Article:

Zhang, H., Li, J., Liu, Q. et al. (13 more authors) (2022) A novel red-edge spectral index for retrieving the leaf chlorophyll content. Methods in Ecology and Evolution. ISSN 2041-210X

https://doi.org/10.1111/2041-210x.13994

Received: 13 January 2022 | Accepted: 5 September 2022

DOI: 10.1111/2041-210X.13994

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

A novel red-edge spectral index for retrieving the leaf chlorophyll content

Hu Zhang[1,2] | Jing Li[1,2] | Qinhuo Liu[1,2] | Shangrong Lin[3] | Alfredo Huete[4] | Liangyun Liu[5] | Holly Croft[6] | Jan G. P. W. Clevers[7] | Yelu Zeng[8] | Xiaohan Wang[1,2] | Chenpeng Gu[1,2] | Zhaoxing Zhang[1] | Jing Zhao[1] | Yadong Dong[1] | Faisal Mumtaz[1,2] | Wentao Yu[1]

[1]State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing Normal University, Beijing, China; [2]University of Chinese Academy of Sciences, Beijing, China; [3]School of Atmospheric Sciences, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-sen University, Zhuhai, China; [4]School of Life Sciences, University of Technology Sydney, Broadway, New South Wales, Australia; [5]Key Laboratory of Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China; [6]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK; [7]Laboratory of Geo-Information Science and Remote Sensing, Wageningen University & Research, Wageningen, The Netherlands and [8]College of Land Science and Technology, China Agricultural University, Beijing, China

Correspondence
Jing Li
Email: lijing200531@aircas.ac.cn

Funding information
National Key Research and Development Program, Grant/Award Number: 2019YFE0126700; National Natural Science Foundation of China, Grant/Award Number: 41871265

Handling Editor: Jessica Royles

Abstract

1. The leaf chlorophyll content (Chl_{leaf}) is a crucial vegetation parameter in carbon cycle modelling and agricultural monitoring at local, regional and global scales. The red-edge spectral region is sensitive to variations in Chl_{leaf}. An increasing number of sensors are capable of sampling red-edge bands, providing opportunities to estimate Chl_{leaf}. However, the contributions of canopy/foliar/soil factors are always combined in the reflectance signal, which limits the generalizability of vegetation index (VI)-based Chl_{leaf} inversions. This study aims to propose a new red-edge chlorophyll index to decouple the effects of the canopy and soil background from the Chl_{leaf} estimation.
2. The chlorophyll sensitive index (CSI) was proposed, and the regression equations between the CSI and Chl_{leaf} were acquired using PROSAIL (PROSPECT + SAIL) and the 4-Scale-PROSPECT model.
3. Sensitivity analyses showed that the CSI is resistant to variations in the canopy structure and soil background. Validation results obtained using 308 ground-measured samples over nine sites world-wide revealed that CSI improves the Chl_{leaf} retrieval accuracy (root mean square error (RMSE = $9.39\,\mu g\,cm^{-2}$) compared with the existing Medium Resolution Imaging Spectrometer (MERIS) terrestrial chlorophyll index (MTCI; RMSE = $13.00\,\mu g\,cm^{-2}$). Moreover, the CSI method steadily achieves a highly accurate inversion under different LAI and Chl_{leaf} conditions. Based on the CSI regression method, a Chl_{leaf} product with a 30-m/10-day resolution across China was generated.

4. The CSI is sensitive to Chl_{leaf} but resistant to canopy structure and soil moisture parameters, and it has the potential to explicitly retrieve leaf-scale biochemistry in ecosystem modelling and ecological applications.

KEYWORDS
chlorophyll content estimation, chlorophyll product, PROSAIL model, remote sensing, spectral vegetation index

1 | INTRODUCTION

The leaf chlorophyll content (Chl_{leaf}) is a key indicator of the physiological condition of vegetation and is integral for the harvesting of solar radiation required to drive photosynthesis (Evans, 1989; Vernon & Seely, 1966). Retrievals of Chl_{leaf} are crucial for providing important information on plant stress and diseases, modelling plant productivity and serving as a proxy for the photosynthetic capacity within terrestrial biosphere models (Croft et al., 2017; Luo et al., 2019). The provision of accurate and spatially and temporally continuous Chl_{leaf} data at a user-relevant spatial resolution is very important for ecological science.

Remote sensing provides a practical approach to obtaining Chl_{leaf} across large spatial swaths. The red-edge wavelength reflectance (680 ~ 750 nm), which sharply increases from the red band absorption maxima to the near-infrared (NIR) shoulder, is most sensitive to chlorophyll and experiences less saturation in the presence of high chlorophyll contents (Croft & Chen, 2018). An increasing number of satellite sensors have sampled this spectral region since the 2000s (e.g. Environmental Satellite (ENVISAT) Medium Resolution Imaging Spectrometer (MERIS), Sentinel-2 Multispectral Instrument (MSI), Sentinel-3 Ocean and Land Colour Instrument (OLCI), RapidEye, WorldView-2 and Gaofen-6). These data provide opportunities to estimate chlorophyll contents at different temporal and spatial scales.

Two methods have been widely used to estimate Chl_{leaf} from remote sensing data: physically based radiative transfer modelling (RTM) and empirical vegetation index (VI)-based approaches. The RTM approach using inversion methods of look-up tables (LUTs) (Croft et al., 2020; Zarco-Tejada et al., 2019) and machine learning methods (Verrelst et al., 2012) allows us to model the physical mechanisms underpinning the light interaction with leaves or canopies. Given an accurate parameterization, the scattering contributions from different scenes or leaf components can be simulated and then retrieve Chl_{leaf}. Based on the PROSPECT+SAIL (PROSAIL) model (Jacquemoud et al., 2009) and the 4-Scale-PROSPECT model (Chen & Leblanc, 1997; Jacquemoud & Baret, 1990), a 300 m-resolution global Chl_{leaf} product, was generated using LUT methods from ENVISAT-MERIS data (Croft et al., 2020). However, the generally used red-edge, red and NIR wavelengths are sensitive to both leaf and canopy parameters, such as the leaf area index (LAI) and soil background optical properties. Different combinations of Chl_{leaf}, LAI and other leaf/canopy/soil variables, as well as the solar-observation geometry, can produce the same reflectance, and ill-posed inversion is the prime issue of model-based inversion and limits its accuracy, especially when applied globally (Combal et al., 2003).

Empirical VI-based methods represent efficient and accessible tools to estimate plant structural and biochemical traits. Many VIs have been developed to estimate the chlorophyll content (Croft et al., 2014), and VIs constructed with red-edge wavelengths generally exhibit better performance. VI-based methods achieve high accuracy in estimating the chlorophyll content at the canopy scale (Chl_{canopy}). The red-edge chlorophyll index (CIre) was found to be strongly correlated with Chl_{canopy} in maize and soybean (Gitelson et al., 2005). The inverted red-edge chlorophyll index (IRECI) exhibits the strongest performance in estimating Chl_{canopy} in experiments performed in situ (Frampton et al., 2013). Furthermore, the MERIS terrestrial chlorophyll index (MTCI) was produced as an official MERIS level 2 product (Dash & Curran, 2004). Nevertheless, the VI-based retrieval of Chl_{leaf} is compounded by the information coupling of leaf and canopy scales, such as LAI, leaf angle distribution (LAD), soil background reflectance and vegetation type (Croft et al., 2014; Demarez & Gastellu-Etchegorry, 2000; Viña et al., 2011). A VI that is sensitive only to Chl_{leaf} but resistant to the canopy structure and soil background is thus crucial for retrieving Chl_{leaf}.

For the extraction of the leaf chlorophyll concentration based on canopy reflectance of vegetation, a primary goal is to decouple the effect of canopy structure parameters from the Chl_{leaf} inversion process. Several efforts to reduce the effect of LAI on the VI-based inversion of Chl_{leaf} have been reported, since most VIs are sensitive to LAI. One approach is to combine VIs with different responses to Chl_{leaf} and canopy parameters. The ratio of the transformed chlorophyll in reflectance index (TCARI, sensitive to chlorophyll) to the optimized soil-adjusted vegetation index (OSAVI, sensitive to LAI), TCARI/OSAVI, achieved the highest accuracy among multiple VIs for retrieving potato Chl_{leaf} (Clevers et al., 2017). The ratio of the modified chlorophyll absorption ratio index (MCARI) to OSAVI (MCARI/OSAVI) was strongly correlated with the chlorophyll content of winter wheat (Wu et al., 2008). Another matrix-based VI combination approach was developed to remove the effect of the LAI on the retrieval of Chl_{leaf}, which was less sensitive to LAI than the VI ratio approach but the performance of the VI pair varied across different vegetation types, growth stages and leaf angles (Xu et al., 2019). Therefore, ratio and matrix approaches are still limited to specific vegetation types and regional areas, which limits their application in large-scale Chl_{leaf} product generation.

The objectives of this research are as follows: (1) to develop a chlorophyll sensitivity index (CSI) that is highly sensitive to Chl_{leaf}

and resistant to canopy structure and soil background parameters, (2) to develop a CSI-based Chl_{leaf} estimation method and validate it with ground measurements and compare it with the existing Chl_{leaf} product and (3) to generate a China-wide Chl_{leaf} product using the CSI empirical model from Sentinel-2 MSI data.

2 | MATERIALS AND METHODS

2.1 | Ground measurements

The measured Chl_{leaf} data from different studies were collected for validation. The measured data comprised 308 measurements from nine different sites within four vegetation types: cropland (CRP), deciduous broadleaf forest (DBF), evergreen needleleaf forest (ENF) and grassland (GRA). Information on each experiment is reported in Table 1. Canopy reflectance spectra used in the validation analysis to calculate CSI were measured using a field spectrometer or a satellite sensor. Chl_{leaf} was measured through a laboratory analysis (Lab) or a field spectrometer-based retrieval method (Spec; Uddling et al., 2007). Detailed descriptions of the experiments at some sites are shown in (Supporting Information 1).

2.2 | Satellite data

2.2.1 | Sentinel-2 MSI data

Sentinel-2 MSI images were used to validate the Chl_{leaf} estimate and to generate the Chl_{leaf} product (https://scihub.copernicus.eu/). The European Space Agency Sentinel-2 Earth observation mission consists of two satellites (Sentinel-2A and Sentinel-2B) with a revisit frequency of 5 days. The MSI onboard Sentinel-2 has 13 bands, including three red-edge bands (RE1: central wavelength = 705 nm; RE2: central wavelength = 740 nm; and RE3: central wavelength = 783 nm). The spatial resolution of Sentinel-2 is 10 m for visible and near-infrared (NIR) bands and 20 m for the red-edge bands. Images were downloaded over the Reusel, Borden and Huailai sites at the times nearest the surface measurements to validate the inversion. Sentinel-2 data from 2019 to 2020 were processed on the Google Earth Engine to derive the Chl_{leaf} product across China. The product was resampled to 30 m using the nearest neighbour method to reduce the time needed for downloading and publishing the product, as well as the storage resource.

