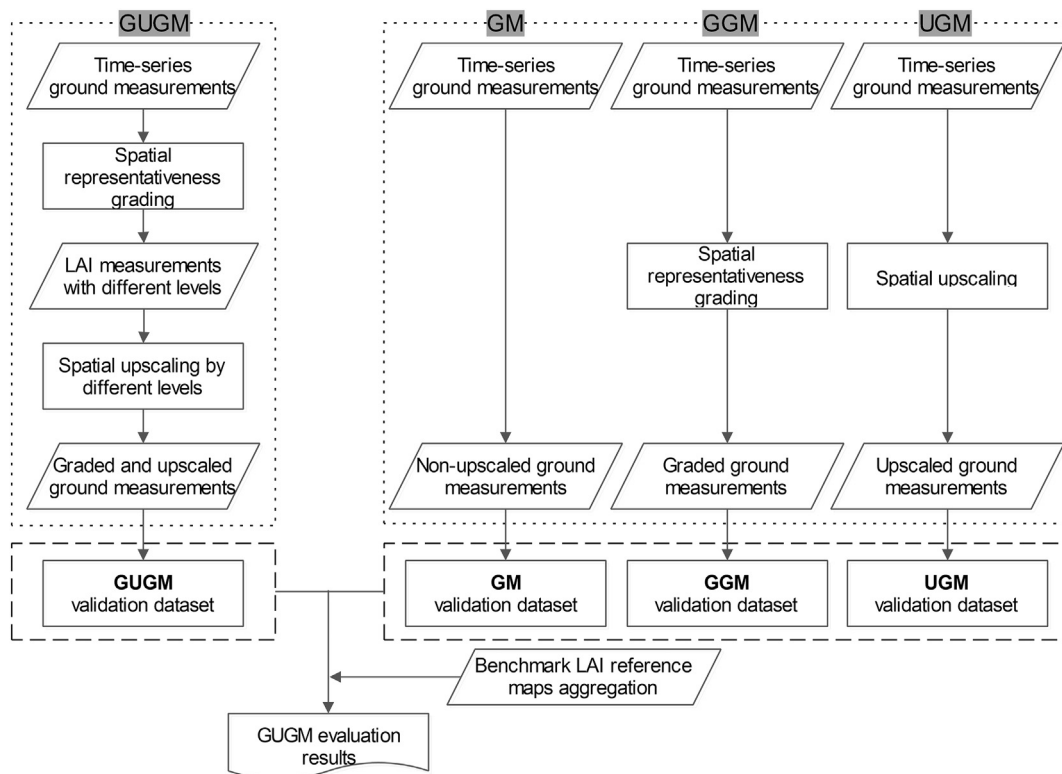


Fig. 3. Comparisons of GUGM with other three methods (GM, GGM and UGM) for processing site-based LAI measurements.

trained and calibrated neural networks were separately built from the vegetation types because the training dataset was not from the same RTM, and then these neural networks were applied to all the available TM/ETM+ datasets to generate LAI_{HSI} maps. It is also critical to confirm that the LAI_{HSI} from the trained and calibrated neural networks is accurate enough. Our additional evaluation of the generated Landsat LAI_{HSI} (Supplementary material, Fig. S3) revealed that the overall bias of the neural network-based LAI_{HSI} was minimal ($= 0.06$ LAI unit), indicating only an unnoticeable overestimation across various LAI ranges. Moreover, the comparison between $LAI_{site-TS}$ and LAI_{HSI} also showed that the performance of the LAI retrievals ($R^2 > 0.75$, bias < 0.3 and RMSE < 0.70) was reasonable for different vegetation types. Acceptable agreement between the Landsat LAI_{HSI} and $LAI_{site-TS}$ confirms that the Landsat LAI_{HSI} maps can be reliably used in the GUGM framework. Although some noise-influenced LAI retrievals occurred, we considered this issue in the spatial upscaling strategy as mentioned in Section 2.2 (Eq. (5)). Finally, we carefully implemented the proposed GUGM method in Section 2 into the $LAI_{site-TS}$ and the produced Landsat LAI_{HSI} (at 30-m resolution) to generate the global LAI time-series validation dataset (at 1-km resolution). Overall assessments of the GUGM-based LAI validation dataset over all available network sites are provided. Thresholds of DVTP, RAE and CS of 60%, 7% and 6% were used in this study, respectively. Further details on the threshold determination for these indicators are reported in Xu et al. (2016).

3.4. Performance evaluation of the GUGM

The performance of the proposed GUGM method should be assessed prior to further implementation. The evaluation first consisted in selecting field campaign sites where the maximal number of HSI reflectance data and corresponding LAI reference benchmark maps are available. Note that given LAI reference benchmark maps (hereafter, LAI_{HSI-BM}) used here were independently generated by building a transfer function between ground LAI and radiometric measurements (i.e. based on the “bottom-up” upscaling approach). The use of

independent benchmark maps representing local conditions precisely enabled us to evaluate both the LAI_{HSI} from the RTMs and neural networks and the final output of the GUGM. Thereafter, the proposed GUGM method was applied to the reflectance data for LAI_{HSI} retrievals. The performance of GUGM can be assessed through the comparison with the LAI benchmark map at the target spatial grid scale (i.e. 1-km).

In this study, we chose three measurement sites representing forest (Järvelja, Estonia, 27° 15' E, 58° 18' N), cropland (Toulouse, France, 1° 9' E, 43° 27' N) and grassland (Kansas, USA, 96° 34' W, 39° 5' N) vegetation types. The Järvelja, Toulouse and Kansas sites are components of the VALERI (<http://w3.avignon.inra.fr/valeri/>), ImagineS (<http://fp7-imagines.eu/>) and BigFoot (<http://www.fsl.orst.edu/larse/bigfoot/>) projects, respectively. Based on the available SPOT and Landsat images, the LAI_{HSI} derived from the GUGM were compared to the given LAI_{HSI-BM} maps to ensure the accuracy of the LAI_{HSI} . The evaluation results show the good agreement between LAI_{HSI} and LAI_{HSI-BM} for different vegetation types, with $R^2 > 0.75$, bias < 0.1 and RMSE < 0.55 (Supplementary material, Fig. S4(a–c)). Note that the performance of GUGM will be evaluated in the 1-km pixel grid, thus the uncertainty of LAI_{HSI} will also be highly reduced (Supplementary material, Fig. S4(d–f)). Then, the spatial representativeness grading and spatial upscaling procedures described in Section 2 (i.e. GUGM) were carefully implemented on all the LAI_{HSI} and their associated ground measurements to generate the LAI validation dataset ($LAI_{site-HSI}$) at 1-km spatial resolution. The comparison of a single acquisition date between $LAI_{site-HSI}$ and aggregated LAI_{HSI-BM} was performed across 500 randomly distributed samples within the area of the LAI benchmark maps, thus a total of 3500 1-km samples ($= 7$ SPOT \times 500 samples) for forest and 2000 1-km samples ($= 4$ Landsat \times 500 samples) for cropland and grassland were used in this evaluation. Note that this evaluation practice was independent of the global LAI validation dataset described in Section 3.3.

In turn, to evaluate the performance of GUGM method, we compared with GUGM three widely used conventional methods for processing $LAI_{site-TS}$ for the validation of the remote sensing LAI products:

(i) using ground measurements without spatial representativeness grading and upscaling, i.e. direct comparison (GM); (ii) using ground measurements with spatial representativeness grading but not upscaling (only measurements satisfying Level 0 spatial representativeness were used) (GGM); and (iii) using upscaled ground measurements without spatial representativeness grading (proposed by Qin et al. (2013) (UGM). All the prepared validation datasets based on these four methods including GUGM, GM, GGM and UGM were compared to the prepared 1-km benchmark LAI samples, and Fig. 3 summarizes a procedural description of the overall evaluation. Note that two evaluation criteria were employed to assess the performance of these methods: RMSE and the total number of validation dataset (N_m). RMSE and N_m represent the uncertainty and size of the prepared LAI validation datasets, respectively. Finally, we also investigated (1) how the GUGM benefits from its spatial representativeness grading and spatial upscaling strategies, and (2) how the temporal availability of the HSI impacts the N_m of each approach in practice.

3.5. Global LAI products and validation under GUGM framework

In this study, three global LAI products, i.e. MODIS, GLASS and GEOV1, were validated using globally available LAI_{site-TS}. Since the performance of the MODIS product was also analyzed by comparison with that of GLASS and GEOV1 at 1-km spatial resolution, we used the MODIS Collection (C5) product in this study. The MODIS LAI product (MOD15A2) C5 was generated at 1-km spatial resolution with an 8-day interval since 2000 (<http://reverb.echo.nasa.gov/>). The main algorithm is based on the biome-specific look-up tables (LUTs) simulated from the three-dimensional (3D) radiative transfer (RT) model (Knyazikhin et al., 1998; Shabanov et al., 2005). Given the daily MODIS red and NIR atmospherically corrected reflectances and the corresponding sun-view geometries as inputs, the mean LAI value is calculated from all the LAI elements for which the corresponding simulated reflectances in the LUTs are close to the MODIS reflectances within specific uncertainties. When the main algorithm fails due to cloud effects or too low sun/view zenith angles, the back-up algorithm based on the LAI-NDVI relationships for each biome is used to retrieve LAI (Myneni et al., 2002). Finally, the LAI corresponding to the maximum Fraction of Photosynthetically Active Radiation absorbed by vegetation (FPAR) over an 8-day period is selected as the product value.

The 8-day GLASS LAI product (<http://glass-product.bnu.edu.cn/en/>) was estimated based on AVHRR from 1981 to 2000 and on MODIS from 2001 to 2015. The spatial resolution of the GLASS LAI product is 0.05-degree in a geographic latitude/longitude projection from 1981 to 2000, and 1-km in a sinusoidal projection from 2001 to 2015 (Xiao et al., 2016). The reflectance was first processed to remove cloud-contamination using a temporal-spatial filtering algorithm and to generate continuous and smooth data using an interpolation algorithm. The key part of this product algorithm is the general regression neural networks (GRNNs), which were trained using fused LAI time-series from MODIS and CYCLOPES products and reprocessed MODIS reflectance (MOD09A1) for each vegetation type over BELMANIP sites from 2001 to 2003 (Xiao et al., 2014). Then, the reprocessed 8-day AVHRR and MODIS reflectance data of a whole year were used as inputs into the GRNNs to retrieve the one-year LAI values.

