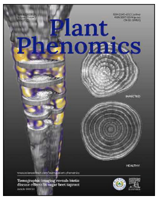
Plant Phenomics 7 (2025) 100059

Contents lists available at ScienceDirect

Plant Phenomics

journal homepage: www.sciencedirect.com/journal/plant-phenomics



Research Article

A novel constrained spectral matching approach for extending UAV multispectral reflectance measurements and estimating nitrogen and phosphorus contents in wetland vegetation species

Zhinan Lao [a,b,1], Bolin Fu [a,1,*], Weiwei Sun [c], Yeqiao Wang [d], Yuyu Zhou [e,f], Hongchang He [a], Tengfang Deng [a], Ertao Gao [a]

[a] College of Geomatics and Geoinformation, Guilin University of Technology, Guilin, 541006, China
[b] School of Geography, South China Normal University, Guangzhou, 510631, China
[c] Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo, 315211, China
[d] Department of Natural Resources Science, University of Rhode Island, Kingston, RI, USA
[e] Department of Geography, The University of Hong Kong, Hong Kong, China
[f] Institute for Climate and Carbon Neutrality, The University of Hong Kong, Hong Kong, China

ARTICLE INFO

Keywords:
Wetlands vegetation species
Nitrogen and phosphorus estimation
Spectral matching
Hybrid retrieval strategy
UAV multispectral image

ABSTRACT

Canopy nitrogen content (CNC) and canopy phosphorus content (CPC) of vegetation in wetlands are key physiological traits, which can be associated with the process of wetland ecosystems. Because of the spectral signals obscured by pigments and water content, it is challenging to accurately estimate CNC and CPC of vegetation species in wetlands using multispectral images. Therefore, we developed the constrained PROSAIL-PRO spectra matching (CPSM) approach to extend multispectral reflectance of unmanned aerial vehicle measurements to 400 ~ 2500 nm. We verified the matched accuracy and spectral reliability of CPSM's spectra from two aspects of reflectance and vegetation spectral characteristic based on field-measured spectral data. We proposed a novel hybrid retrieval strategy to achieve the high-precision estimation of CNC and CPC for seven karst wetland vegetation species. Finally, we evaluated the applicability of combining CPSM with our strategy to estimate CNC and CPC for two typical species in mangrove wetlands. Our results proved that CPSM-based spectra had good consistency with original reflectance of UAV images ($R^2 = 0.82 \sim 0.86$), and they could maintain similar spectral characteristics to measured spectra. Besides, this study found that the optimal spectral features of CNC and CPC were distributed near the red edge position and water-absorption valley of vegetation spectra. We obtained high-precision estimation of CNC and CPC in karst wetland using CPSM and our hybrid retrieval strategy ($R^2 = 0.60 \sim 0.98$, MRE = 5.91 % ~ 26.25 %). The approach also showed a better transferring performance in estimating CNC and CPC of mangrove species ($R^2 = 0.77 \sim 0.89$, MRE = 9.65 % ~ 16.87 %). The CPSM approach is effective to achieve high-precision estimation of vegetation CNC and CPC.

1. Introduction

Wetlands are one of the most important and valuable ecosystems on Earth. Wetlands play important roles in purifying water quality, reducing floods, protecting coastlines and regulating climate [1]. Wetlands provide the habitats for many endemic and unique species and maintain biodiversity on Earth [2]. Wetlands have the great potential to restore natural ecosystems through lakes, marshes, rivers, and vegetation, thereby contributing to the achievement of various Sustainable Development Goals (SDGs) ([3]). Vegetation, as one of the three key components of wetland ecosystems, not only helps prevent soil erosion but also sustains the hydrological balance within wetlands, providing a sensitive indicator of changes in the ecosystem [4]. Nitrogen and phosphorus, as vital nutrients, are crucial for plant growth and a range of physiological regulatory functions [5]. Previous studies confirmed that the nitrogen and phosphorus levels in wetlands can limit the biomass production of vegetation [6]. Koerselman and Meuleman [7] reported that if the nitrogen/phosphorus radio of the aboveground

---

* Corresponding author.
E-mail address: fubolin@glut.edu.cn (B. Fu).

[1] These authors contributed equally to this paper.

https://doi.org/10.1016/j.plaphe.2025.100059
Received 17 February 2025; Received in revised form 9 May 2025; Accepted 26 May 2025

biomass of the vegetation community is < 14 or >16, the biomass production is almost always limited by nitrogen or phosphorus, respectively. Moreover, the nitrogen content in vegetation, which is closely linked to photosynthesis and net primary production and is regarded as one of the most essential nutrients affecting its growth and quality [8]. Phosphorus content is a component of several important compounds, e.g., nucleic acids, phosphoric acid, and adenosine triphosphate, which can improve the resistance and adaptation of vegetation. Therefore, accurate estimation of canopy nitrogen content (CNC) and canopy phosphorus content (CPC) of different vegetation species is vital for ecological conservation and preservation of biodiversity in wetlands.