2.2.2 | ENVISAT MERIS data

Before 2015, ENVISAT MERIS was the main instrument for observing the chlorophyll-sensitive red-edge bands. MERIS sampled the global surface reflectance between 2002 and 2012 at a resolution of 300 m. In this paper, a full-resolution (FR) surface reflectance product with a 300-m, 7-day resolution was used. The FR surface reflectance product provides one red-edge band (central band = 708 nm). MERIS FR data are produced by a series of preprocessing steps, including radiometric, geometric and atmospheric corrections. MERIS FR surface reflectance images were used to calculate Chl_{leaf} and compared with the ground-measured Chl_{leaf} sampled before 2012.

2.2.3 | Land cover map

The Global Land Cover with Fine Classification System product in 2020 (GLC_FC30-2020) (Liu et al., 2020) was used to define the vegetation types and derive the 30-m Chl_{leaf} product in China. Based on the GLC_FCS30-2020 land cover product, plants in China were classified into five types in this study: CRP, broadleaf forest (BF), needleleaf forest (NF), shrubland (SHR) and GRA. An empirical regression relationship between Chl_{leaf} and VIs for each vegetation type was constructed to produce the Chl_{leaf} product across China.

2.3 | PROSAIL and 4-Scale-PROSPECT model simulations

Satellite-derived canopy reflectance was simulated using different radiative transfer models and the spectral response function. The PROSAIL model was utilized to simulate canopy reflectance of CRP and GRA, whose canopies are considered turbid media with homogeneous horizontal layers. For DBF, ENF and SHR canopies, the 4-Scale-PROSPECT model was used, which accounts for the spatial distribution of vegetation groups, crown shape and leaf clumping. Parameters used in the two models are listed in Table 2. The PROSAIL model is also used to carry out the sensitivity analysis of different VIs and the evaluation of uncertainties brought by carotenoids.

2.4 | Construction of the CSI

The responses of each simulated Sentinel-2 MSI band reflectance to changes in Chl_{leaf} and LAI are studied. Normalized reflectance, which is calculated as the band reflectance in different LAI (Chl_{leaf}) divided by its max reflectance, is depicted in Figure 1a,b. Derivatives of normalized reflectance to the two parameters were calculated to quantitatively analyse the response of normalized reflectance to Chl_{leaf} and LAI (Figure 1c,d).

In the construction of the CSI, the RE1 band was selected primarily for its high saturation threshold to Chl_{leaf} (Figure 1a,c) and high sensitivity to Chl_{leaf}. The red-edge normalized difference vegetation index (NDVIre) was calculated first.

$$NDVIre = \frac{\rho_{NIR} - \rho_{RE1}}{\rho_{NIR} + \rho_{RE1}}. \quad (1)$$

As shown in the first colour bars from Figure 2a,c, NDVIre increases with LAI and Chl_{leaf} simultaneously, especially when LAI < 3.5. The

TABLE 1 Ground measurements of Chl_{leaf} and canopy reflectance spectra

Site name	Country	Location	Time (year/month)	Number of samples	Vegetation type	Species	Chl_{leaf} method	Canopy spectral method	Reference
Xiao Tangshan	China	40°11′44′′N, 116°26′3′′E	2002/42004/4	92	CRP	Winter wheat	Lab	ASD spectrometer	Liu et al. (2010)
Nebraska-2, 3	USA	41°9′54′′N, 96°28′12′′W	2002/6–92,004/6–9	73	CRP	Soybean	Lab	Ocean Optics USB2000	Gitelson et al. (2005)
Reusel	Netherlands	51°59′48″N, 5°9′35″ E	2016/6–8	25	CRP	Potato	Spec	Sentinel-2 MSI	Clevers et al. (2017)
Borden	Canada	44°19′12′′N, 79°55′48′′W	2016/5–10	19	DBF	Red maple, White Ash, Bigtooth Aspen	Lab	Sentinel-2 MSI	Croft et al. (2017)
Sudbury Zhang	Canada	47°9′36′′N, 81°44′24′′W	2003–04 summer	18	ENF	Black spruce	Lab	ENVISAT MERIS	Zhang et al. (2008)
Sudbury Simic	Canada	47°10′48′′N, 81°44′24′′W	2007 summer	10	ENF	Black spruce	Lab	ENVISAT MERIS	Simic et al. (2011)
Huailai ENF	China	40°20′54′′N, 115°47′2′′E	2020/10–12	24	ENF	Chinese pine	Lab	Sentinel-2 MSI	
Huailai DBF	China	40°21′18′′N, 115°47′45′′E	2020/10	19	DBF	Aspen	Lab	Sentinel-2 MSI	
Huailai GRA	China	40°21′56′′N, 115°48′21′′E	2021/6–8	28	GRA		Lab	Sentinel-2 MSI	

TABLE 2 Input parameters of the PROSAIL and the 4-Scale-PROSPECT model to obtain the VI-Chl_{leaf} regression equations

	PROSAIL		4-Scale+PROSPECT		
Parameter	**CRP**	**GRA**	**DBF/EBF**	**DNF/ENF**	**SHR**
Chlorophyll a+b content (μg/cm^2)	5–70, step:5	5–70, step:5	5–100, step:5	5–100, step:5	5–100, step:5
Carotenoid content, Car (μg/cm^2)	8–16, step:4	4–12, step:4	Chl_{leaf}/7	Chl_{leaf}/7	Chl_{leaf}/7
Brown pigment C_{brown}	0	0	0	0	0
Equivalent water thickness, C_w (g/cm^2)	0.01	0.01	0.01	0.01	0.01
Dry matter content, C_m (g/cm^2)	0.009	0.005	0.005	0.05	0.005
Leaf structure parameter, *N*	1–2	1–1.5	DBF:1.2, EBF:1.8	2.8	1.8
Leaf area index, LAI (m^2/m^2)	0.5–6.5, step:0.5	0.5–6.5, step:0.5	0.5–8 step:0.5	0.5–8, step:0.5	0.5–8, step:0.5
Solar zenith angle, θ_s (°)	30–70, step:10	30–70, step:10	30–70, step:10	30–70, step:10	30–70, step:10
View zenith angle, θ_v (°)	0 (nadir)	0 (nadir)	0 (nadir)	0 (nadir)	0 (nadir)
Leaf angle distribution, LAD	5 types	5 types	–	–	–
Soil moisture, psoil	0–1	0–1	–	–	–
Hotspot parameter, S_L	0.1	0.1	–	–	–
Stick height (m)	–	–	10	10	3
Crown height (m)	–	–	8	10	7
Crown shape	–	–	Spheroid	Cone and cylinder	Spheroid
Tree density (trees/ha)	–	–	1400	3000	1000
Crown radius	–	–	1.25	1.00	1.25
Neyman grouping	–	–	3	4	3
Clumping index	–	–	0.9	0.8	0.8
Needle to shoot ratio	–	–	1	1.4	1
Foliage element width	–	–	0.15	0.1	0.15
Background composition	–	–	Green vegetation and soil	Green vegetation and soil	Dry grass and soil

strategy to reduce the sensitivity to LAI is utilizing the different response characteristics for bands shown in Figure 1: (1) the decrease in the blue-band reflectance (ρ_{blue}) with increasing LAI is more dramatic than that with increasing Chl_{leaf}, and (2) ρ_{RE1} decreases more sharply than ρ_{blue} with increasing Chl_{leaf}, but (3) the decreases in ρ_{RE1} and ρ_{blue} with LAI are more similar. These characteristics cause the calculator ρ_{blue} / ρ_{RE1} to increase with Chl_{leaf} but decrease with LAI (second colour bar in Figure 2a,c). Additionally, due to the saturation resistance of RE1, the calculator ρ_{blue} / ρ_{RE1} has a wider Chl_{leaf} sensitivity range. The chlorophyll sensitive index (CSI) was constructed as follows:

$$\text{Chlorophyll Sensitive Index (CSI)} = K \times \frac{\rho_{NIR} - \rho_{RE1}}{\rho_{NIR} + \rho_{RE1}} \times \frac{\rho_{blue}}{\rho_{RE1}}. \quad (2)$$

In the simulation, the NDVIre ranges from 0 to 1, and the value domain of ρ_{RE1} (0.05–0.35) is two to three times ρ_{blue} (<0.15) with increasing LAI and Chl_{leaf}. Thus, a gain factor $K = 2.5$ is used to adjust the range of the index mainly between 0 and 1 in vegetative areas. Figure 2a illustrates that the normalized response of CSI to LAI ranges from 0.7274 to 1, which is much smaller than that of NDVIre (0.3522–1) and ρ_{blue} / ρ_{RE1} (0.4348–1). The two parts of the CSI also have an opposite response to the increased soil moisture (psoil; Figure 2b). Regarding the response to Chl_{leaf}, the CSI has a larger range than NDVIre and ρ_{blue} / ρ_{RE1}. However, new uncertainties may arise from the blue band. The blue band is sensitive to carotenoid and atmospheric conditions, which limits the applicability of the blue-band-based VI. Thus, the effects of carotenoids and atmospheric conditions on the CSI are analysed quantitatively in Section 3.5.

The sensitivity of the CSI to Chl_{leaf} and other variables was compared with some existing red-edge indices (Table 3) in the present study.

2.5 | VI-based Chl_{leaf} inversion and product

Table 4 describes the physical model-simulated vegetation type-specific regression equations with the fitting accuracy (coefficient of determination (R^2) and root mean square error (RMSE)) for the Sentinel-2 MSI band. The regression models are validated in Section 3.2, and the CSI-based regression model is applied to Sentinel-2 data in Section 3.4.