The GEOV1 LAI product (<http://land.copernicus.eu/global/>) was derived from the SPOT/VEGETATION sensor data at a 10-day interval and 1/112° (approximately 1-km at the equator) spatial resolution in the Plate Carrée projection (Baret et al., 2013). A neural network was used to estimate LAI values for all biome types. First, MODIS and CYCLOPES LAI products were fused and scaled using different weights of these two products (the weight is $\min(1, LAI_{CYC}/4)$ for CYCLOPES and $1 - \min(1, LAI_{CYC}/4)$ for MODIS) over the BELMANIP2 sites for 2003–2004 to generate a new LAI training dataset (Baret et al., 2006), which was combined with the corresponding SPOT/VEGETATION reflectance data to train the neural network. Then, the calibrated neural network was used to generate LAI

products from the SPOT/VEGETATION directionally normalized top of canopy reflectance data (Baret et al., 2013).

The generated LAI time-series validation dataset (LAI_{site-HSI} at 1-km spatial resolution described in Section 3.3) from LAI_{site-TS}, Landsat HSI and GUGM was used to validate the three global LAI products. To investigate the temporal performance of the global LAI products, only valid LAI values of these products were used for validation. For the MODIS LAI product, the main and back-up algorithm can be categorized based on the quality control (QC) layer. Due to the low accuracy of inputs caused by residual clouds or poor atmospheric correction (Wang et al., 2001; Yang et al., 2006a), only the LAI estimates produced using the main algorithm were analyzed in this study. Additionally, the quality flags (i.e., cloud, cloud shadow, cirrus, aerosol, snow, etc.) for each LAI product were used to exclude invalid retrievals from this analysis (see Supplementary material Fig. S5 for details regarding the quality flags applied here). A series of statistical metrics, i.e. R^2 , bias and RMSE, were used to evaluate the accuracy and uncertainty of the product time-series in different years and seasons (Widłowski, 2015).

4. Results and discussion

4.1. Performance evaluation of the GUGM method

Table 2 shows the performance metrics (RMSE and N_m) of the GUGM and the other three methods (GM, GGM and UGM) based on the benchmark 1-km LAI (i.e. LAI_{HSI-BM}) at the forest, cropland and grassland sites, respectively. The DVTP of the randomly selected samples were greater than the DVTP_{Threshold} (= 60%), thus all the measurements were categorized into Levels 0–3. As GUGM and GGM directly used the LAI measurements graded as Level 0 without the spatial upscaling, they had the same RMSE for Level 0. Recall that GUGM applies a rigorous spatial upscaling process to Levels 1–3 cases to consider the heterogeneity, whereas GGM discards any data that does not satisfy the best spatial representativeness (i.e. Level 0). For additional quantification of the benefits of GUGM, the RMSE of Levels 1–3 without upscaling were given (see Levels 1–3 of GGM in Table 2). Compared to the significantly increased RMSE (forest: 0.28 to 0.51, cropland: 0.36 to 0.69, grassland: 0.33 to 0.64) from Levels 1–3 without upscaling, the relatively stable RMSE (forest: 0.21 to 0.25, cropland: 0.12 to 0.20, grassland: 0.13 to 0.18) of the GUGM indicate that the proposed approach can reduce the uncertainty caused by spatial heterogeneity by up to 0.5 LAI units. Note that results also show the degree of uncertainty reduction via GUGM varied among landscape conditions. Moreover, all the ground measurements ($N_m = 3500$) at the forest site could be useable via the GUGM approach, whereas only a few observations were valid in GGM ($N_m = 51$) accounting for < 1.5% of all the ground measurements. This discrepancy between GUGM ($N_m = 2000$) and GGM ($N_m = 0$) is more obvious in both the cropland and grassland sites.

For GM and UGM, both of them were able to use all sets of ground measurements. Overall, the RMSE of GM and UGM at the forest site were 0.46 and 0.26 LAI units, whereas those at the cropland (grassland) site were 0.63 (0.59) and 0.40 (0.39). The RMSE difference between GM and UGM suggests that spatial upscaling without grading can reduce the potential uncertainty from spatial heterogeneity by > 0.2 LAI units. Additionally, we observed relatively larger reduction of RMSE at the cropland than the forest, probably because of the more heterogeneous nature of this vegetation type (Garrigues et al., 2006). Our result also shows that UGM exhibited reasonable performance but was still inferior to GUGM. This result is apparent at the cropland site, which showed the RMSE decreased from 0.40 (UGM) to 0.19 (GUGM). The observed discrepancy between UGM and GUGM was primarily caused by the difference in the spatial representativeness of the ground measurements. GUGM introduces an additional representativeness grading process to divide LAI_{site-TS} into different levels and each level of each sample employs a specific upscaling coefficient (W) to upscale LAI_{site-TS}. In contrast, UGM uses only one W for all the observational dates of each sample, resulting in larger RMSE because the same W

Table 2

The mean RMSE at each level over 500 samples based on GUGM and the other three methods at the forest, cropland and grassland sites. N_m indicates the total number of measurements at each level. The Level 4 results were unavailable because the *DVIP* of all the ground measurements were > 60%. The gray color in the table indicates the final uncertainty of the validation datasets using this method in practice. The “–” indicates that this method does not include this case.

Site type	Method	Mean RMSE					Nm				
		Level 0	Level 1	Level 2	Level 3	All	Level 0	Level 1	Level 2	Level 3	All
Forest	GUGM	0.21	0.21	0.24	0.25	0.24	51	1008	8	2433	3500
	GM	–	–	–	–	0.46	–	–	–	–	3500
	GGM	0.21	0.28	0.50	0.51	0.46	51	1008	8	2433	3500
	UGM	–	–	–	–	0.26	–	–	–	–	3500
Cropland	GUGM	–	0.12	–	0.20	0.19	0	646	0	1354	2000
	GM	–	–	–	–	0.63	–	–	–	–	2000
	GGM	–	0.36	–	0.69	0.63	0	646	0	1354	2000
	UGM	–	–	–	–	0.40	–	–	–	–	2000
Grassland	GUGM	–	0.13	–	0.18	0.18	0	657	0	1343	2000
	GM	–	–	–	–	0.59	–	–	–	–	2000
	GGM	–	0.33	–	0.64	0.59	0	657	0	1343	2000
	UGM	–	–	–	–	0.34	–	–	–	–	2000

introduces larger uncertainty for both high and low representativeness, which will be further discussed below.

To further explore the benefits of the spatial representativeness grading of the $LAI_{site-TS}$ for reducing the uncertainty of the validation dataset, here we show the comparison of errors between UGM and GUGM at the forest site as an example in Fig. 4. We calculated the reduced error (RE) for UGM and GUGM by comparing their upscaled LAI dataset with the LAI_{HSL-BM} ($RE = 100 \times (\Delta UGM - \Delta GUGM) / \Delta UGM$, $\Delta = |LAI_{site-TS} - LAI_{HSL-BM}|$) at the different levels for all the ground measurements ($N_m = 3500$), as shown in Fig. 4(a). The Levels 0 and 2 had relatively greater REs (mean RE = 58.9% for Level 0 and 62.5% for Level 2), whereas Level 3 had the lowest RE (mean RE = 30.5%). This pattern can

be clearly explained by the compositional differences in the graded levels. The upscaling coefficient W is derived from all the available time-series of observations by minimizing the cost function between the LAI measurements (i.e. $LAI_{site-TS}$) and the upscaled LAI (i.e. $LAI_{site-HSL}$) (Eq. (5) in Section 2.2). The upscaling algorithm searches for W by maximizing the degree of explainable variance by the given inputs, so W largely depends on the condition of the input samples (i.e. the consistency of their spatial representativeness, see Fig. S1), and thus the degree of the advantage from the grading process also varies. For instance, all time-series inputs that are graded at a certain level yield identical performance from both UGM and GUGM as the retrieved single upscaling coefficient is sufficient for consistent representativeness. When combining different grade levels, the RE

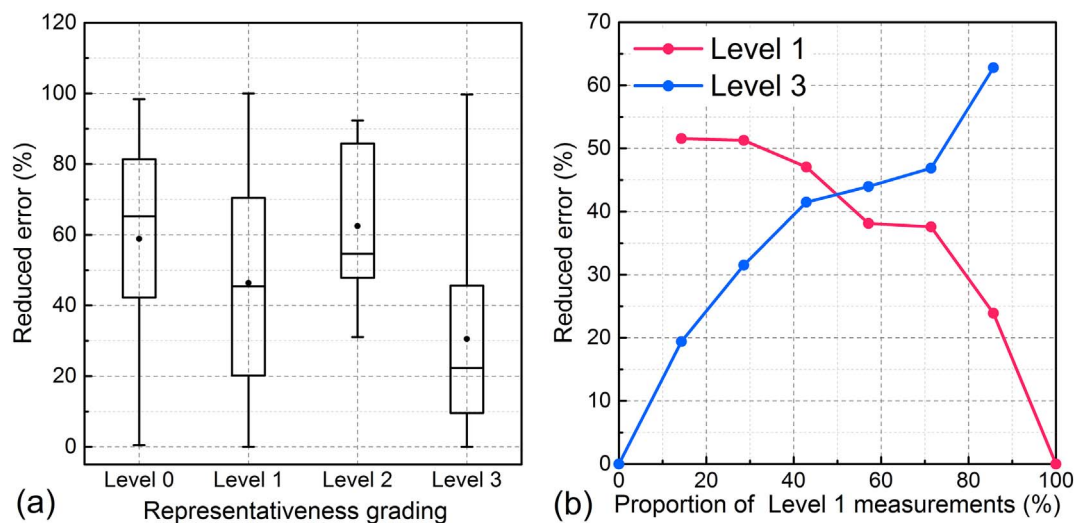


Fig. 4. (a) The boxplot of the reduced error ($RE = 100 \times (\Delta UGM - \Delta GUGM) / \Delta UGM$, $\Delta = |LAI_{site-TS} - LAI_{HSL-BM}|$) for the different spatial representativeness levels. The box stretches from the 25th percentile to the 75th percentile, and the bars indicate the range with in 1.5 IQR (interquartile range). Median and mean values are shown as lines and points, respectively. (b) The variation in the mean RE at certain levels (Levels 1 and 3) decreases as a function of the proportion of Level 1 measurements (100% - proportion of Level 3 measurements) in the inputs when combining different levels.

of certain levels decreases as a function of their dominance in the inputs. Fig. 4(b) obviously shows this tendency from the evaluation datasets. In this study, > 85% of sites had different representativeness levels during the acquisition dates, and Levels 1 and 3 were dominant (Table 2). Thus, the single upscaling coefficient in UGM was more likely to upscale Levels 1 or 3 than Levels 0 and 2 causing relatively lower RE to be observed in Levels 1 and 3 (Fig. 4(a)). Nevertheless, the RE of Levels 1 and 3 were also obvious due to the different degrees of representativeness among the 7 observation dates. These promising results suggest that GUGM significantly improved both the size and accuracy of the validation dataset by combining the strategies of spatial grading and upscaling, compared to the effects of applying only the spatial representativeness grading (GGM) or upscaling (UGM). Note that the N_m of GUGM was ideally quantified in this evaluation practice because all the observation dates for each site could access corresponding HSI. However, the limited availability of HSI in reality is critical when determining the size of the validation dataset, especially for the GUGM.