Currently, unmanned aerial vehicle (UAV) technology with the ultra-high resolution, high efficiency and versatility has been utilized to estimate vegetation's physiological traits [9]. However, the sensitive spectral bands of nitrogen and phosphorus in vegetation are usually distributed in short-wave infrared bands within 1000 ~ 2500 nm [10–12]. Furthermore, most of UAV multispectral sensors (DJI Phantom 4 Multispectral UAV, DJI MAVIC 3 Multispectral UAV, RedEdge-MX, etc.) are not cover the sensitive short-wave infrared region. While satellite-based multispectral sensors do operate within the SWIR range, the inherent spectral resolution of broadband configurations may diminish the diagnostic contribution of narrow sensitive bands to CNC and CPC. Consequently, accurately capturing spectral signatures critical to CNC and CPC estimation remains challenging under these constraints. Previous researchers have demonstrated that hyperspectral remote sensing with various continuous narrow bands can identify the absorption characteristics of nitrogen and phosphorus in vegetation [13]. However, the high cost of UAV-based hyperspectral sensors currently limits their large-scale application. To address this issue, spectral super-resolution technology has emerged as a promising solution. Currently, the spectral super-resolution reconstruction of multispectral images has been accomplished based on existing hyperspectral images and deep learning (DL) algorithms, which possess powerful spatial-spectral extraction and nonlinear projection capabilities [14]. The intensity of energy required by optical remote sensor to form a reflectance image varies not only by time and place, but also by the internal structure and component of vegetation, which have different responses to different wavelengths. Therefore, compared with DL algorithms, radiative transfer models (RTMs, such as PROSPECT and SAIL, etc.) can forward simulate spectral reflectance across the entire vegetation spectrum, furthering estimate the signals of spaceborne or airborne spectral sensors based on the physical laws of the interaction between electromagnetic radiation and vegetation structures [15]. However, the reflectance of vegetation in RTMs is strongly influenced by biophysical and biochemical parameters. For example, the chlorophyll content significantly affects the response of visible light reflectance [16]; the change of leaf water content causes the spectral absorption position to significantly shift [17]; the leaf area index (LAI) greatly affects the reflectance of the simulated spectrum. Therefore, if the reflectance of the vegetation canopy within the 400 ~ 2500 nm range can be inverted using RTMs, given the known energy intensity, vegetation structure, and biochemical composition of a specific area on one day, it could offer a novel approach for extending the vegetation spectrum through RTMs to estimate physiological traits.

Currently, researches on the retrieval of CNC and CPC primarily involve empirical models, machine learning (ML) or deep learning (DL) algorithms, and RTMs. Empirical models, which rely on statistical relationships between spectral reflectance and CNC or CPC, have been widely utilized due to their simplicity. However, challenges in retrieval of CNC and CPC arise from the variation in sensitive spectral bands, as well as the confounding and overlapping effects of biophysical and biochemical parameters on CNC and CPC. For example, the spectral reflectance response of vegetation nitrogen content was masked by strong water absorption, soil background, and canopy structure of vegetation species in near-infrared and short-wave near infrared bands [18]. Recurring water absorption features, caused by the overtones of the fundamental excitation frequency of O-H bonds in water, predominantly influence the spectral reflectance of vegetation in the NIR range [19]. Additionally, variations in soil water content affect the background, leading to changes in the brightness of spectral properties [20]. Therefore, derivative transform [21], logarithmic transform [22], continuous wavelet transforms [23] and spectral index approaches have been used to identify the reflection characteristics and reduce the influence of other biophysical and biochemical parameters on spectral reflectance. Jin and Wang [24] showed that derivative transforms can extract valuable spectral information from the original spectrum. However, the first and second derivatives may also result in the loss of important spectral data or an increase in noise [25]. Fractional order derivatives (FOD) can help eliminate spectral peak overlap and baseline shifts, effectively reducing the impacts of various radiation and scattering [26]. This approach also enhances the robustness of empirical models and provides optimal retrieval features for machine learning (ML) or deep learning (DL) algorithms.