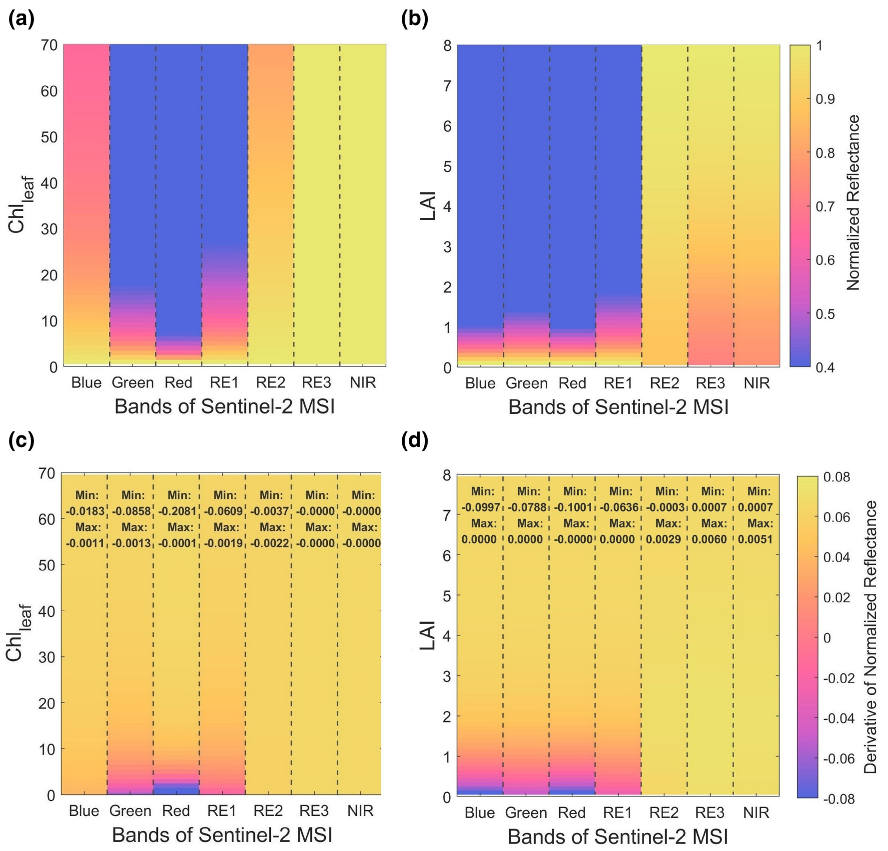


FIGURE 1 PROSAIL-simulated reflectance for different bands with increasing Chl_{leaf} (a) and LAI (b) and the variations in the derivative of the normalized reflectance with increasing Chl_{leaf} (c) and LAI (d)

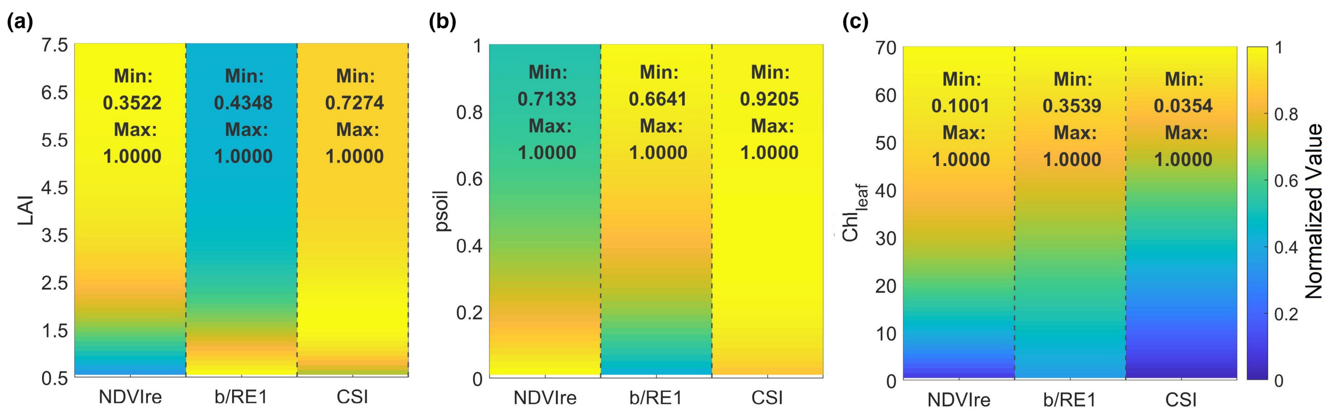


FIGURE 2 Variations in NDVIre, the calculator ρ_B / ρ_{RE1} and CSI with increasing LAI (a), soil moisture (psoil) (b) and Chl_{leaf} (c). Before the comparison, the values of each index or calculator were normalized by dividing their values by their maximum value.

2.6 | Sensitivity analysis of different VIs

Three indicators were used in the sensitivity analysis. The first is the variable coefficient (CV) to express the sensitivity of a particular VI to a parameter v (v is LAI or Chl_{leaf}):

$$CV_v = \frac{\sigma_v}{\mu_v}, \tag{3}$$

where σ_v is the standard deviation and μ_v is the mean value. A higher CV_v indicates that the VI is more sensitive to the change in v. The $CV_{Chl_{leaf}} / CV_{LAI}$ ratio is calculated to evaluate the extent to which a VI is sensitive to Chl_{leaf} and insensitive to LAI. A higher $CV_{Chl_{leaf}} / CV_{LAI}$ indicates a stronger ability to capture the Chl_{leaf} variation and to remain constant when LAI varies. The second indicator is the R^2 value of the linear regression equations between VIs and Chl_{leaf}. The best linear regression equations for each VI and Chl_{leaf} were acquired, and then the linear R^2 was calculated. The third indicator is the saturation point for Chl_{leaf} (SP), which is defined as the starting point for the region where the VI does not respond to the increasing Chl_{leaf}. VIs were normalized first to compare VIs with different value ranges; then, the SP is calculated as the point where the absolute value of the first derivative to Chl_{leaf} is less than a defined threshold (0.005 in this paper). Reasons for the selection of these indicators are shown in (Supporting Information 2).

TABLE 3 Existing VIs proposed in the literature. The specific wavelength constructing an index was transformed to the closest band of the Sentinel-2 MSI. In the formula, the blue (B), green (G), red (R), near-infrared (NIR), red-edge band 1 (RE1, 690–730 m), red-edge band 2 (RE2, 730–770 nm) and red-edge band 3 (RE3, 770–790 nm) denote the bands in Sentinel-2 MSI

Index	Formula	References
Red-edge normalized difference vegetation index (NDVIre)	$\left(\mathrm{NIR} - \frac{\mathrm{RE1}}{\mathrm{NIR}+\mathrm{RE1}}\right)$	Gitelson and Merzlyak (1994)
Red-edge ratio normalized difference vegetation index (RERNDVI)	$\frac{\mathrm{NIR}-\mathrm{R}}{\mathrm{NIR}+\mathrm{R}} * \sqrt{\mathrm{RE2}/\mathrm{RE1}}$	Chang and Shoshany (2016)
Red-edge chlorophyll index (CIre)	$\mathrm{NIR}/\mathrm{RE1} - 1$	Gitelson et al. (2005)
Novel inverted red-edge chlorophyll index (IRECI)	$(\mathrm{NIR}-\mathrm{R})/(\mathrm{RE1}/\mathrm{RE2})$	Frampton et al. (2013)
Modified chlorophyll absorption ratio index (MCARI)	$[(\mathrm{RE2}-\mathrm{RE1}) - 0.2(\mathrm{RE2}-\mathrm{R}) * \mathrm{RE2}/\mathrm{RE1}$	Daughtry et al. (2000)
MERIS terrestrial chlorophyll index (MTCI)	$(\mathrm{NIR}-\mathrm{RE1})/(\mathrm{RE1}-\mathrm{R})$	Dash and Curran (2004)
Transformed chlorophyll absorption in reflectance index (TCARI)	$3\left[(\mathrm{RE1}-\mathrm{R}) - 0.2(\mathrm{RE1}-\mathrm{G})(\mathrm{RE1}/\mathrm{R})\right]$	Haboudane et al. (2002)
Maccioni 2001 (Macc01)	$(\mathrm{RE3}-\mathrm{RE1})/(\mathrm{RE3}-\mathrm{R})$	Maccioni et al. (2001)
Modified normalized difference (MND)	$(\mathrm{RE2}-\mathrm{B})/(\mathrm{RE2}+\mathrm{RE1}-2\mathrm{B})$	Sims and Gamon (2002)
Datt 99 (Datt99)	$(\mathrm{NIR}-\mathrm{RE1})/(\mathrm{NIR}-\mathrm{R})$	Datt (1999)
TCARI/optimized soil-adjusted vegetation index (TCARI/OSAVI)	$\frac{3[(\mathrm{RE1}-\mathrm{R}) - 0.2(\mathrm{RE1}-\mathrm{G})(\mathrm{RE1}/\mathrm{R})]}{(1+0.16)(\mathrm{NIR}-\mathrm{R})/(\mathrm{NIR}+\mathrm{R}+0.16)}$	Wu et al. (2008)

TABLE 4 Regression equations between Chl_{leaf} and VIs for Sentinel-2 MSI images. y: Chl_{leaf}, x: simulated VI. The numbers in the parentheses represent the R^2 and RMSE (unit: $\mu g\ cm^{-2}$). The bold values represent the three highest R^2 or the lowest RMSE for each type and the underlined values represent the higest R^2 or the lowest RMSE

	CRP	DBF, EBF	DNF, ENF	GRA
NDVIre	$y = 73.46x - 3.19$ (0.59, 13.66)	$y = 121.57x - 15.28$ (0.55, 17.60)	$y = 92.64x + 8.12$ (0.36, 21.08)	$y = 81.31x - 10.74$ (0.61, 12.63)
RERNDVI	$y = 21.01x + 4.66$ (0.43, 15.97)	$y = 32.71x + 7.54$ (0.32, 21.64)	$y = 22.82x + 26.11$ (0.18, 23.84)	$y = 22.68x - 0.63$ (0.48, 14.57)
CIre	$y = 4.89x + 27.75$ (0.56, 16.87)	$y = 10.17x + 16.14$ (0.66, 13.73)	$y = 10.99x + 22.74$ (0.54, 15.82)	$y = 5.29x + 15.25$ (0.62, 12.42)
IRECI	$y = 16.06x + 16.53$ (0.39, 16.59)	$y = 14.43x + 36.13$ (0.17, 24.00)	$y = 25.62x + 39.39$ (0.12, 24.66)	$y = 12.22x + 16.04$ (0.43, 15.21)
MCARI	$y = 33.07x + 17.78$ (0.44, 15.90)	$y = 33.76x + 35.40$ (0.20, 23.42)	$y = 58.95x + 39.14$ (0.15, 24.19)	$y = 25.74x + 16.52$ (0.48, 14.53)
MTCI	$y = 4.97x + 10.75$ (**0.73**, 11.16)	$y = 5.73x + 13.87$ (**0.86**, **8.86**)	$y = 7.83x + 11.81$ (0.80, **10.30**)	$y = 4.14x + 12.41$ (0.76, 9.83)
MND	$y = 163.69x - 90.80$ (0.69, 11.75)	$y = 265.68x - 161.05$ (0.69, 14.54)	$y = 212.51x - 110.92$ (0.48, 18.97)	$y = 183.09x - 109.94$ (0.72, 10.69)
Macc01	$y = 117.25x - 52.09$ (**0.80**, **9.59**)	$y = 216.36x - 123.42$ (0.82, 11.08)	$y = 221.86x - 120.85$ (0.79, 12.03)	$y = 131.00x - 65.90$ (**0.77**, **9.64**)
TCARI/OSAVI	$y = 9.98/(0.12 + x)$ (0.51, 12.04)	$y = 9.66/(0.04 + x)$ (**0.86**, **8.90**)	$y = 7.57/(0.02 + x)$ (**0.83**, **9.55**)	$y = 7.66/(0.06 + x)$ (0.72, 9.96)
Datt99	$y = 122.39x - 56.84$ (<u>0.81</u>, **9.33**)	$y = 226.79x - 133.20$ (0.83, 10.69)	$y = 237.76x - 135.66$ (**0.81**, 11,47)	$y = 135.85x - 70.26$ (**0.79**, **9.33**)
CSI	$y = 76.92x + 2.00$ (0.64, <u>8.28</u>)	$y = 99.31x - 9.78$ (<u>0.93</u>, <u>6.04</u>)	$y = 121.99x - 15.97$ (<u>0.93</u>, <u>6.18</u>)	$y = 89.18x + 0.03$ (<u>0.99</u>, <u>6.61</u>)