Here, we additionally investigated the performance on N_m of the four different approaches using all the available $LAI_{site-TS}$ across the sites (Fig. 2 and Table S1). Fig. 5 shows the valid number of $LAI_{site-HSI}$ calculated based on the four approaches and the number of available HSI. Since the GM method does not require HSI and the UGM method requires only several HSI to calculate individual upscaling coefficients, the number of HSI would not significantly affect the valid numbers of $LAI_{site-HSI}$ for these two methods. For the GGM method, only a few measurements at four sites, i.e. TH SKR, US MMS, US HO and TH MKL, could be directly used in the validation of the products, and they were not temporally sufficient to evaluate the temporal performance of the LAI products. Similarly, many observation dates in the $LAI_{site-TS}$ were discarded when aggregating only HSI to obtain the validation dataset. In contrast, GUGM integrates most $LAI_{site-TS}$ and available HSI to generate a more upscaled LAI dataset (40% more HSI and 25% fewer $LAI_{site-TS}$), which is effective not only in assessing the time-series of LAI products but also in providing more reliable validation results than other conventional methods (GM, GGM and UGM). The difference between the number of HSI and GUGM for each site in Fig. 5 indicates the gain from the GUGM back-up strategy, which utilized the same interannual representativeness level for adjacent dates (< 4 days) in different years if no HSI were available.

4.2. Global LAI time-series validation dataset derived from the GUGM method

The first step in the GUGM is to grade the spatial representativeness of

all the $LAI_{site-TS}$. The spatial representativeness of the observed vegetation types in the product pixel grid was evaluated using the DVTP indicator to identify Level 4 measurements. Fig. 6(a) shows the fraction of each site-observed vegetation type within the 1-km pixel grid in 2000 and 2010, respectively. The results present that the fraction of the vegetation type observed at the site was similar between 2000 and 2010 for most sites, suggesting that the vegetation type around the site did not vary substantially. Among these sites, 6 sites from CERN and 4 sites from FLUXNET were graded as Level 4, and thus the LAI measurements from these sites could not be processed by GUGM for the validation of products. The DVTP of the other 28 sites were > 60%, and therefore their representativeness levels were then determined based on RAE and CS indicators. Fig. 6(b) shows the value range of RAE and CS at each site. The boxplot in Fig. 6(b) indicates that the RAE and CS of the different measurements were quite different at the same site due to variations in the vegetation properties at the different growth stages, which suggests that temporal variations in spatial heterogeneities should be carefully considered even at the same site. Based on the thresholds of RAE and CS, Fig. 6(c) shows a total of 1693 $LAI_{site-TS}$ over the 38 sites with providing the respective number of graded each spatial representativeness level. Due to the lack of HSI, GUGM was unable to use approximately 15% of the ground measurements for grading, as shown in gray in Fig. 6(c). As mentioned before, the $LAI_{site-TS}$ at a site displays varying degree of spatial heterogeneity due to the evolving vegetation growth stages, even when it is covered by a single vegetation type. Therefore, the grading results of all the $LAI_{site-TS}$ were generally different at a site. As a result, Level 0, 1, 2, 3 and 4 accounted for 4.5%, 21.0%, 0.3%, 45.0% and 29.2% of all graded measurements, respectively. The low proportion of Level 0 suggests that only a few of the ground measurements could be directly used to validate the LAI products without spatial upscaling. In other words, additional upscaling procedures to reduce the uncertainty induced by lower spatial representativeness are required for those measurements graded in to Levels 1–3.

Fig. 7 shows the obvious differences between $LAI_{site-TS}$ and the spatially upscaled LAI (i.e. $LAI_{site-HSI}$) derived from the GUGM. The degree of LAI discrepancy between these two datasets was determined by the level of spatial representativeness (see both Figs. 7(a) and 6(b–c)). For example, the LAI significantly changed at the US NE and US UMB sites because most measurements were graded as Level 3 (more heterogeneous), whereas the LAI slightly changed at the US MMS site due to the large proportion of Level 0 and 1 measurements. Fig. 7(b) and (c) show $LAI_{site-TS}$ and $LAI_{site-HSI}$ in the temporal trajectory at the cropland site CH YCA and the forest site US MMS, respectively. The $LAI_{site-TS}$ changed considerably after using the GUGM method. The

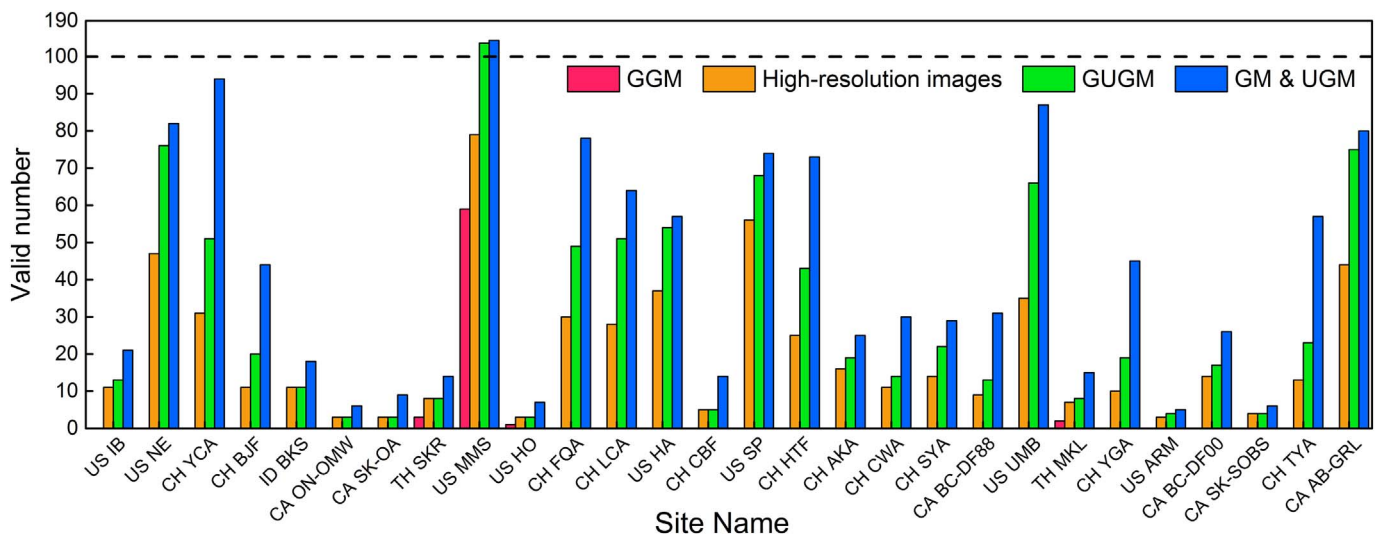


Fig. 5. The valid number of the validation dataset over all the available sites for the four methods and the aggregated high-resolution images. As 10 sites were graded into Level 4, they would not be used in this analysis (see Section 4.2 and Fig. 8(a) for more details).