Although ML and DL algorithms have demonstrated strong performance in estimating nitrogen and phosphorus levels [27,28], their results often lack interpretability. Previous studies have shown that radiative transfer models (RTMs), which are based on well-defined physical principles, outperform empirical models in terms of transferability and stability [29]. However, challenges remain in retrieving CNC using radiative transfer models (RTMs) due to ill-posed inversion problems and the need to correct absorption coefficients [17]. To enhance retrieval accuracy for vegetation physiological traits, previous studies have integrated RTMs with machine learning (ML) or deep learning (DL) algorithms to develop hybrid models [30]. This approach leverages the ability of ML and DL algorithms to identify complex nonlinear relationships [31] between dependent and independent variables within large training datasets. The interpretability and predictive capability of these models were further improved by embedding the physical laws governing the relationship between vegetation canopy reflectance and biophysical and biochemical parameters from RTMs into ML or DL algorithms. However, in most studies, hybrid model construction primarily involves increasing the training dataset size by incorporating simulation data generated by RTMs [32]. It is often overlooked that the knowledge learned from hybrid models trained on simulated sample data may not align with the inherent characteristics of real-world scenarios, leading to limited generalization ability. Additionally, when RTMs are used to simulate the spectral reflectance of multispectral sensors such as Landsat TM/OLI or Sentinel-2 MSI, their inherent advantage is diminished. This is because RTMs are capable of simulating vegetation reflectance across 2101 bands at 1 nm intervals from 400 to 2500 nm, whereas multispectral sensors capture reflectance in only a few broad bands. However, integrating RTM-simulated spectral data (400 ~ 2500 nm) with measured vegetation traits could introduce a novel retrieval strategy for vegetation physiological traits. At present, most studies have evaluated estimation models either within a single scene or across multiple species within a single scene, with limited research assessing models for multiple species across different wetland ecosystems.

In this study, we combined a novel constrained spectra matching approach with a new hybrid retrieval strategy to achieve high-precision retrieval of CNC and CPC of vegetation species. Meanwhile, we evaluated the transfer ability and applicability of our method in different wetlands. The main contributions of this paper are as follows:

(1) We introduced an innovative constrained PROSAIL-PRO spectra matching (CPSM) method, which incorporates field-measured physiological traits to constrain the PROSAIL-PRO model (PROSPECT-PRO + 4SAIL), enabling the extension of UAV multispectral reflectance from the 450–840 nm range to 400 ~ 2500 nm;

(2) We verified the consistency of CPSM-based spectral data and the field-based spectra in reflectance and vegetation spectral characteristics;

(3) We proposed a new hybrid retrieval strategy for estimating CNC and CPC by combining the benefits of ultra-high spatial resolution UAV imagery with the high spectral resolution data from the CPSM approach;

(4) This study further evaluated the transfer ability and robustness of combining CPSM with the hybrid strategy to estimate CNC and CPC of mangrove species.

2. Materials

2.1. Study area

2.1.1. Karst wetland

Huixian Karst Wetland of International Importance (HKWI) (Fig. 1A, B) (25°05 ′20 "N ~ 25°06′ 55″ N, 110°10 ′50 "E ~ 110°14′ 21″ E) is the largest and most representative karst wetland in subtropical peak forest plain regions globally. The region belongs to a subtropical monsoon humid climate with an average annual temperature and precipitation of 19.2 °C and 1863.2 nm [33], respectively. It is composed of lakes, marshes, permanent rivers, karst underground rivers, ponds, artificial canals, farmland, and typical karst vegetation species, representing a complex ecosystem of wetland-farmland-forests, harboring representative vegetation species of karst wetlands, such as *Cladium chinense Nees*, *Phragmites australis (Cav.) Trin. ex Steud*, *Typha domingensis Pers.* and *Carex baccans Nees*. The maximum difference in water level between the wet season and dry season is 1.814 m [34].

2.1.2. Mangrove wetland

In order to verify the applicability of our method in other wetland types, two mangrove wetlands were selected as the study area, including Jinhaiwan Mangrove Natural Reserve (JMNR) (21°24′35″ ~ 21°25′30″ N, 109°9′30″~ 109°13′40″E) and Dangjiang Mangrove Natural Reserve (DMNR) (21°33′50″ ~ 21°35′10″ N, 109°5′50″~ 109°7′10″ E), both located in Beihai, China (Fig. 1C, D). There are obvious differences in the growing environment between karst wetland's vegetation and mangroves. We confirmed that there were significant differences in vegetation traits between the two wetlands types using ANOVA analysis (Table S1 in Supplementary Materials 1). The mangrove species are dominated by the *Avicennia marina* (*AM*) in JMNR and the *Aegiceras corniculatum* (*AC*) in DMNR, with a minor presence of *Kandelia candel*, *Rhizophora stylosa*, *Sonneratia caseolaris*, *Bruguiera gymnorrhiza*. A medicinal species of *Acanthus ilicifolius* also grow in DMNR.