3 | RESULTS

3.1 | Sensitivity analysis of different VIs using simulated data

3.1.1 | Sensitivity of VIs to Chl_{leaf} and LAI

The sensitivity of CSI and some existing red-edge chlorophyll indices to variations in LAI and Chl_{leaf} were examined using PROSAIL-simulated canopy reflectance (Table 5). The CV_Chl_{leaf} values are the highest for TCARI/OSAVI, MTCI, CIre and CSI (>50%), indicating high sensitivity to chlorophyll. The CV_LAI values for TCARI/OSAVI, MTCI and CIre (>20%) are higher than those for the CSI, which indicates that they are also sensitive to LAI. The CV_Chl_{leaf} /CV_LAI suggests the ability of a VI to decouple Chl_{leaf} from LAI. The CV_Chl_{leaf} /CV_LAI of the CSI (3.43) and Datt99 (3.22) is larger than those of TCARI/OSAVI (3.08), Macc01 (2.88), MTCI (2.63) and other VIs (<2), suggesting that the CSI and Datt99 are more capable of

TABLE 5 Sensitivity and linearity results. The bold numbers indicate the two best-performing indices, and the underline numbers indicate the worst. In the simulation, except for the LAI and Chl_{leaf}, the other variables were set to constants (carotenoid (Car) = 8 μg cm^{-2}, N = 1.5, LAD is the spherical distribution, solar zenith angle = 10°, the other variables are set to constants as described in Table 2)

	NDVIre	RERNDVI	CIre	IRECI	MCARI	MTCI	MND	Macc01	Datt99	TCARI/OSAVI	CSI
CV_Chl$_{leaf}$	27.96%	26.62%	54.19%	38.56%	44.26%	59.46%	11.05%	17.79%	17.14%	**111.52%**	50.16%
CV_LAI	24.37%	31.93%	39.18%	40.23%	43.39%	22.64%	7.11%	**6.18%**	**5.32%**	36.18%	14.62%
CV_Chl$_{leaf}$/CV_LAI	1.15	0.83	1.38	0.96	1.02	2.63	1.55	2.88	**3.22**	3.08	**3.43**
SP (μg cm^{-2})	42.97	46.61	69.94	66.87	64.00	**88.88**	23.56	27.58	26.03	39.23	**89.41**
R^2	0.82	0.85	**0.99**	0.94	0.94	**0.99**	0.83	0.77	0.77	0.58	**0.99**

sensing changes in Chl_{leaf} in the presence of different LAI values. The SPs of the regression models for Datt99, MND, Macc01 and TCARI/OSAVI with values less than 40 μg cm^{-2} suggest severe saturation problems when estimate Chl_{leaf}. The higher SP for the CSI indicates that the saturation problem is significantly mitigated. In Table 5, the CSI also shows the highest linear R^2 values (0.99) among the 11 indices.

3.1.2 | Sensitivity of VIs to the soil moisture content and LAD

As shown in Figure 3, when the soil changes from wet (psoil = 0) to dry (psoil = 1), with the exception of TCARI/OSAVI, the other five well-performing VIs in Section 3.1.1 fluctuate by approximately 5 %, indicating these indices perform well at removing the effect of the soil moisture content. Figure 3b shows the sensitivity of individual VIs to different LAD types. For sparse vegetation (Figure 3b (1), (2)), all six indices fluctuate within 7% for all LAD patterns. As LAI and Chl_{leaf} increased, Macc01, MND, Datt99 and CSI still changed little with LAD (Figure 3b (3), b(4)). However, TCARI/OSAVI and MTCI became more sensitive to LAD as the canopy became dense (Figure 3b (4)), and the TCARI/OSAVI and MTCI values fluctuated by approximately 19.62% and 9.38% for different LADs respectively.

3.1.3 | Sensitivity of VIs to leaf-scale parameters

In Figure 4a, the CSI decreases by approximately 20% when Car increased from 4 to 16 μg cm^{-2}. The five existing VIs in Section 3.1.2 do not fluctuate substantially with Car. In Figure 4b, Macc01, MND and Datt99 are less sensitive to N. The three indices decrease by less than 5% of their maximum values with changes in N, followed by the CSI at less than 20%. MTCI and TCARI/OSAVI show the strongest sensitivity to N, and fluctuate by approximately 40% of its maximum value when N changed from 1 to 3. The CSI shows strong and similar sensitivity to Car and N. Based on this result, the CSI displays weaker sensitivity to canopy-scale parameters but a stronger response to leaf-scale parameters. The effects of Car and N on the CSI-based method will be discussed in Section 4.

3.1.4 | Sensitivity of CSI-estimated Chl_{leaf} to leaf/canopy/background parameters

Figure 5 shows that the changes in the soil background (psoil) and solar-observation geometry (solar zenith angle, SZA and view zenith angle, VZA) induce little difference in the CSI-estimated Chl_{leaf}. The change in the canopy structural parameters (LAI) also slightly influences the estimated Chl_{leaf}: the differences fluctuated between −1.80 μg cm^{-2} and +2.44 μg cm^{-2}. The effects of the leaf biochemical

TABLE 6 Value ranges of different parameters in the sensitivity analysis

	Chl_{leaf} ($\mu g\,cm^{-2}$)	C_{ar} ($\mu g\,cm^{-2}$)	*N*	LAI ($m^2\,m^{-2}$)	Psoil	SZA (°)	VZA (°)
Value range	10–70	4–14	1–3	0.5–5.5	0–1	0–60	0–20
Reference value	35	8	1.5	3	0.5	30	10

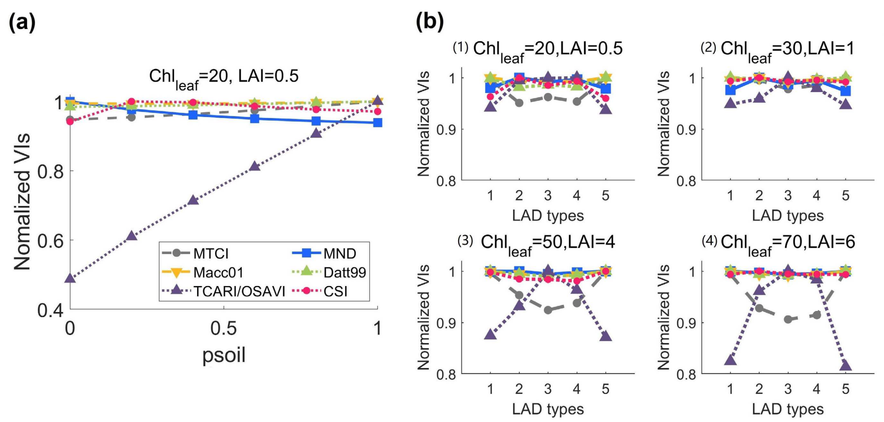


FIGURE 3 The normalized values of six VIs were simulated using the PROSAIL model under different soil moisture (psoil) (a) and LAD (b). In (b), LAD types 1–5 represent planophile, plagiophile, extremophile, uniform and spherical distributions of the leaf angle.

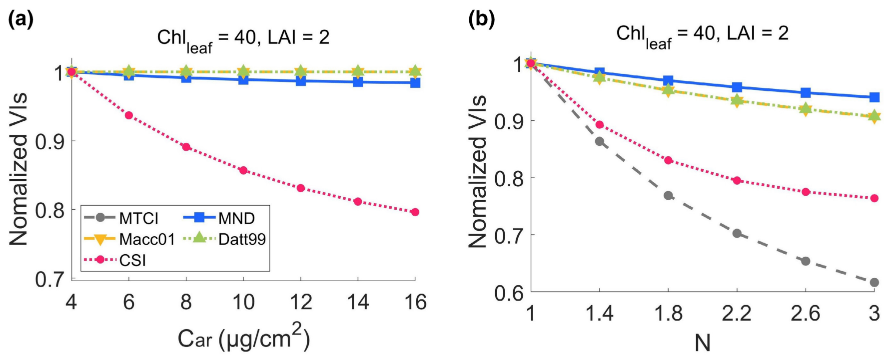


FIGURE 4 The normalized values of six VIs simulated using the PROSAIL model under different Car (a) and leaf structure parameter (N) (b) scenarios.

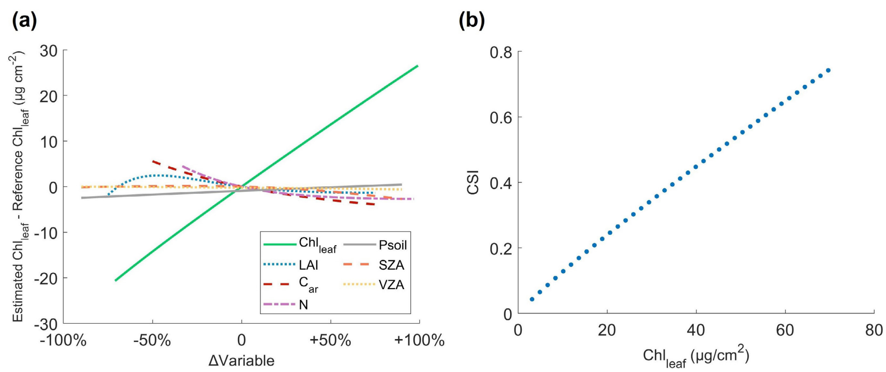


FIGURE 5 Sensitivity of the CSI-estimated Chl_{leaf} to variations in different parameters. The *X*-axis represents the differences in a variable from its reference values. The parameter settings are listed in Table 6.