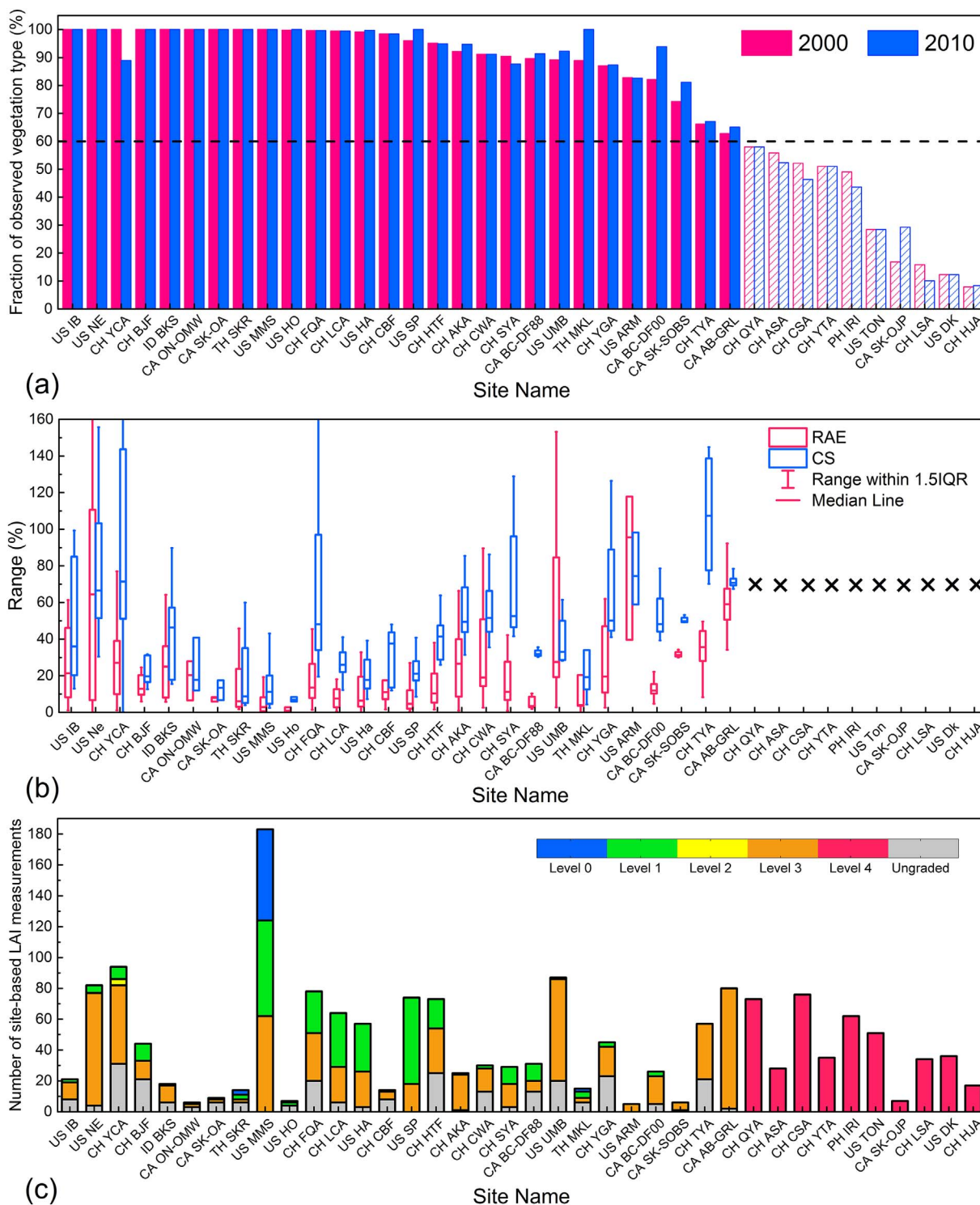


Fig. 6. (a) The DVTP of 38 sites in the 1-km grid in 2000 and 2010. (b) The distributions of RAE (red box) and CS (blue box), which were calculated for all measurements with corresponding HSI at each site. Note that only 28 sites were needed to calculate their RAE and CS because their DVTP are > 60%. The box stretches from the 25th percentile to the 75th percentile. The median value is shown as the line, while the bars indicate the range with in 1.5 IQR (interquartile range). (c) The number of site-based LAI measurements ($LAI_{site-TS}$) from Level 0 to Level 4 and ungraded results over the 38 sites based on the DVTP, RAE and CS indicators in the GUGM method. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mean ratio of the LAI variation between $LAI_{site-TS}$ and $LAI_{site-HSI}$ was 14.4% for Level 1, 41.1% for Level 2 and 175.7% for Level 3 at the CH YCA site, whereas the variation was 5.0% for Level 1 and 15.5% for Level 3 at the US MMS site. The main reason for the LAI variation at each site is that the $LAI_{site-TS}$ with lower spatial representativeness had greater error in representing the 1-km pixel grid, which can cause

greater changes in the LAI after upscaling using the GUGM method. Note that the abnormally large variations in the LAI for Level 3 at the CH YCA site were observed on some dates because the corresponding $LAI_{site-TS}$ were too small (< 0.3), suggesting that the small changes in $LAI_{site-TS}$ can cause abnormally large variations. Additionally, the variation in the LAI for different sites depends on the observed vegetation

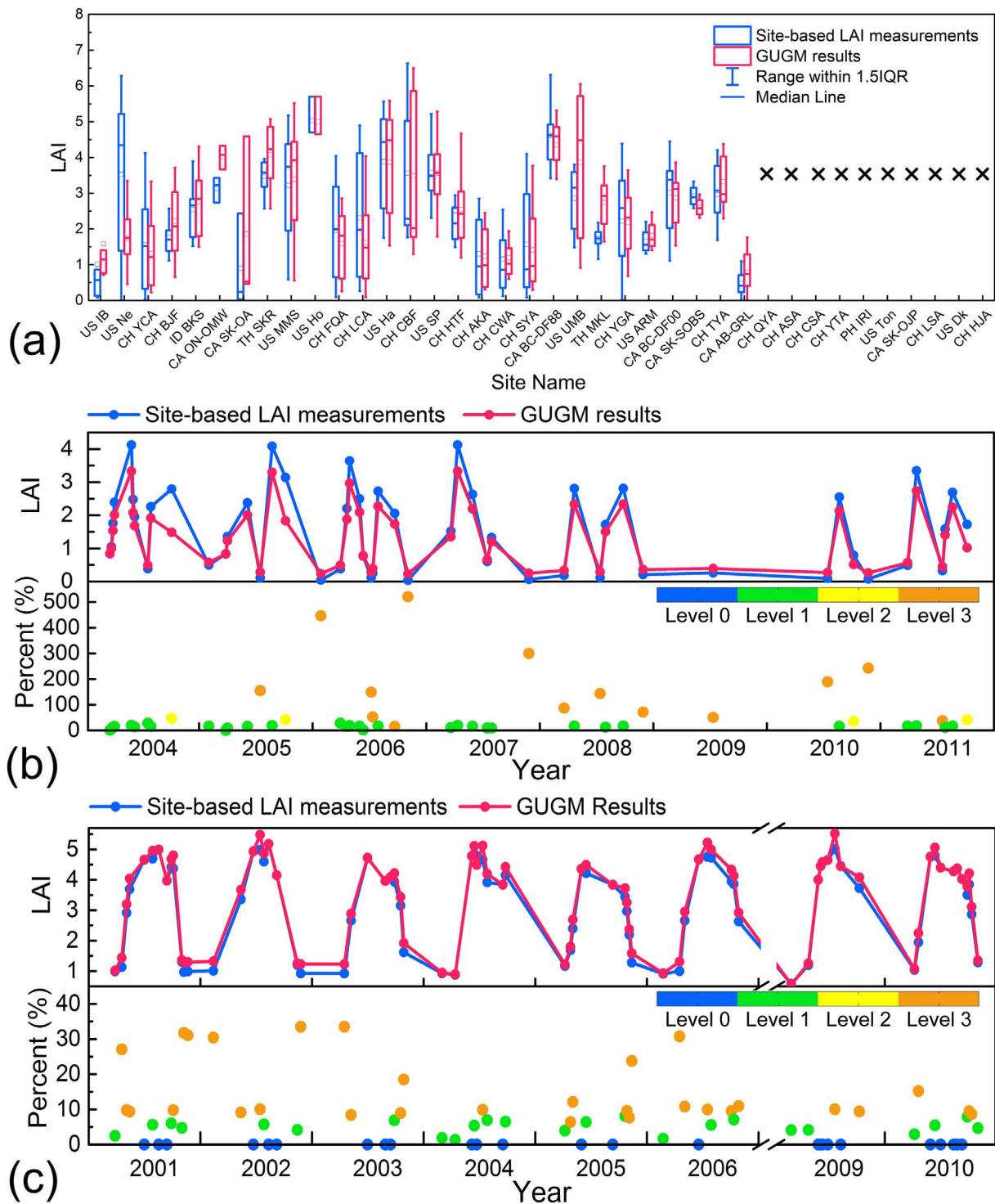


Fig. 7. (a) The differences in the LAI range between the site-based LAI measurements and the spatially upscaled LAI validation dataset derived using the GUGM method for each site and a comparison of the LAI changes between the site-based LAI measurements and the GUGM results on a temporal trajectory at (b) the CH YCA cropland site and (c) the US MMS forest site. The upper portion of (b–c) displays the LAI change for each year. The points in the lower portion show the relative changes (%) in the LAI values, and their colors stand for the different representativeness levels.

type, which will be further discussed in Table 3. Moreover, the LAI_{site-TS} generally exhibited good representativeness during the growing season, with inferior representativeness for other growth stages. This is because transitional growth stages (i.e. green-up or senescence) or the dormant season enhanced the spatial heterogeneity due to variations in developmental speed or the mixture of plant functional types (e.g., deciduous and evergreen) (Ding et al., 2014). The observed inconsistency in representativeness levels for each site demonstrates that the upscaling can

benefit from the spatial representativeness grading process, which is pivotal in GUGM. Finally, the upscaled LAI (LAI_{site-HSI}) could be either higher (Fig. 7(c)) or lower (Fig. 7(b)) than LAI_{site-TS} and the degree of the difference also varied as both the surrounding vegetation condition and its heterogeneity are determinant factors.

The difference between LAI_{site-TS} and LAI_{site-HSI} for each grading level was quantitatively analyzed for all sites and for the different vegetation types. Table 3 shows the statistical indicators associated with

Table 3

The statistical results of the change in LAI from the application of the GUGM method for all site-based LAI measurements at different levels. The “–” indicates no available data.

Type	Indicators	Total	Level 0	Level 1	Level 2	Level 3
All	N	924	65	271	4	584
	Mean Δ LAI	0.59	0	0.51	1.05	0.70
	Mean relative Δ LAI	51.9%	0	29.9%	49.8%	67.8%
Cropland	N	341	–	76	4	261
	Mean Δ LAI	0.82	–	0.67	1.05	0.86
	Mean relative Δ LAI	91.7%	–	56.3%	49.8%	100.3%
Forest	N	508	65	195	–	248
	Mean Δ LAI	0.47	0	0.44	–	0.62
	Mean relative Δ LAI	21.6%	0	19.7%	–	28.7%
Grassland	N	75	–	–	–	75
	Mean Δ LAI	0.38	–	–	–	0.38
	Mean relative Δ LAI	76.0%	–	–	–	76.0%

the LAI difference between the data before and after the application of the GUGM method for all 924 ground measurements at each level, which were divided among cropland, forest and grassland, with 341, 508 and 75 measurements, respectively. Results show that the mean absolute difference in the LAI values (Δ LAI) for the forest site was smaller than that for the cropland site, which implies that the spatial heterogeneity of the forest was generally less than that of the cropland. This conclusion is consistent with the findings of previous research by Garrigues et al. (2006). Moreover, the most notable discrepancies in LAI (mean Δ LAI = 0.86 for cropland, 0.62 for forest and 0.38 for grassland) and its relative form (mean relative Δ LAI = 100.3% for cropland, 28.7% for forest and 76.0% for grassland, relative Δ LAI = $100 \times |LAI_{\text{site-TS}} - LAI_{\text{site-HSI}}| / LAI_{\text{site-TS}}$) were observed for Level 3, which induced larger changes due to greater heterogeneity. As shown in Table 2, the difference between LAI_{site-HSI} (GUGM) and the original LAI_{site-TS} (GM) was caused by the spatial heterogeneity in the product pixel grid. Therefore, the observed significant change shows the uncertainty introduced by the spatial heterogeneity was largely reduced by GUGM compared with LAI_{site-TS} (Section 4.1), which confirms the greater merit of GUGM in reducing the uncertainty of validation datasets caused by the observation scale mismatch between LAI_{site-TS} and the product pixel.

4.3. Product validation using GUGM-based LAI time-series dataset

4.3.1. Overall validation of the global LAI products

Based on the upscaled site-based LAI measurements derived from GUGM (LAI_{site-HSI}) and the quality control (Fig. S5 in Supplementary material) for each LAI product, 572, 922 and 677 valid pixels were selected from MODIS, GLASS and GEOV1, respectively. To achieve a comparable validation result among different products, only pixels with valid LAI ($N = 440$) from all three products were selected for further validation in this study. The overall performance of each product for all biome types is presented in Fig. 8. The results show that GLASS exhibited a little overestimation (bias = 0.11), whereas MODIS and GEOV1 presented a slight discrepancy in LAI values. The lowest RMSE was achieved by GLASS (= 0.72), followed by GEOV1 (= 0.82). The results of this comparison indicate relatively greater uncertainty for MODIS retrievals (RMSE = 0.93), which is partly due to the high uncertainty of the MODIS retrievals over densely forested regions (Shabanov et al., 2005), especially when the LAI was > 4, as shown in Fig. 8(a). Overall, the MODIS validation results are similar to those of similar studies at the global scale, in which the RMSE ranged from 0.78 to 1.17 (Camacho et al., 2013; Fang et al., 2012; Yan et al., 2016). However, using the time-series validation dataset at the global scale, the uncertainties associated with GLASS and GEOV1 are slightly higher than those reported in previous studies ($N = 25$, RMSE = 0.64 for GLASS by Xiao et al. (2014) and $N = 48$, RMSE = 0.74 for GEOV1 by Camacho et al. (2013)).