(Car) and structural parameters (N) on the CSI-estimated Chl_{leaf} are slightly larger than those on the LAI. The variance decreases from +5.57 to $-3.74\,\mu g\,cm^{-2}$ when Car increases from $4\,\mu g\,cm^{-2}$ to $16\,\mu g\,cm^{-2}$ and decreases from +4.48 to $-2.70\,\mu g\,cm^{-2}$ when N changes from 1 to 3. Chl_{leaf} exerts the greatest effect on the CSI-estimated results. A −70% or +100% change in Chl_{leaf} from its reference value causes the estimated Chl_{leaf} to decrease or increase by more than $20\,\mu g\,cm^{-2}$. Therefore, CSI is highly sensitive to Chl_{leaf} compared with the other parameters.

3.2 | Validation of the VI-based Chl_{leaf} inversion approach for different vegetation types

In Table 7 and Figure 6, CSI, MTCI, MND, Macc01 and Datt99 perform better than the other indices with lower RMSE and relative RMSE (rRMSE) values and higher R^2 values. The CSI has the highest accuracy (RMSE = $9.39\,\mu g\,cm^{-2}$, R^2 = 0.49) for all four vegetation types. MTCI (RMSE = $13.00\,\mu g\,cm^{-2}$, R^2 = 0.19) and Macc01 (RMSE = $13.76\,\mu g\,cm^{-2}$, R^2 = 0.23) have the next highest accuracies.

TABLE 7 Accuracy of different VIs in estimating Chl_{leaf}. The bold numbers indicate the five best-performing indices and the underline number indicates the best performing index.

	NDVIre	RERNDVI	CIre	IRECI	MCARI	MTCI	MND	Macc01	Datt99	TCARI/OSAVI	CSI
R^2	0.05	0.00	0.07	0.00	0.00	**0.19**	**0.19**	**0.23**	**0.14**	0.00	0.49
RMSE ($\mu g\,cm^{-2}$)	16.05	17.31	16.31	17.65	17.48	**13.00**	**14.21**	**13.76**	**14.31**	20.43	9.39
rRMSE	40.75%	43.94%	41.39%	44.81%	44.36%	**33.00%**	**36.06%**	**34.92%**	**36.33%**	51.86%	23.83%
Bias ($\mu g\,cm^{-2}$)	3.39	0.38	4.91	1.92	**1.85**	**−1.17**	2.34	4.41	6.04	**1.40**	**−0.58**

Table 7 illustrates the validation results of each VI for individual plant types. For CRP, the Chl_{leaf} values modelled by the CSI (RMSE = 9.51 μg cm^{-2}, rRMSE = 21.98% and R^2 = 0.40), Macc01 (RMSE = 10.49 μg cm^{-2}, rRMSE = 24.24%, R^2 = 0.29) and Datt99 (RMSE = 10.68 μg cm^{-2}, rRMSE = 24.68%, R^2 = 0.27) show higher accuracy. However, the Chl_{leaf} estimated by Macc01 and Datt99 changes little when Chl_{leaf} > 40 μg cm^{-2}, indicating the saturation problem (Figure 6b,d). For DBF, the accuracy of the CSI-based inversion (RMSE = 7.04 μg cm^{-2}, rRMSE = 25.84% and R^2 = 0.7) is significantly higher than that of the other VIs (RMSE ≥15.00 μg cm^{-2}, rRMSE ≥55.06%). Existing indices tend to overestimate the Chl_{leaf} of ENF, with biases larger than 5 μg cm^{-2}. CIre (RMSE = 11.08 μg cm^{-2}, bias = 8.79 μg cm^{-2}) and MTCI (RMSE = 12.49 μg cm^{-2}, bias = 7.02 μg cm^{-2}) perform better than the other currently used indices. The CSI has a lower RMSE (9.52 μg cm^{-2}) and bias (−1.22 μg cm^{-2}) than all the present VIs. For GRA, the CSI also has the lowest RMSE (11.01 μg cm^{-2}), and bias (−0.29 μg cm^{-2}). The accuracies of MND (RMSE = 14.07 μg cm^{-2}, R^2 = 0.02) and Datt99 (RMSE = 14.31 μg cm^{-2}, R^2 = 0.03) are next after the CSI, but the underestimation is significant (bias = −13.65 μg cm^{-2}, −10.09 μg cm^{-2}), especially when Chl_{leaf} > 40 μg cm^{-2}.

3.3 | Validation of the VI-based models under different LAI and Chl_{leaf} conditions

The effects of LAI and Chl_{leaf} on the different VI regression methods were analysed using 127 ground measurements from in situ experiments in Xiaotangshan and Nebraska (Figures 7 and 8). As shown in Figure 8a, the RMSE of the CSI-based estimate remains relatively low and stable in 6.13–10.19 μg cm^{-2} with increasing LAI. The RMSE of TCARI/OSAVI-based estimations decreases from 12.82 μg cm^{-2} (LAI = 3.15) to 6.30 μg cm^{-2} (LAI = 4.95). For the other VIs, RMSE increases significantly as LAI increases. Figure 8b illustrates that the accuracy of the CSI and MTCI methods remains stable under different Chl_{leaf} conditions. The TCARI/OSAVI method has a high RMSE when Chl_{leaf} is large. For the other VIs, RMSE tends to decrease with increasing Chl_{leaf}. The CSI-based method is capable of maintaining stable accuracy for different LAI and Chl_{leaf} values and is almost impervious to the LAI and Chl_{leaf} conditions.

3.4 | Spatial and temporal trends of the CSI-estimated Chl_{leaf}

The Chl_{leaf} distribution with a resolution of 30 m across China on 8–28 August 2020 is presented in Figure 9a using CSI-based inversion approach (Table 4). Figure 9b illustrates the seasonal phenologies for six different plant types at specific sites. As shown in Figure 9b, the Chl_{leaf} values of DBF, GRA, CRP and SHR exhibit strong seasonal phenologies, increasing in spring and remaining high in summer. In

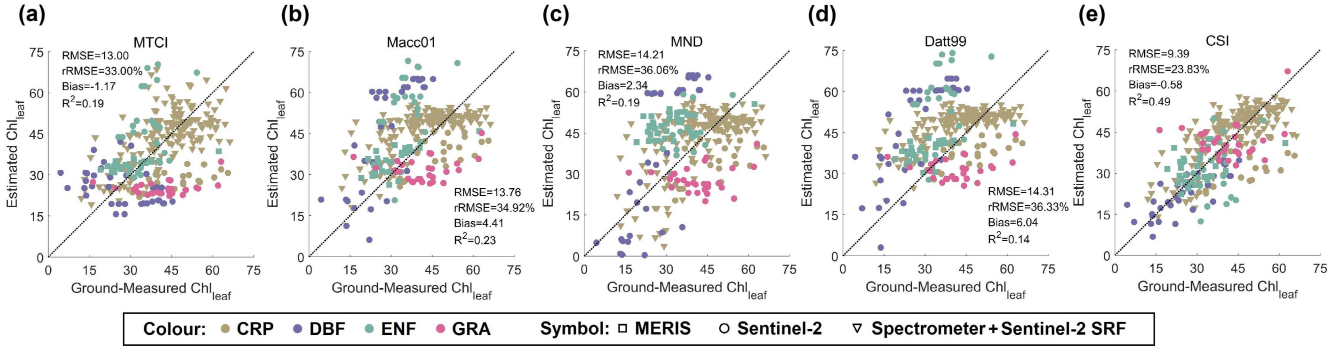


FIGURE 6 Relationships between the measured Chl_{leaf} and VI-derived Chl_{leaf} ((a–e) represent MTCI, Macc01, MND, Datt99, CSI, respectively). Regression models are shown in Table 4 and Table S1 (in the Supporting Information).

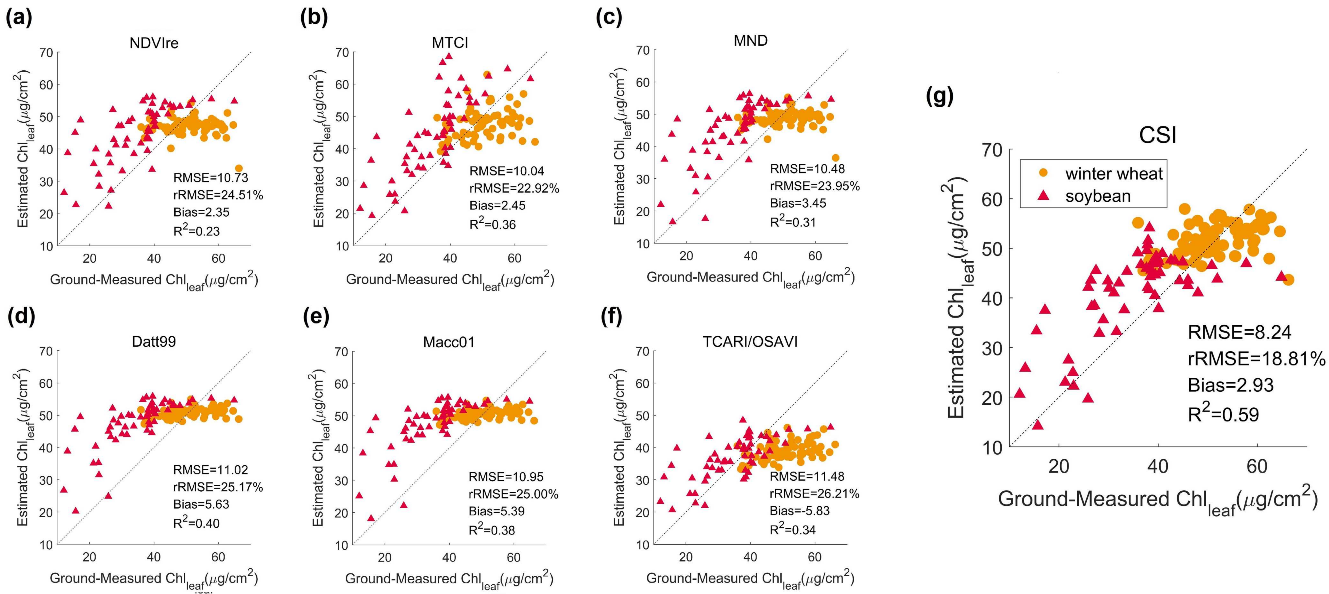
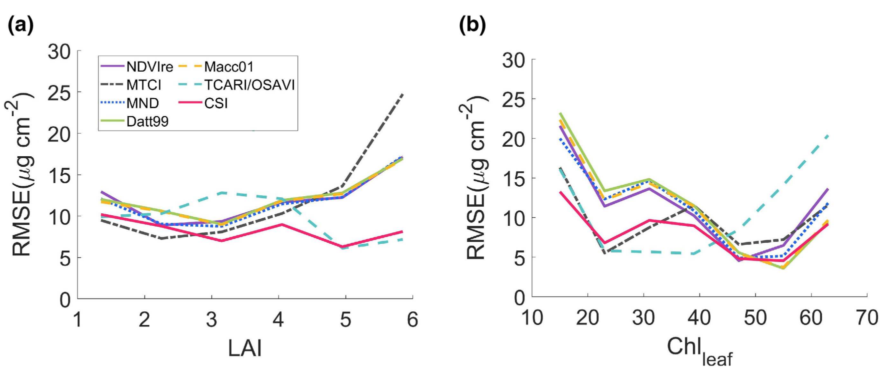


FIGURE 7 Validation of different methods for determining Chl_{leaf} using soybean and winter wheat samples. (a–g) represent the results using NDVIre, MTCI, MND, Datt99, Macc01, TCARI/OSAVI, CSI regression methods.

FIGURE 8 RMSE of estimated Chl_{leaf} under different LAI (a) and Chl_{leaf} (b) conditions.



autumn, Chl_{leaf} decreases sharply to less than 10 $\mu g\,cm^{-2}$. EBF and ENF also show the temporal variance in Chl_{leaf}, but their minimum values exceed 20 $\mu g\,cm^{-2}$ in winter. The highest Chl_{leaf} of CRP and EBF reaches more than 70 $\mu g\,cm^{-2}$, followed by the values of DBF, whose maximum values are approximately 60 $\mu g\,cm^{-2}$. The CSI-based Chl_{leaf} product is compared with the RTM-based product (Croft et al., 2020) and other indices retrieved Chl_{leaf} in Supporting Information 4 (Figures S2–S4).

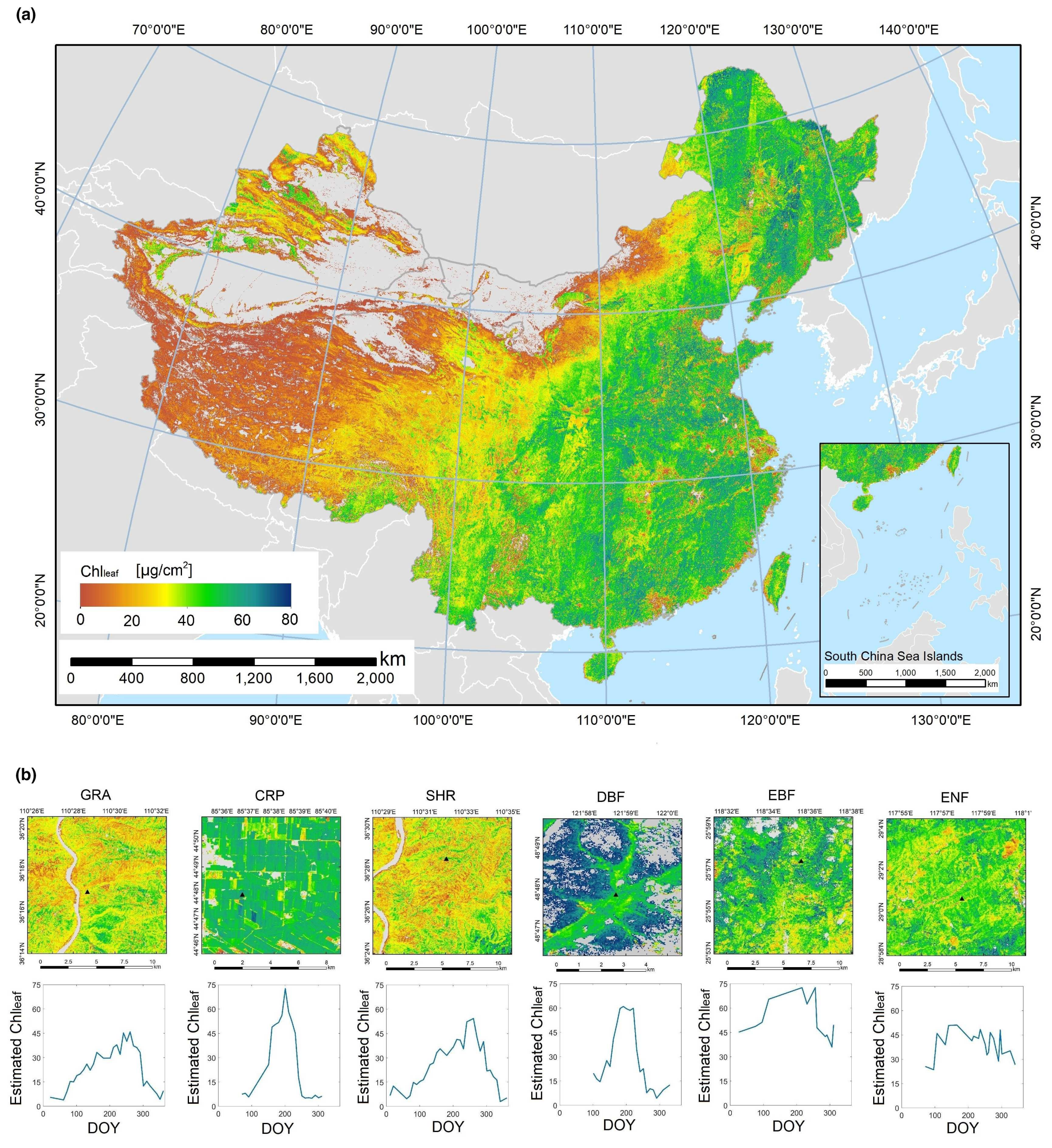


FIGURE 9 (a) Map of Chl_{leaf} across China. In addition to the regression models in Table 4, the Chl_{leaf} of SHR was calculated using $Chl_{leaf} = 130.34 * CSI - 25.37$ (RMSE = 10.21 μg cm^{-2}, $R^2 = 0.88$). (b) Chl_{leaf} product images on July 2019 and the seasonal variation for each type.

3.5 | Uncertainties generated by the blue band in CSI-based methods

3.5.1 | Effects of carotenoids

The Car content exhibits strong absorption in the blue band; thus, CSI displays higher sensitivity to Car, as shown in Figure 4. The effect of the Car content on CSI-estimated Chl_{leaf} differs with the time when the ratio Car/Chl_{leaf} (Table 9) changes. In summer, when Car/Chl_{leaf} ranges from 0.15 to 0.3, the CSI-estimated Chl_{leaf} only changes by 4.27 μg cm^{-2} with an absolute error (AE) less than 5 μg cm^{-2}. In autumn and spring, the variance in the retrieved Chl_{leaf} caused by Car increases: AE reaches 7.59 (autumn) and 6.29 μg cm^{-2} (spring). In winter, the Car/Chl_{leaf} ranges from 0.4 to 3.0, and the AE

TABLE 8 Accuracy of different VIs in estimating Chl_{leaf} for each vegetation type. The bold numbers indicate the best three performance indices and the underline number indicates the best index.

	CRP				DBF				EBF				GRA			
	RMSE	rRMSE	R^2	Bias	RMSE	rRMSE	R^2	Bias	RMSE	rRMSE	R^2	Bias	RMSE	rRMSE	R^2	Bias
NDVIre	11.35	26.23%	0.23	−1.28	25.72	94.42%	**0.67**	22.98	18.62	56.52%	0.30	17.38	20.47	49.22%	0.00	−17.47
RERNDVI	13.19	30.48%	0.16	−6.01	26.28	96.47%	**0.67**	24.04	19.06	57.85%	0.04	17.32	22.46	54.01%	0.00	−19.83
CIre	11.79	27.25%	0.15	2.93	31.15	114.35%	**0.67**	26.56	**11.08**	**33.63%**	0.28	**8.79**	20.88	50.21%	0.00	−18.28
IRECI	15.28	35.31%	0.05	−3.32	22.38	82.16%	0.66	21.44	19.52	59.25%	0.05	17.91	21.35	51.34%	0.00	−18.76
MCARI	15.15	35.01%	0.06	−3.12	22.73	83.44%	0.64	21.68	18.47	56.06%	0.08	16.82	21.59	51.92%	0.01	−19.11
MTCI	11.49	26.55%	0.23	**−1.05**	**15.00**	**55.06%**	0.16	−1.59	**12.49**	**37.91%**	0.20	**7.02**	19.19	46.14%	0.06	−16.53
MND	11.09	25.63%	0.28	**−0.16**	20.90	76.72%	0.64	**10.47**	16.36	49.66%	0.09	14.13	**14.07**	**33.83%**	0.02	−13.65
Macc01	**10.49**	**24.24%**	**0.29**	3.24	22.78	83.62%	**0.69**	12.96	15.26	46.32%	**0.26**	10.09	17.22	41.41%	**0.04**	**−9.79**
TCARI/OSAVI	12.63	29.19%	0.22	−7.22	**20.79**	**76.32%**	0.10	12.41	37.50	113.82%	0.05	22.45	17.64	42.42%	0.01	**5.94**
Datt99	**10.68**	**24.68%**	**0.27**	3.62	22.68	83.26%	0.59	19.89	17.63	53.51%	0.23	13.44	**14.31**	**34.41%**	**0.03**	−10.09
CSI	9.51	21.98%	0.40	0.11	7.04	25.84%	0.70	**−3.39**	9.52	28.89%	0.07	−1.22	11.01	26.47%	**0.03**	−0.29

of the CSI-estimated Chl_{leaf} reaches 6.58 μg cm^{-2} ($Car/Chl_{leaf} = 0.4$) and −1.63 μg cm^{-2} ($Car/Chl_{leaf} = 3.0$). The effect of the Car on the CSI-estimated Chl_{leaf} is slight, with a tendency of overestimation in summer, and becomes more obvious in spring, autumn and winter, with a high probability of overestimation.

3.5.2 | Effects of the atmosphere

Table 10 shows that the ground-measured blue-band and RE1-band reflectance reported in a previous study (Sola et al., 2018) decrease sharply as canopies become denser, but the NIR band increases slightly. According to previous research (Sola et al., 2018), the overestimations are obvious for the blue band of the Sentinel-2 atmospheric uncorrected Level-1C (L1C) reflectance product. In different canopies, the relative error (RE) of the blue band reaches more than 90% compared with the ground measurements. The RE of the other two bands in the L1C data is lower, ranging from −12.5% to 16.56%. The atmospheric effects on the CSI calculated based on L1C reflectance are substantial, with an overestimation ranging from 59.72% to 295.59% under different vegetation conditions. For the atmospheric corrected Level-2A (L2A) product, which was generated using the SEN2COR atmosphere correction processor (Main-Knorn et al., 2017), the error of the blue band decreases considerably from >90% to 18.7%–20.3%, although the error of the other band increases slightly. The atmospheric effects on the CSI calculated using L2A data are substantially reduced to −11.48% to −3.96%, suggesting that the atmospheric effects tend to make CSI underestimate Chl_{leaf}, especially under high vegetation coverage conditions. Precise atmospheric correction is essential for Chl_{leaf} estimation with CSI-based methods.