To further assess the products, each of them was compared to the LAI_{site-HSI} for individual biome types, i.e. grasses, crops, EBF (Evergreen Broadleaf Forest), DBF (Deciduous Broadleaf Forest) and ENF (Evergreen Needleleaf Forest), as shown in Table 4. The mean values and standard deviations of the GUGM LAI and product LAI were also provided. Both the GUGM and products showed the highest mean LAI for EBF, followed by comparable LAI between DBF and ENF. Crops and grasses showed the relatively lower vegetation density, with a mean LAI below 2. For grasses, all the products showed the highest R^2 and lowest uncertainties of all the biomes, with only a slight overestimation. Note that only one grass site was used to validate the products, and the low-magnitude LAI values could be one reason for the low uncertainty of the products. For crops, MODIS showed some underestimation, as was previously reported by Camacho et al. (2013), whereas GLASS and GEOV1 presented minor overestimations. This difference between MODIS and the other two LAI products for crops is possibly because the CYCLOPES LAI values were given higher weight than MODIS in the training dataset in the retrieval process for GLASS and GEOV1, especially for low LAI values of GEOV1 (Baret et al., 2013). Moreover, most of the GLASS LAI values were < 3, and the R^2 between GLASS and GUGM LAI was lower than those between other products, which is partly because some high reflectance may have been removed when using temporal-spatial filtering algorithms to obtain smoother reflectance data (Xiao et al., 2014). Compared to the relatively low uncertainty for grasses and crops, MODIS and GEOV1 had relatively larger RMSEs for forest. For EBF, the number of validation pixels was too small to comprehensively demonstrate the uncertainty for this biome. Nevertheless, it still shows that MODIS and GEOV1 presented quite different LAI because MODIS overestimated LAI while GEOV1 underestimated LAI obviously, which was reported by Camacho et al. (2013) and Yan et al. (2016). The DBF, which is the main biome type for forest, showed the lowest uncertainty achieved by GLASS (RMSE = 0.75) followed by GEOV1 (RMSE = 0.90) and MODIS (RMSE = 1.08). In Fig. 8, MODIS exhibited increasing LAI uncertainties for DBF with increasing LAI, which is an intrinsic limitation of the LAI retrieval algorithm because small variations in reflectances may result in large variation in LAI at condition of saturation (Shabanov et al., 2005). For ENF, MODIS and GLASS exhibited the similar performance, whereas GEOV1 displayed better correspondence with LAI_{site-HSI}.

4.3.2. Analysis of global LAI time-series products

Fig. 9 displays the time-series comparisons among the MODIS, GLASS, GEOV1 and LAI_{site-HSI} from 2001 to 2011 over four selected ground sites representing four biomes (i.e. grasses, crops, DBF and ENF). For the grasses at CA AB-GRL site (49.71° N, 112.94° W) in Canada, MODIS and GLASS tracked the seasonal variation of LAI_{site-HSI} well, but GEOV1 significantly overestimated the LAI at the peak of the growing season (Fig. 9(a)). In terms of the crops at CH FQA site (35.02° N, 114.55° E) in China, all the products well captured the seasonality for double-cropping LAI (Fig. 9(b)). As expected, the benefit of pre-processing in GLASS enabled the retention of most observations across different seasons during 2001–2011, but MODIS provided the fewest LAI, ensuring good observational quality. As shown in the analysis of grass, MODIS and GLASS presented quite similar uncertainties, whereas GEOV1 exhibited greater uncertainty due to the overestimation at the peak of the crop yield stage, especially in 2008 and 2010. For DBF at the US MMS site (39.32° N, 86.41° W) in the US, all the products displayed the distinct seasonality of a mid-latitude temperate forest and a slight overestimation compared to LAI_{site-HSI} (Fig. 9(c)). At this site, MODIS did not provide valid LAI retrievals during summer for most years. This gap in valid retrievals between MODIS and the other products is likely due to unfavorable atmospheric conditions, such as cloud contaminations. Under these challenging conditions, the short composite period (8-day) of MODIS cannot resolve observational gaps without gap-filling or further rigorous processing used in GLASS and GEOV1. This obvious discrepancy also could be found at the ENF site (CA BC-

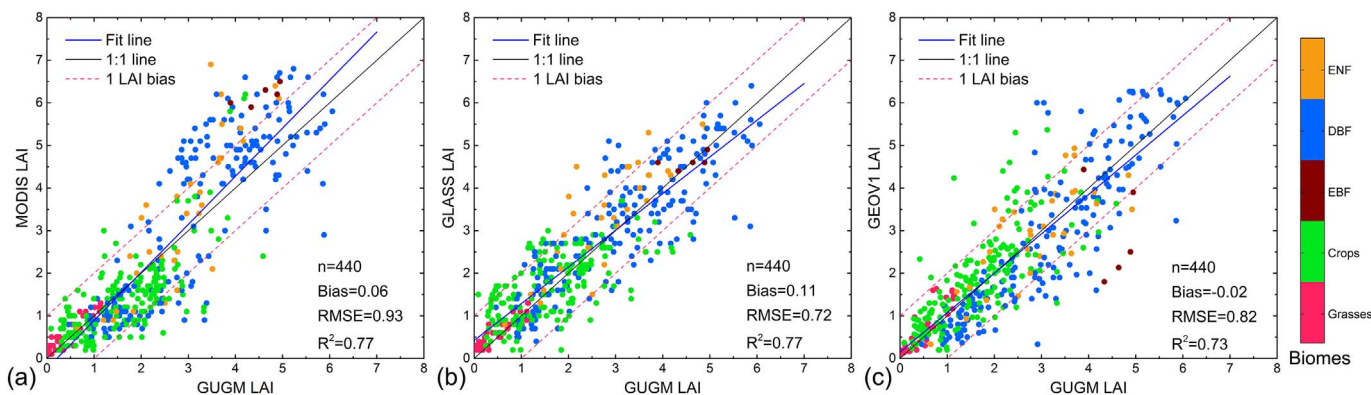


Fig. 8. Comparisons between upscaled site-based LAI measurements derived from the GUGM (GUGM LAI) method and the (a) MODIS, (b) GLASS and (c) GEOV1 products. The different biomes are depicted by different colors. EBF, DBF and ENF stand for “Evergreen Broadleaf Forest”, “Deciduous Broadleaf Forest” and “Evergreen Needleleaf Forest”, respectively.

Table 4

Direct validation results of three global LAI products for different biome types. The mean values and standard deviations of both the GUGM and product LAI are also provided (mean value ± standard deviation). “N” indicates the number of all valid pixels.

Product	Biomes	N	GUGM LAI	Product LAI	R ²	Bias	RMSE
MODIS	Grasses	36	0.50 ± 0.39	0.63 ± 0.39	0.82	0.13	0.21
	Crops	178	1.66 ± 0.93	1.42 ± 0.96	0.52	-0.24	0.74
	EBF	5	4.54 ± 0.43	6.18 ± 0.24	0.61	1.64	1.66
	DBF	189	3.08 ± 1.45	3.28 ± 1.99	0.73	0.20	1.08
	ENF	32	2.86 ± 1.03	3.39 ± 1.80	0.73	0.53	1.17
GLASS	Grasses	36	0.50 ± 0.39	0.60 ± 0.33	0.82	0.10	0.19
	Crops	178	1.66 ± 0.93	1.79 ± 0.77	0.42	0.13	0.74
	EBF	5	4.54 ± 0.43	4.62 ± 0.18	0.28	0.08	0.34
	DBF	189	3.08 ± 1.45	3.09 ± 1.47	0.76	0.01	0.75
	ENF	32	2.86 ± 1.03	3.41 ± 1.26	0.65	0.55	0.92
GEOV1	Grasses	36	0.50 ± 0.39	0.67 ± 0.54	0.80	0.17	0.31
	Crops	178	1.66 ± 0.93	1.86 ± 1.07	0.56	0.20	0.75
	EBF	5	4.54 ± 0.43	2.95 ± 1.15	0.08	-1.59	1.99
	DBF	189	3.08 ± 1.45	2.83 ± 1.82	0.78	-0.25	0.90
	ENF	32	2.86 ± 1.03	2.96 ± 1.13	0.57	0.10	0.75

DF00, 49.87° N, 125.29° W) in Canada, but this result was mostly due to the negative impact of snow (Fig. 9(d)). GLASS and GEOV1 displayed temporally smoother LAI profiles than did MODIS due to the aforementioned pre-processing of the input reflectance. Additionally, unlike DBF, the three products showed more divergent LAI retrievals for the ENF site and MODIS had the lowest uncertainty.

With the strong advantages of GUGM for the validation of product time-series, we further quantified the overall consistency of the uncertainty across different years. Fig. 10 shows the distribution of the RMSE calculated from LAI_{site-HSI} and the three products from 2001 to 2011 (see Fig. S6 in the Supplementary material for the inter-annual variations in the RMSE). Across all the biomes for the overall uncertainty (RMSE) (Fig. 10(a)), GLASS generally presented the lowest magnitude, followed by GEOV1 and MODIS. However, the variation in RMSE in each year indicates that MODIS had more consistent uncertainty through the years for overall biome types than GLASS and GEOV1 (Fig. 10(a)). For the individual biome types, EBF, DBF and ENF were combined into forest due to the small number of valid product pixels for EBF and ENF in different years. All the products showed that the RMSE magnitude of the forest biome from 2001 to 2011 was higher than those of crops and grasses, as shown in Fig. 10(b). This result can be explained by the larger uncertainty of forest for MODIS due to the saturation of high LAI (Shabanov et al., 2005), and both GLASS and GEOV1 also partly adopted MODIS retrievals as training inputs (Baret et al., 2013; Xiao et al., 2014). Moreover, whether the uncertainty of the products is similar in each year also depends on the biome types. MODIS provided more stable RMSE for grass than for crops and forest, whereas GLASS and GEOV1 had more stable uncertainty in the forest biome than in the other two biomes across multi-year validation

practices.