4 | DISCUSSION

The CSI is an index slightly affected by LAI, LAD and the soil moisture content (Table 5, Figures 3 and 5). Uncertainties arising from LAI in Chl_{leaf} estimates can be significantly reduced using CSI. The reduced dependency on psoil also ensures that the Chl_{leaf} estimation accuracy remains high and stable under low LAI conditions (Figure 8a). The decreased accuracy at the beginning or end of the growing season due to interference by the soil can be avoided using the CSI-based algorithm (Figure 8). Due to the high insensitivity to canopy structure and soil background, the accuracies of CSI-estimated Chl_{leaf} improve more significantly for forest samples whose scenarios are more complicated (Table 8). The RMSE of the CSI estimates for the DBF samples decreases by at least 7.96 μg cm^{-2} compared with the other existing indices.

The CSI is an index showing the cross-type ability. The effect of Chl_{leaf} on the CSI is much higher than the effects of the canopy structure, solar-observation geometry, background, leaf structure and biochemistry (Figure 5). Therefore, the effects of vegetation type characterized by these factors on Chl_{leaf} inversion are

Chl_{leaf} ref. ($\mu g\,cm^{-2}$)	Spring: 25		Summer: 55		Autumn: 15		Winter: 5	
Car/Chl_{leaf} ref.	0.15	0.50	0.15	0.30	0.20	0.60	0.40	3.00
Chl_{leaf} inv. ($\mu g\,cm^{-2}$)	31.29	21.89	54.79	50.52	22.59	14.23	11.58	3.37
ΔChl_{leaf} inv. ($\mu g\,cm^{-2}$)	9.40		4.27		8.36		8.21	
AE ($\mu g\,cm^{-2}$)	+6.29	−3.11	−0.21	−4.48	+7.59	−0.77	+6.58	−1.63

TABLE 9 Effects of the Car content on Chl_{leaf} estimates in different seasons. Chl_{leaf} ref. represents the reference values; Car/Chl_{leaf} ref. are set according to previous studies (Gamon et al., 2016; Wong et al., 2019); Chl_{leaf} inv. represents the Chl_{leaf} inverted using the CSI-based method; and ΔChl_{leaf} inv. represents the change in the retrieved Chl_{leaf} under different Car conditions. AE represents the absolute error

Vegetation condition	Ground-measured reflectance/VI		Relative error of L1C reflectance product		Relative error of L2A reflectance product	
	Sparse	Dense	Sparse	Dense	Sparse	Dense
Blue	0.120	0.074	93.2%	107.27%	20.3%	18.7%
RE1	0.350	0.138	−12.5%	16.56%	12.5%	21.28%
NIR	0.420	0.472	1.24%	0%	10.42%	5.00%
CSI	0.078	0.737	294.59%	59.72%	−3.96%	−11.48%

TABLE 10 Ground-measured reflectance of each band used in the CSI and the relative errors of the L1C and L2A products

largely reduced using CSI. The Chl_{leaf} values inverted with the CSI regression method display similarly high accuracies for the four vegetation types (Table 8). In contrast, accuracies for different vegetation types obtained using the other indices fluctuate more dramatically. For example, Datt99 performs well in CRP and GRA, but the RMSE and rRMSE of DBF are high. The CSI also achieves higher and more stable accuracy for different species of one type (Figure 7). Thus, the CSI-based empirical relationship method has the potential to expand the applicability from local areas to larger scales. In future research, the algorithm can also be applied to more sensors with red-edge bands to generate Chl_{leaf} products. Because of the high correlations between Chl_{leaf} and the light, drought stress (Khayatnezhad, 2012; Park & Matsumoto, 2018), the CSI-based method would potentially yield valuable information concerning the presence of biotic stress factors and abiotic stresses. It provides a convenient approach to better understand leaf-scale biochemistry in ecosystem modelling and ecological applications.

The CSI shows high sensitivity and a strong linear correlation with Chl_{leaf} through the careful selection of bands and ratios in the construction of the CSI. It utilizes a red-edge band and the NIR band, similar to many other VIs (Gitelson et al., 2003). Additionally, the CSI crucially incorporates the blue band, which has different responses to Chl_{leaf} and LAI. The potential of the blue band to improve Chl_{leaf} inversion has also been confirmed in recent studies (Jin & Wang, 2019; O'Reilly & Werdell, 2019). Multiplying NDVIre by ρ_{blue} / ρ_{RE1} increases the sensitivity to Chl_{leaf} and improves the resistance to saturation. Problems of saturation under high Chl_{leaf} conditions for some indices, such as TCARI/OSAVI and Datt99, are successfully reduced in the CSI (Figure 7d–g).

Due to the sensitivity of the blue-band reflectance to the carotenoids and atmosphere conditions, errors can be caused when applying the CSI to satellite images. The errors vary under different Chl_{leaf} conditions (Table 9), but the accuracy is better than the existing indices (Table 8). The underestimation in summer is slight when Chl_{leaf} is a more important independent factor affecting reflectance (Lewandowska & Jarvis, 1977; Thomas & Gausman, 1977). In winter, Car become more decisive in determining reflectance and overestimate the CSI-estimated Chl_{leaf}. Figures 7g and 8 show the overestimation and higher RMSE when Chl_{leaf} is low, but the results also illustrate that the errors of all VI-based methods increase. The large errors of other indices may result from the large proportion of soil information in canopy reflectance. The low Chl_{leaf} and LAI conditions are still a challenge for estimating Chl_{leaf} due to the weak information available for leaves. Further research to decouple the canopy/soil information and Car contents from Chl_{leaf} has the potential to improve the Chl_{leaf} estimation. However, the results of the sensitivity analysis (Figure 4) indicate a similar effect of Car contents on Chl_{leaf} estimations to that of the leaf structure parameter *N*, and both of them are far less sensitive than Chl_{leaf}. Based on this finding, the effect of Car contents on the CSI-based regression method is limited. In the CSI definition, sensitivity to Car was introduced only by ρ_{blue}, but the product of the calculator ρ_{blue} / ρ_{RE1} times NDVIre, which are both sensitive to Chl_{leaf}, doubles the sensitivity to Chl_{leaf} in the CSI.

The blue band shows sensitivity to atmospheric conditions, generating uncertainties in the CSI-estimated Chl_{leaf}. The scattering effects of aerosols increase the reflectance of blue band and CSI. If the pixel is still contaminated by clouds after atmospheric correction, the CSI has a higher value and significantly overestimates Chl_{leaf}. The L2A reflectance data tend to produce an underestimated Chl_{leaf} (3.96% - 11.48%), especially when the canopies are dense (Table 10). The validation results (Figure 6) also prove that CSI-derived Chl_{leaf} tends to be underestimated (bias = −4.954 $\mu g\,cm^{-2}$) when Chl_{leaf} is high (>45 $\mu g\,cm^{-2}$) thus, accurate atmospheric correction will further improve the accuracy.

The leaf structure parameter (N) is another influential factor in CSI-based Chl_{leaf} estimations (Figure 4). N represents the complexity of the leaf internal structure. A larger N causes multiple scattering and more absorption of pigments inside the leaf. The red-edge bands are capable of capturing absorption information. The definition of the CSI enlarges the information in the RE1 band and enlarges the effect of absorption on Chl_{leaf} and N. Recent research suggests that the N values vary significantly in different phenological stages (Boren et al., 2019). Thus, the empirical relationship between the CSI and Chl_{leaf} trained over more specific plant functional types in different growing periods will improve the Chl_{leaf} estimation accuracy. Additionally, the mixed pixel effect is widespread in the vegetative area (Yu et al., 2018), and the understorey vegetation plays an important role in the forest ecosystem (Nilsson & Wardle, 2005). Thus, the CSI–Chl_{leaf} relationship over mixed pixels also deserves further study.

The validation experiments were performed by independent researchers, whose methods to measure Chl_{leaf} varied, and the dataset was located at mid-high latitudes. Therefore, the accuracy reported in this study may be influenced by the validation dataset, and more ground measurements covering wider geographic areas and more vegetation species will be helpful to better evaluate the accuracy of the retrieved Chl_{leaf}.

5 | CONCLUSIONS

A new chlorophyll-sensitive index, CSI, was proposed in this research. Based on the strong chlorophyll absorption at blue wavelengths, the ρ_{blue} / ρ_{RE1} calculator was designed to strengthen the positive response to Chl_{leaf} and the negative response to LAI. By multiplying ρ_{blue} / ρ_{RE1} with the index NDVIre, which has a positive response to both LAI and Chl_{leaf}, the CSI displays a weaker response to LAI and stronger response to Chl_{leaf}. The CSI empirical regression method was derived to calculate Chl_{leaf}. The validation with ground measurements for four vegetation types showed that the CSI method has the highest overall accuracy (RMSE = 9.39 $\mu g\,cm^{-2}$, rRMSE = 23.83%, R^2 = 0.49) among the 11 VI regression methods. The CSI also performs best for each of the four vegetation types (RMSE = 9.51 $\mu g\,cm^{-2}$ for CRP; RMSE = 7.04 $\mu g\,cm^{-2}$ for DBF; RMSE = 9.52 $\mu g\,cm^{-2}$ for ENF; RMSE = 11.01 $\mu g\,cm^{-2}$ for GRA). The CSI-estimated Chl_{leaf} shows high and stable accuracy under different LAI and Chl_{leaf} (larger than 20 $\mu g\,cm^{-2}$) conditions. Due to the sensitivity to Car contents, the CSI-estimated Chl_{leaf} tends to be overestimated when Chl_{leaf} is lower than 20 $\mu g\,cm^{-2}$. A 30-m and 10-day Chl_{leaf} product across China was also generated based on the CSI regression method. It has the potential to be applied in generating continental or global Chl_{leaf} products. Future studies should focus on training the CSI-based regression models over more specific plant functional types in different growing periods and over mixed pixels with different land cover/vegetation types. Further validation in wider regions with more vegetation functional types is beneficial to evaluate the accuracy of the CSI-based method.

AUTHOR CONTRIBUTIONS

Hu Zhang, Jing Li and Qinhuo Liu conceived the ideas and designed the methodology. Liangyun Liu, H. Croft, Jan. Clevers and Chenpeng Gu collected the data. Xiaohan Wang, Zhaoxing Zhang, Jing Zhao, Yadong Dong and Wentao Yu analysed the data. Hu Zhang, Jing Li and Shangrong Lin led the writing of the manuscript. Alfredo Huete and Yelu Zeng reviewed and edited the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

This work was supported by the National Key Research and Development Program (2019YFE0126700), and the National Natural Science Foundation of China (No. 41871265). We appreciate the Center for Advanced Land Management Information Technologies (CALMIT), University of Nebraska–Lincoln, for providing us with the in situ spectra and biochemical datasets.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13994.