In addition, we assessed the seasonal performance of the three products by dividing the validation dataset (LAI_{site-HSI}) into four seasons, i.e. March–May, June–August, September–November and December–February. Fig. 11(a) shows the seasonal performances of MODIS, GLASS and GEOV1 across different biome types. First, the statistics for the number of validation pixels (N) shows that the valid product pixels were mostly situated from March to November, which corresponds to the critical growth phases of vegetation and the non-snowy period in the northern hemisphere. The RMSE indicates that MODIS and GEOV1 presented the similar variation in their uncertainties across the four seasons, whereas GLASS presented consistent uncertainties except in December–February. For bias, no obvious deviations were observed between MODIS and LAI_{site-HSI}. However, GLASS exhibited significant overestimation in June–August and GEOV1 also displayed overestimation in June–August but underestimation in September–November and December–February. The high R² (≥0.6) in the different seasons indicates that the variations were similar between products and LAI_{site-HSI}.

Fig. 11(b)–(d) displays the seasonal performance of the products for individual biome types. For grasses (Fig. 11(b)), all products had similar and consistent uncertainties in different seasons, with a slight overestimation for LAI in June–August. However, the R² values were quite different for each product in the different seasons, likely due to the limited numbers of valid pixels. For crops (Fig. 11(c)), the number of ground measurements in December–February was too small to statistically analyze their uncertainties. The uncertainties of each product in the other seasons were comparable, but their deviations described by the bias were quite different in the different seasons. MODIS and GLASS presented obvious underestimation in March–May, whereas GLASS and GEOV1 displayed evident overestimation in June–August. The R² had similar seasonal variation for each product, but sudden divergence was observed from GLASS in December–February. For the forest biome (Fig. 11(d)), the uncertainties of each product were quite different in the four seasons. MODIS and GLASS tracked similar seasonal variations in the RMSE across seasons, but MODIS showed relatively larger RMSEs than GLASS. Both MODIS and GLASS had similar uncertainties from March to November, whereas GEOV1 showed relatively consistent uncertainties from June to February. In addition, MODIS exhibited LAI overestimation across entire seasons, especially spring and summer. Conversely, GEOV1 showed obvious underestimation except in June–August. The observed changes in the product performance in the different years and seasons emphasize that assessing the temporal uncertainty should be a pre-requisite to more comprehensively understand the global LAI products. The GUGM approach proposed here provides an effective way to prepare larger and higher quality validation dataset, which in turn enables us to assess the uncertainty of global LAI products with higher confidence.

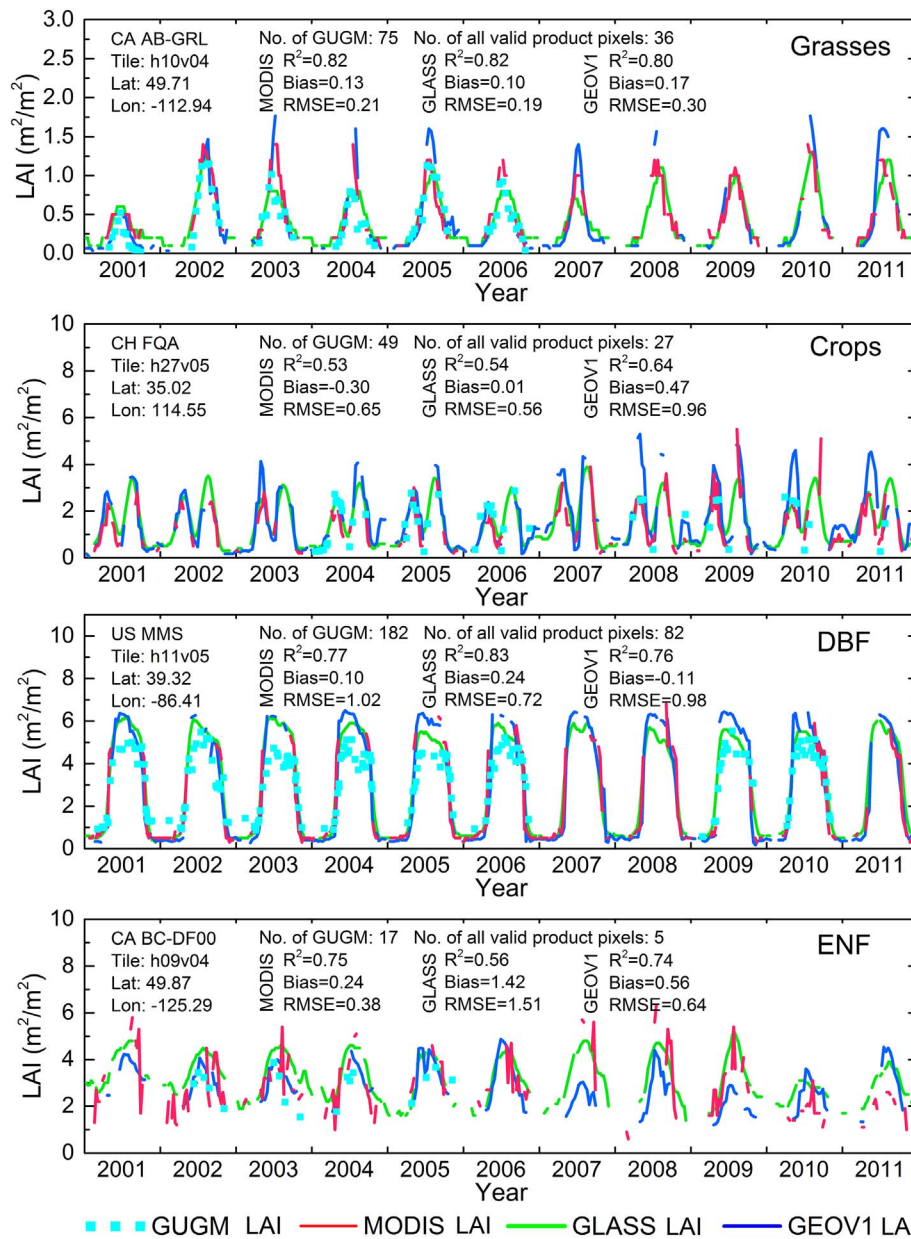


Fig. 9. The temporal trajectories of the MODIS (red line), GLASS (green line) and GEOV1 (blue line) LAI products for (a) grasses, (b) crops, (c) DBF and (d) ENF from 2001 to 2011. The square points indicate the GUGM LAI results based on the site-based LAI measurements. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Conclusions

The spatial scale mismatch between the area of ground measurements and the product pixel grid limits the utilization of site-based LAI measurements for validating satellite LAI products. Our GUGM approach, which integrates a spatial representativeness grading and a spatial upscaling strategy, is proposed here to resolve this scale-mismatch issue and maximize the use of time-series of ground measurements. This approach first ingests both high-resolution images and site-based LAI measurements to capture the spatiotemporal variability in the product pixel grid. Then, the approach calculates the spatial upscaling coefficient based on site-based LAI measurements and aggregated high-resolution reference maps to derive a reliable LAI time-series validation dataset. In this study, we carefully evaluated the performances of the GUGM by comparing to the benchmark LAI and other widely used approaches. Based on the LAI time-series validation dataset derived from the GUGM, we additionally assessed the uncertainty of

three global LAI products, including MODIS, GLASS and GEOV1. Considering all evaluation results together, GUGM can (a) significantly reduce the uncertainty from spatial scale mismatch and (b) increase the size of the available validation dataset. These results are strongly corroborated as GUGM outperforms other widely used approaches across various biomes in the above two respects. By implementing the GUGM to validate three global LAI products, we also confirmed that the proposed approach is useful to assess the temporal performance of the products. We believe that this result is critical to better understand the structure of uncertainty and its evolution along different seasonal or annual contexts. The quantitative validation practice for three global products shows GLASS performs better with GUGM LAI than GEOV1 and MODIS for overall biome types. All products can well capture the seasonal variation witnessed by GUGM LAI. MODIS LAI exhibits noisy temporal variation at forest sites, which may be primarily caused by the short compositing period and unsmoothed input reflectances. Across multiple years, MODIS displays more consistent performance than other

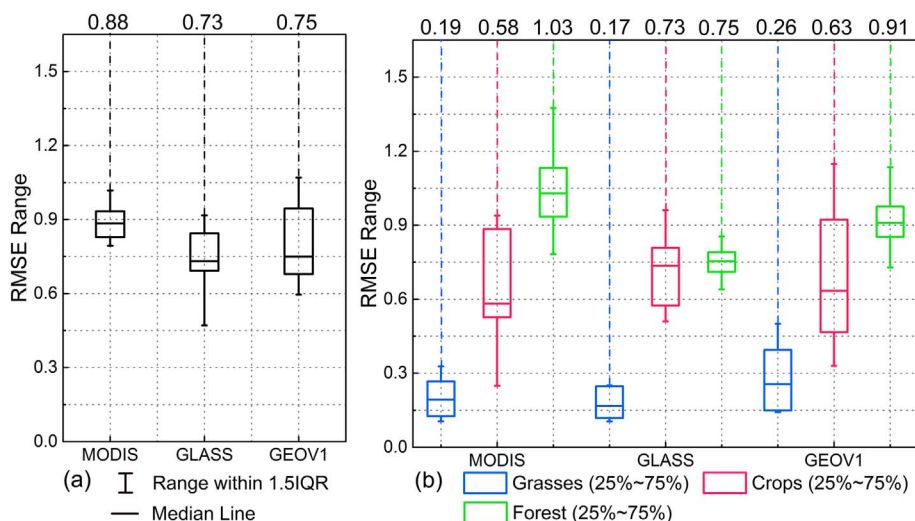


Fig. 10. The distribution of RMSE derived from comparisons among three global LAI products and GUGM LAI for (a) overall biome types and (b) each biome type from 2001 to 2011. The box stretches from the 25th percentile to the 75th percentile. The numbers at the top of figure correspond to the median RMSE within each boxplot. The bars indicate the range with in 1.5 IQR (interquartile range).

products for overall biome types and all products at different biome types show the variation of uncertainty. At seasonal scale, both MODIS and GEOV1 present the similar variation of uncertainty, but their magnitudes are higher than GLASS for overall biome types. The promising GUGM approach shows great potentials in the realization of Stage 3 validation using temporally continuous ground LAI

measurements from global networks and also provides fundamental information to further improve LAI algorithms.

Additionally, future studies can focus on several issues to improve the potential applications of the proposed GUGM. First, due to the different sensors and processing chains for various products, the effect of the point spread function (PSF) should be properly considered in the 1-km pixel grid to

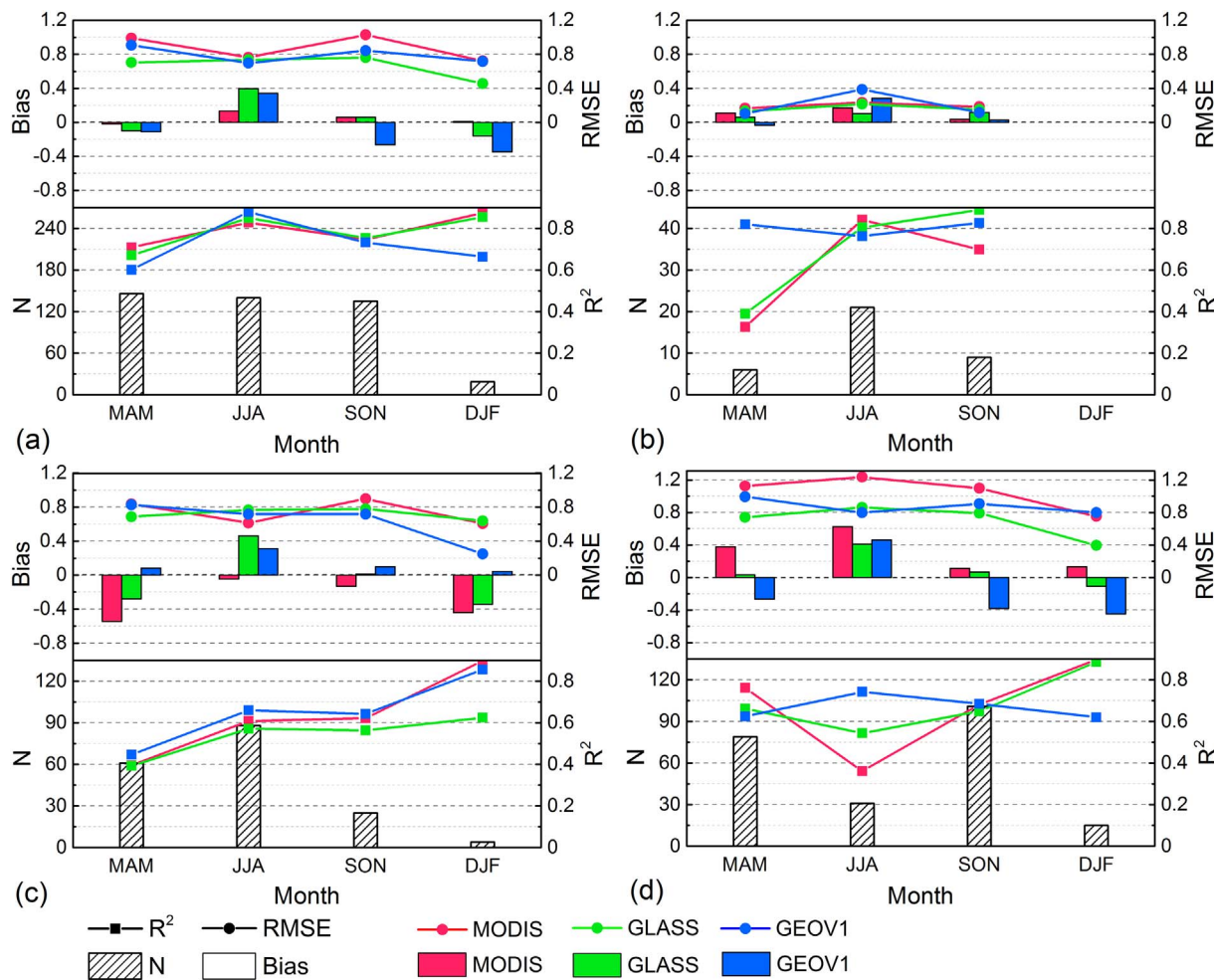


Fig. 11. The number of valid validation pixels and the derived statistical metrics (R^2 , bias and RMSE) across four seasons in MODIS (red), GLASS (green) and GEOV1 (blue) for (a) overall biome types, (b) grasses, (c) crops and (d) forest. The “MAM”, “JJA”, “SON” and “DJF” in the x-axis stand for March–May, June–July, September–November and December–February, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

achieve a robust result for the validation of different global products. Second, GUGM has a large potential to validate the current and future products in the next coming years with the help of newly developed techniques for site-based LAI measurements and high-resolution images. For example, the new instruments (LAINet) have been developed to temporally monitor the LAI based on a wireless sensor network (WSN) (Qu et al., 2014a, 2014b). The recent availability of high temporal frequency high-resolution images (Harmonized Landsat Sentinel-2 product) should enable us to utilize more site-based LAI measurements by providing more time-matched surface reflectance products. Third, the GUGM can perform an essential role in synthesizing global site-based LAI measurements from various research networks, such as FLUXNET, the Terrestrial Ecosystem Research Network (TERN), the National Ecological Observatory Network (NEON), the Integrated Carbon Observation Network (ICOS), etc. The synthesized datasets can offer a unique opportunity to more comprehensively quantify the uncertainties of satellite products and provide the ecosystem-level information regarding vegetation structural changes. We believe that the universality of this approach will be beneficial to scientific communities to evaluate their products using global networks of site-based measurements.

Acknowledgments

This work was supported by the National Basic Research Program of China (973 Program) (No. 2013CB733401) and the National Natural Science Foundation of China (No. 41671374, No. 41271366). Additional funding was provided by the National Aeronautics and Space Administration Earth Science Division (Grant No. NNX14AP80A, NNX14AI71G, NNX16AO34H). The authors thank FLUXNET and CERN for providing the field LAI datasets. The 30-m land cover datasets were provided by the National Geomatics Center of China. We also are grateful for the Landsat Surface Reflectance products courtesy of the U.S. Geological Survey Earth Resources Observation and Science Center for providing the Landsat TM/ETM+ surface reflectance data. The MODIS Collection 5 LAI products were downloaded from the Land Processes Distributed Archive Center (LP DAAC). The GLASS LAI products were downloaded from the BNU Center for Global Change Data Processing and Analysis. The GEOV1 LAI products were downloaded from the Copernicus Global Land Service. The authors would like to thank the three anonymous reviewers for their constructive comments and suggestions that helped to improve the paper.

Appendix A. Abbreviations, symbols, and descriptions

Table A
Abbreviations or symbols in this study and their descriptions.

Abbreviation or symbol	Description	Abbreviation or symbol	Description
AVHRR	Advanced very high resolution radiometer	LAI _{HSI}	High spatial resolution (e.g. 30-m) LAI map
BELMANIP	Benchmark Land Multisite Analysis and Intercomparison of Products	LAI _{HSI-AGG}	Aggregated high-resolution LAI map (e.g. 30-m spatial resolution) to the product pixel grid (e.g. 1-km)
CEOS	Committee on Earth Observation Satellites	LAI _{HSI-BM}	High-resolution LAI benchmark map provided by field campaigns (e.g. VALERI, BigFoot, ImagineS...)
CERN	Chinese ecosystem research network	LAI _{site-HSI}	Generated LAI validation dataset in the product pixel grid (e.g. 1-km) based on site measurements and high-resolution LAI maps
CS	Coefficient of sill	LAI _{site-TS}	Time-series of site-based LAI measurements
C5	Collection 5	LEDAPS	Landsat ecosystem disturbance adaptive processing system
DBF	Deciduous broadleaf forest	LPV	Land product validation
DVTP	Dominant vegetation type percent	LST	Land surface temperature
EBF	Evergreen broadleaf forest	LUT	Look-up table
ENF	Evergreen needleleaf forest	MODIS	Moderate resolution imaging spectroradiometer
EROS	Earth Resources Observation and Science	NDVI	Normalized differential vegetation index
ESPA	Center Science Processing Architecture	NIR	Near infra-red
ESUs	Elementary sampling units	N _m	Number of measurements
ETM +	Enhanced thematic mapper plus	QC	Quality control
FPAR	Fraction of photosynthetically active radiation absorbed by vegetation	RAE	Relative absolute error
GEOV1	Geoland2 version 1	RE	Reduced error from UGM to GUGM
GGM	Grading ground measurements	RMSE	Root mean square error
GLASS	Global land surface satellite	RTMs	Radiative transfer models
GM	Ground measurements	SAIL	Scattering from arbitrarily inclined leaves
GRNN	General regression neural networks	TM	Thematic mapper
GUGM	Grading and upscaling Ground Measurements	UGM	Upscaling ground measurements
HSI	High spatial resolution image	USGS	US Geological Survey
LAI	Leaf Area Index	VALERI	Validation of Land European Remote sensing Instruments

Appendix B. Supplementary data

Supplementary data associated with this article can be found in the online version, at <https://doi.org/10.1016/j.rse.2018.02.049>. These data include the Google map of the site locations described in this article.