DATA AVAILABILITY STATEMENT

Chl_{leaf} data are available from the Science Data Bank (DOI: https://doi.org/10.11922/sciencedb.j00001.00265, https://www.scidb.cn/en/detail?dataSetId=846695127865884672#) (Li et al., 2021). Code for figures and validation data are available at https://doi.org/10.5281/zenodo.7088530 (Zhang, 2022).

ORCID

Hu Zhang https://orcid.org/0000-0003-2874-9219
Jing Li https://orcid.org/0000-0001-9736-5732
Qinhuo Liu https://orcid.org/0000-0002-3713-9511
Shangrong Lin https://orcid.org/0000-0001-7153-2184
Liangyun Liu https://orcid.org/0000-0002-7987-037X
Jan G. P. W. Clevers https://orcid.org/0000-0002-0046-082X
Yelu Zeng https://orcid.org/0000-0003-4267-1841
Chenpeng Gu https://orcid.org/0000-0002-5049-5262
Jing Zhao https://orcid.org/0000-0001-7221-3556
Faisal Mumtaz https://orcid.org/0000-0003-1322-126X
Wentao Yu https://orcid.org/0000-0002-2348-5050

REFERENCES

Boren, E. J., Boschetti, L., & Johnson, D. M. (2019). Characterizing the variability of the structure parameter in the PROSPECT Leaf Optical properties model. *Remote Sensing*, *11*, 1236.

Chang, J. S., & Shoshany, M. (2016). Red-edge ratio normalized vegetation index for remote estimation of green biomass. *2016 Ieee International Geoscience and Remote Sensing Symposium (Igarss)*, 1337–1339. https://ieeexplore.ieee.org/document/7729340

Chen, J. M., & Leblanc, S. G. (1997). A four-scale bidirectional reflectance model based on canopy architecture. *IEEE Transactions on Geoscience and Remote Sensing*, *35*, 1316–1337.

Clevers, J., Kooistra, L., & van den Brande, M. (2017). Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. *Remote Sensing*, *9*(5), 405.

Combal, B., Baret, F., Weiss, M., Trubuil, A., Macé, D., Pragnère, A., Myneni, R., Knyazikhin, Y., & Wang, L. (2003). Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem. *Remote Sensing of Environment*, *84*, 1–15.

Croft, H., & Chen, J. M. (2018). Leaf pigment content. In S. Liang (Ed.), *Comprehensive remote sensing* (pp. 117–142). Elsevier.

Croft, H., Chen, J. M., Luo, X., Bartlett, P., Chen, B., & Staebler, R. M. (2017). Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Global Change Biology*, *23*, 3513–3524.

Croft, H., Chen, J. M., Wang, R., Mo, G., Luo, S., Luo, X., He, L., Gonsamo, A., Arabian, J., Zhang, Y., Simic-Milas, A., Noland, T. L., He, Y., Homolová, L., Malenovský, Z., Yi, Q., Beringer, J., Amiri, R., Hutley, L., … Bonal, D. (2020). The global distribution of leaf chlorophyll content. *Remote Sensing of Environment*, *236*, 111479.

Croft, H., Chen, J. M., & Zhang, Y. (2014). The applicability of empirical vegetation indices for determining leaf chlorophyll content over different leaf and canopy structures. *Ecological Complexity*, *17*, 119–130.

Dash, J., & Curran, P. J. (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing*, *25*, 5403–5413.

Datt, B. (1999). Visible/near infrared reflectance and chlorophyll content in eucalyptus leaves. *International Journal of Remote Sensing*, *20*, 2741–2759.

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., de Colstoun, E. B., & McMurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, *74*, 229–239.

Demarez, V., & Gastellu-Etchegorry, J. P. (2000). A modeling approach for studying Forest chlorophyll content. *Remote Sensing of Environment*, *71*, 226–238.

Evans, J. R. (1989). Photosynthesis and nitrogen relationships in leaves of C3 plants. *Oecologia*, *78*, 9–19.

Frampton, W. J., Dash, J., Watmough, G., & Milton, E. J. (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, *82*, 83–92.

Gamon, J. A., Huemmrich, K. F., Wong, C. Y. S., Ensminger, I., Garrity, S., Hollinger, D. Y., Noormets, A., & Peñuelas, J. (2016). A remotely sensed pigment index reveals photosynthetic phenology in evergreen conifers. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 13087–13092.

Gitelson, A., & Merzlyak, M. N. (1994). Spectral reflectance changes associated with autumn senescence of aesculus-Hippocastanum L and acer-Platanoides L leaves - spectral features and relation to chlorophyll estimation. *Journal of Plant Physiology*, *143*, 286–292.

Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, *160*, 271–282.

Gitelson, A. A., Vina, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, *32*, L08403.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*, *81*, 416–426.

Jacquemoud, S., & Baret, F. (1990). PROSPECT: A model of leaf optical properties spectra. *Remote Sensing Environ*, *34*, 75–91.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C., & Ustin, S. L. (2009). PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, *113*, S56–S66.

Jin, J., & Wang, Q. (2019). Selection of informative spectral bands for PLS models to estimate foliar chlorophyll content using hyperspectral reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, *57*, 3064–3072.

Khayatnezhad, M. (2012). The effect of drought stress on leaf chlorophyll content and stress resistance in maize cultivars (Zea mays). *African Journal of Microbiology Research*, *6*(12), 2844–2848.

Lewandowska, M., & Jarvis, P. G. (1977). Changes in chlorophyll and carotenoid content, specific leaf area and dry weight fraction in Sitka spruce, in response to shading and season. *New Phytologist*, *79*, 247–256.

Li, J., Zhang, H., Wang, X., Zhang, Z., Gu, C., Wen, Y., Chu, T., & Liu, Q. (2021). *MuSyQ 30m/10days leaf chlorophyll content product (from 2019 to 2020 across China version 01)*. Science Data Bank.

Liu, L., Wang, J., Huang, W., & Zhao, C. (2010). Detection of leaf and canopy EWT by calculating REWT from reflectance spectra. *International Journal of Remote Sensing*, *31*, 2681–2695.

Liu, L., Zhang, X., Chen, X., Gao, Y., & Mi, J. (2020). GLC_FCS30-2020: Global land cover with fine classification system at 30m in 2020 (v1.2). *Zenodo*.

Luo, X., Croft, H., Chen, J. M., He, L., & Keenan, T. F. (2019). Improved estimates of global terrestrial photosynthesis using information on leaf chlorophyll content. *Global Change Biology*, *25*, 2499–2514.

Maccioni, A., Agati, G., & Mazzinghi, P. (2001). New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. *Journal of Photochemistry and Photobiology. B*, *61*, 52–61.

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., & Gascon, F. (2017). *Sen2Cor for Sentinel-2*. SPIE.

Nilsson, M. C., & Wardle, D. A. (2005). Understory vegetation as a Forest ecosystem driver: Evidence from the northern Swedish boreal Forest. *Frontiers in Ecology and the Environment*, *3*, 421–428.

O'Reilly, J. E., & Werdell, P. J. (2019). Chlorophyll algorithms for ocean color sensors - OC4, OC5 & OC6. *Remote Sensing of Environment*, *229*, 32–47.

Park, S. G., & Matsumoto, M. (2018). A study on the effects of light conditions on the longevity and characteristics of *Daphniphyllum macropodum* leaves. *Journal of the Faculty of Agriculture, Kyushu University*, *63*, 15–19.

Simic, A., Chen, J. M., & Noland, T. L. (2011). Retrieval of forest chlorophyll content using canopy structure parameters derived from multi-angle data: The measurement concept of combining nadir hyperspectral and off-nadir multispectral data. *International Journal of Remote Sensing*, *32*, 5621–5644.

Sims, D. A., & Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, *81*, 337–354.

Sola, I., García-Martín, A., Sandonís-Pozo, L., Álvarez-Mozos, J., Pérez-Cabello, F., González-Audícana, M., & Montorio Llovería, R. (2018). Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes. *International Journal of Applied Earth Observation and Geoinformation*, *73*, 63–76.

Thomas, J. R., & Gausman, H. W. (1977). Leaf reflectance vs. leaf chlorophyll and carotenoid concentrations for eight crops. *Agronomy Journal*, *69*, 799–802.

Uddling, J., Gelang-Alfredsson, J., Piikki, K., & Pleijel, H. (2007). Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynthesis Research*, *91*, 37–46.

Vernon, L. P., & Seely, G. R. (1966). *The chlorophylls*. Academic Press.

Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J. P., Camps-Valls, G., & Moreno, J. (2012). Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sensing of Environment*, *118*, 127–139.

Viña, A., Gitelson, A. A., Nguy-Robertson, A. L., & Peng, Y. (2011). Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sensing of Environment*, *115*, 3468–3478.

Wong, C. Y., D'Odorico, P., Bhathena, Y., Arain, M. A., & Ensminger, I. (2019). Carotenoid based vegetation indices for accurate monitoring of the phenology of photosynthesis at the leaf-scale in deciduous and evergreen trees. *Remote Sensing of Environment*, *233*, 111407.

Wu, C. Y., Niu, Z., Tang, Q., & Huang, W. J. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and Forest Meteorology*, *148*, 1230–1241.

Xu, M., Liu, R., Chen, J. M., Liu, Y., Shang, R., Ju, W., Wu, C., & Huang, W. (2019). Retrieving leaf chlorophyll content using a matrix-based vegetation index combination approach. *Remote Sensing of Environment*, *224*, 60–73.

Yu, W., Li, J., Liu, Q., Zeng, Y., Zhao, J., Xu, B., & Yin, G. (2018). Global land cover heterogeneity characteristics at moderate resolution for mixed pixel modeling and inversion. *Remote Sensing*, *10*, 856.

Zarco-Tejada, P. J., Hornero, A., Beck, P. S. A., Kattenborn, T., Kempeneers, P., & Hernandez-Clemente, R. (2019). Chlorophyll content estimation in an open-canopy conifer forest with sentinel-2A and hyperspectral imagery in the context of forest decline. *Remote Sensing of Environment*, *223*, 320–335.

Zhang, H. (2022). Zhanghu927/CSI_codes: CSI codes and data. In Zhanghu927 (ed.): Zenodo.

Zhang, Y., Chen, J. M., Miller, J. R., & Noland, T. L. (2008). Leaf chlorophyll content retrieval from airborne hyperspectral remote sensing imagery. *Remote Sensing of Environment*, *112*, 3234–3247.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhang, H., Li, J., Liu, Q., Lin, S., Huete, A., Liu, L., Croft, H., Clevers, J. G. P., Zeng, Y., Wang, X., Gu, C., Zhang, Z., Zhao, J., Dong, Y., Mumtaz, F., & Yu, W. (2022). A novel red-edge spectral index for retrieving the leaf chlorophyll content. *Methods in Ecology and Evolution*, *00*, 1–17. https://doi.org/10.1111/2041-210X.13994