References

- Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Am. Meteorol. Soc.* 82, 2415–2434.
- Baret, F., Morisette, 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 inter-comparison of land biophysical products: proposition of the CEOS-BELMANIP. *IEEE Trans. Geosci. Remote Sens.* 44, 1794–1803.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J.L., Leroy, M., 2007. LAI, FAPAR and fCover CYCLOPES global products derived from VEGETATION: part 1: principles of the algorithm. *Remote Sens. Environ.* 110, 275–286.
- 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. *Remote Sens. Environ.* 137, 299–309.
- Barr, A.G., Black, T.A., Hogg, E.H., Kljun, N., Morgenstern, K., Nesic, Z., 2004. Inter-annual variability in the leaf area index of a boreal aspen-hazelnut forest in relation to net ecosystem production. *Agric. For. Meteorol.* 126, 237–255.
- Bi, J., Knyazikhin, Y., Choi, S., Park, T., Barichivich, J., Ciaia, P., Fu, R., Ganguly, S., Hall, F., Hilker, T., Huete, A., Jones, M., Kimball, J., Lyapustin, A.I., Möttus, M., Nemani, R.R., Piao, S., Poulter, B., Saleska, S.R., Saatchi, S.S., Xu, L., Zhou, L., Myneni, R.B., 2015. Sunlight mediated seasonality in canopy structure and photosynthetic activity of Amazonian rainforests. *Environ. Res. Lett.* 10, 064014.
- Bonan, G.B., 1995. Land-atmosphere interactions for climate system models: coupling biophysical, biogeochemical, and ecosystem dynamical processes. *Remote Sens. Environ.* 51, 57–73.
- 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. *Remote Sens. Environ.* 137, 310–329.
- Chen, J.M., Black, T.A., 1992. Defining leaf area index for non-flat leaves. *Plant Cell Environ.* 15, 421–429.
- Chen, J.M., Leblanc, S.G., 1997. A four-scale bidirectional reflectance model based on canopy architecture. *IEEE Trans. Geosci. Remote Sens.* 35, 1316–1337.
- Chen, T., Martin, E., 2009. Bayesian linear regression and variable selection for spectroscopic calibration. *Anal. Chim. Acta* 631, 13–21.
- Chen, B., Coops, N.C., Fu, D., Margolis, H.A., Amiro, B.D., Black, T.A., Arain, M.A., Barr, A.G., Bourque, C.P.A., Flanagan, L.B., Lafleur, P.M., McCaughey, J.H., Wofsy, S.C., 2012. Characterizing spatial representativeness of flux tower eddy-covariance measurements across the Canadian Carbon Program Network using remote sensing and footprint analysis. *Remote Sens. Environ.* 124, 742–755.
- Chen, J., Ban, Y., Li, S., 2014. China: open access to earth land-cover map. *Nature* 514, 434.
- Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30 m resolution: a POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* 103, 7–27.
- Claverie, M., Vermote, E.F., Weiss, M., Baret, F., Hagolle, O., Demarez, V., 2013. Validation of coarse spatial resolution LAI and FAPAR time series over cropland in southwest France. *Remote Sens. Environ.* 139, 216–230.
- Ding, Y., Zhao, K., Zheng, X., Jiang, T., 2014. Temporal dynamics of spatial heterogeneity over cropland quantified by time-series NDVI, near infrared and red reflectance of Landsat 8 OLI imagery. *Int. J. Appl. Earth Obs. Geoinf.* 30, 139–145.
- Fang, H., Wei, S., Liang, S., 2012. Validation of MODIS and CYCLOPES LAI products using global field measurement data. *Remote Sens. Environ.* 119, 43–54.
- Fensholt, R., Sandholt, I., Rasmussen, M.S., 2004. Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. *Remote Sens. Environ.* 91, 490–507.
- Fu, B., Li, S., Yu, X., Yang, P., Yu, G., Feng, R., Zhuang, X., 2010. Chinese ecosystem research network: progress and perspectives. *Ecol. Complex.* 7, 225–233.
- Garrigues, S., Allard, D., Baret, F., Weiss, M., 2006. Quantifying spatial heterogeneity at the landscape scale using variogram models. *Remote Sens. Environ.* 103, 81–96.
- Garrigues, S., Lacaze, R., Baret, F., Morisette, J.T., Weiss, M., Nickeson, J.E., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R.B., Knyazikhin, Y., Yang, W., 2008. Validation and intercomparison of global Leaf Area Index products derived from remote sensing data. *J. Geophys. Res.* 113 (G02028).
- Gonsamo, A., Chen, J.M., 2014. Continuous observation of leaf area index at Fluxnet-Canada sites. *Agric. For. Meteorol.* 189–190, 168–174.
- Kim, J., Guo, Q., Baldocchi, D.D., Leclerc, M.Y., Xu, L., Schmid, H.P., 2006. Upscaling fluxes from tower to landscape: overlaying flux footprints on high-resolution (IKONOS) images of vegetation cover. *Agric. For. Meteorol.* 136, 132–146.
- Knyazikhin, Y., Martonchik, J.V., Myneni, R.B., Diner, D.J., Running, S.W., 1998. 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.
- Liu, J., Chen, J.M., Cihlar, J., Park, W.M., 1997. A process-based boreal ecosystem productivity simulator using remote sensing inputs. *Remote Sens. Environ.* 62, 158–175.
- Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Feng, G., Kutler, J., Teng-Kui, L., 2006. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* 3, 68–72.
- Mather, P., Koch, M., 2010. *Computer Processing of Remotely-sensed Images: An Introduction*. John Wiley & Sons.
- Morisette, 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., Kalaeska, M., Sanchez-Azofeifa, G.A., Chubey, M., Rivard, B., Stenberg, P., Rautiainen, M., Voipio, P., Manninen, T., Pilant, A.N., Lewis, T.E., Iames, 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. Remote Sens.* 44, 1804–1817.
- Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.R., Lotsch, A., Friedl, M., Morisette, J.T., Votava, P., Nemani, R.R., Running, S.W., 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.* 83, 214–231.
- Qin, J., Yang, K., Lu, N., Chen, Y., Zhao, L., Han, M., 2013. Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia. *Remote Sens. Environ.* 138, 1–9.
- Qu, Y., Han, W., Fu, L., Li, C., Song, J., Zhou, H., Bo, Y., Wang, J., 2014a. LAInet – a wireless sensor network for coniferous forest leaf area index measurement: design, algorithm and validation. *Comput. Electron. Agric.* 108, 200–208.
- Qu, Y., Zhu, Y., Han, W., Wang, J., Ma, M., 2014b. Crop leaf area index observations with a wireless sensor network and its potential for validating remote sensing products. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 431–444.
- Richardson, A.D., Anderson, R.S., Arain, M.A., Barr, A.G., Bohrer, G., Chen, G., Chen, J.M., Ciaia, P., Davis, K.J., Desai, A.R., Dietze, M.C., Dragoni, D., Garrity, S.R., Gough, C.M., Grant, R., Hollinger, D.Y., Margolis, H.A., McCaughey, H., Migliavacca, M., Monson, R.K., Munger, J.W., Poulter, B., Raczka, B.M., Ricciuto, D.M., Sahoo, A.K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., Xue, Y., 2012. Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. *Glob. Chang. Biol.* 18, 566–584.
- Ross, J., 1989. *The Radiation Regime and Architecture of Plant Stands*. Dr. W. Junk Publ, The Hague, The Netherlands.
- Samanta, A., Knyazikhin, Y., Xu, L., Dickinson, R.E., Fu, R., Costa, M.H., Saatchi, S.S., Nemani, R.R., Myneni, R.B., 2012. Seasonal changes in leaf area of Amazon forests from leaf flushing and abscission. *J. Geophys. Res. Biogeosci.* 117 (G01015).
- Sellers, P.J., Dickinson, R.E., Randall, D.A., Bets, A.K., Hall, F.G., Berry, J.A., Collatz, G.J., Denning, A.S., Mooney, H.A., Nobre, C.A., Sato, N., Field, C.B., Henderson-Sellers, A., 1997. Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science* 275, 502–509.
- Shabanov, N.V., Huang, D., Yang, W.Z., Tan, B., Knyazikhin, Y., Myneni, R.B., Ahl, D.E., Gower, S.T., Huete, A.R., Aragao, L.E.O.C., Shimabukuro, Y.E., 2005. Analysis and optimization of the MODIS leaf area index algorithm retrievals over broadleaf forests. *IEEE Trans. Geosci. Remote Sens.* 43, 1855–1865.
- Tan, B., Hu, J.N., Zhang, P., Huang, D., Shabanov, N., Weiss, M., Knyazikhin, Y., Myneni, R.B., 2005. Validation of Moderate Resolution Imaging Spectroradiometer leaf area index product in croplands of Alpiilles, France. *J. Geophys. Res.-Atmos.* 110 (D01107).
- Tang, S., Chen, J.M., Zhu, Q., Li, X., Chen, M., Sun, R., Zhou, Y., Deng, F., Xie, D., 2007. LAI inversion algorithm based on directional reflectance kernels. *J. Environ. Manag.* 85, 638–648.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics.
- Verger, A., Baret, F., Camacho, F., 2011. Optimal modalities for radiative transfer-neural network estimation of canopy biophysical characteristics: evaluation over an agricultural area with CHRIS/PROBA observations. *Remote Sens. Environ.* 115, 415–426.
- Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. *Remote Sens. Environ.* 16, 125–141.
- Wang, Y.J., Tian, Y.H., Zhang, Y., El-Saleous, N., Knyazikhin, Y., Vermote, E., Myneni, R.B., 2001. Investigation of product accuracy as a function of input and model uncertainties: case study with SeaWiFS and MODIS LAI/fPAR algorithm. *Remote Sens. Environ.* 78, 299–313.
- Weiss, M., Baret, F., Garrigues, S., Lacaze, R., 2007. LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: validation and comparison with MODIS collection 4 products. *Remote Sens. Environ.* 110, 317–331.
- Widlowski, J.-L., 2015. Conformity testing of satellite-derived quantitative surface variables. *Environ. Sci. Pol.* 51, 149–169.
- Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L., Song, J., 2014. Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance. *IEEE Trans. Geosci. Remote Sens.* 52, 209–223.
- Xiao, Z., Liang, S., Wang, J., Xiang, Y., Zhao, X., Song, J., 2016. Long-time-series global land surface satellite leaf area index product derived from MODIS and AVHRR surface reflectance. *IEEE Trans. Geosci. Remote Sens.* 54, 5301–5318.
- Xu, B., Li, J., Liu, Q., Huete, A., Yu, Q., Zeng, Y., Yin, G., Zhao, J., Yang, L., 2016. Evaluating spatial representativeness of station observations for remotely sensed leaf area index products. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 9, 3267–3282.
- Yan, K., Park, T., Yan, G., Liu, Z., Yang, B., Chen, C., Nemani, R., Knyazikhin, Y., Myneni, R., 2016. Evaluation of MODIS LAI/fPAR product collection 6. Part 2: validation and intercomparison. *Remote Sens.* 8, 460.
- 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. Remote Sens.* (In press).
- Yang, W., Huang, D., Tan, B., Stroeve, J.C., Shabanov, N.V., Knyazikhin, Y., Nemani, R.R.,

- Myneni, R.B., 2006a. Analysis of leaf area index and fraction of PAR absorbed by vegetation products from the terra MODIS sensor: 2000–2005. *IEEE Trans. Geosci. Remote Sens.* 44, 1829–1842.
- Yang, W., Tan, B., Huang, D., Rautiainen, M., Shabanov, N.V., Wang, Y.J., Privette, J.L., Huemmrich, K.F., Fensholt, R., Sandholt, I., Weiss, M., Ahl, D.E., Gower, S.T., Nemani, R.R., Knyazikhin, Y., Myneni, R.B., 2006b. MODIS leaf area index products: from validation to algorithm improvement. *IEEE Trans. Geosci. Remote Sens.* 44, 1885–1898.
- Yin, G., Li, J., Liu, Q., Fan, W., Xu, B., Zeng, Y., Zhao, J., 2015. Regional leaf area index retrieval based on remote sensing: the role of radiative transfer model selection. *Remote Sens.* 7, 4604.
- Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R.R., Myneni, B.R., 2013. Global data sets of vegetation leaf area index (LAI)3g and fraction of photosynthetically active radiation (FPAR)3g derived from global inventory modeling and mapping studies (GIMMS) normalized difference vegetation index (NDVI3g) for the period 1981 to 2011. *Remote Sens.* 5, 927–948.
- 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., Penuelas, 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. Chang.* 6, 791–795.