2.2. Data source

2.2.1. UAV image acquisition and processing

The multispectral images were collected using a DJI Phantom 4 Multispectral UAV over the study area on November 27–28, 2021 (Fig. S1a in Supplemental Materials 1) with the clear, cloudless conditions from 11:00 to 15:00 a.m. (UTC+8:00). The spectral bands include blue (450 nm ± 16 nm), green (560 nm ± 16 nm), red (650 nm ± 16 nm), red edge (730 nm ± 16 nm) and near infrared (840 nm ± 26 nm). We set the heading and side overlap rates to be 80 % and 70 %, respectively. The UAV flight height was 105 m with the speed of 5.7 m/s.

Besides, we used the same sensor to collect multispectral images of mangroves in DMNR (April 10 ~ 25, 2023) and JMNR (August 21 ~ 25, 2023). These two sets of multispectral images were used to verify the applicability of combining CPSM and our strategy proposed in this study in other regions. The acquisition progress of UAV multispectral images was the same as that of karst wetland. Finally, the Pix4D software was used to conduct geometric correction, image mosaicking, and radiation calibration of the original UAV images, and the digital orthophoto images with a spatial resolution of 0.05m were generated in the study areas.

2.2.2. Field measurements

We collected 240 samples of seven typical vegetation species with high, moderate, and low inundation frequencies [34] in karst wetland from November 27 to December 3, 2021. The geographic locations were recorded using Hi-Target V90 RTK. The vegetation species of our study in karst wetland included *Eichhornia crassipes (Mart.) Solms* (ES), *Cladium chinense Nees* (CCN), *Carex baccans Nees* (CBN), *Cynodon dactylon (L.) Pers.* (CDP), *Miscanthus* (SR), *Bamboo* (BB), *Triadica sebifera (Linnaeus) Small* (TS). Firstly, we used an SPAD-502 Plus chlorophyll meter (Minolta Camera Co. Ltd., Japan) and an LAI-2200C (Li-Cor, Inc., Lincoln, NE, USA) to measure the leaf chlorophyll content (LCC) [34], leaf area index, and average leaf inclination angle (ALA) of each species (Fig. S1 in Supplemental Materials 1) [21]. Secondly, we randomly picked 20 leaves from the upper crown layer of the vegetation in the four directions. The samples were stored in a dark and cool insulated container, and brought back to the laboratory for measurement of total nitrogen (TN) and total phosphorus (TP) contents. Samples of *Aegiceras corniculatum* (76) and *Avicennia marina* (112) were collected and measured for physiological traits using the same method. The works were synchronized with UAV images acquisition.

2.2.3. Laboratory measurement of physiological traits of vegetation species

We recorded the fresh weight (W_f) of each sampled leaf using an electronic scale with accuracy of 0.01 g and measured the area (LA (cm^2)) of each leaf using Li-Cor 3000C (Li-Cor, Inc., Lincoln, NE, USA) immediately after the field collection. We took 0.05g of leaves to measure carotenoid content (*Car*) and chlorophyll content (*Cab*) using spectrophotometry. Then, leaf samples were kept at 75 °C for 48 h to dry until they reached constant weight, and the dry weight (W_d) of each sample was recorded using the same electronic scale. The formulas of equivalent water thickness (*Cw*), dry matter content (*Cm*), *Car* and *Cab* are shown in the Supplementary Materials 2. Finally, we ground the dried leaves into powder (<0.150 mm) and used Clever Chem 380 (De Chem-Tech Company, Germany) to measure the TN and TP contents in the leaves based on the indophenol-blue method [35] and molybdenum antimony spectrophotometry [36], respectively. The statistical information of the measured data in karst wetland is summarized in Table S2 in Supplemental Materials 1.

3. Methodology

In this study, we firstly used a constrained RTM spectra matching approach to extend the spectral ranges of UAV multispectral sensor for covering the CPC and CNC sensitive spectral region (Section 3.1). Secondly, we verified the consistency between CPSM-based spectral data and the field-based spectra in reflectance and vegetation spectral characteristics, and further used the FOD-based spectra and its spectral indexes to discriminate the spectral features sensitive to CNC and CPC for constructing the retrieval dataset (Section 3.2); Besides, we proposed a new hybrid retrieval strategy to estimate the CNC and CPC of seven vegetation species in karst wetland (Section 3.3). Finally, we combined CPSM with the hybrid retrieval strategy to estimate CNC and CPC of mangrove species for verifying the robustness of our methods when it transferred to other wetland types. The workflow is illustrated in Fig. 2.

3.1. Constrained PROSAIL-PRO spectra matching approach

In this paper, a novel constrained PROSAIL-PRO spectra matching (CPSM) approach (Fig. S2 in Supplementary Materials 1) was developed by combining physiological trait parameters and PROSAIL-PRO to extend UAV's original multispectral reflectance to 400 ~ 2500 nm. The reflectance of each pixel of vegetation species in P4M multispectral images was used to invert to obtain the optimal solution range in the look-up table (LUT), including the simulated 400 ~ 2500 nm spectral

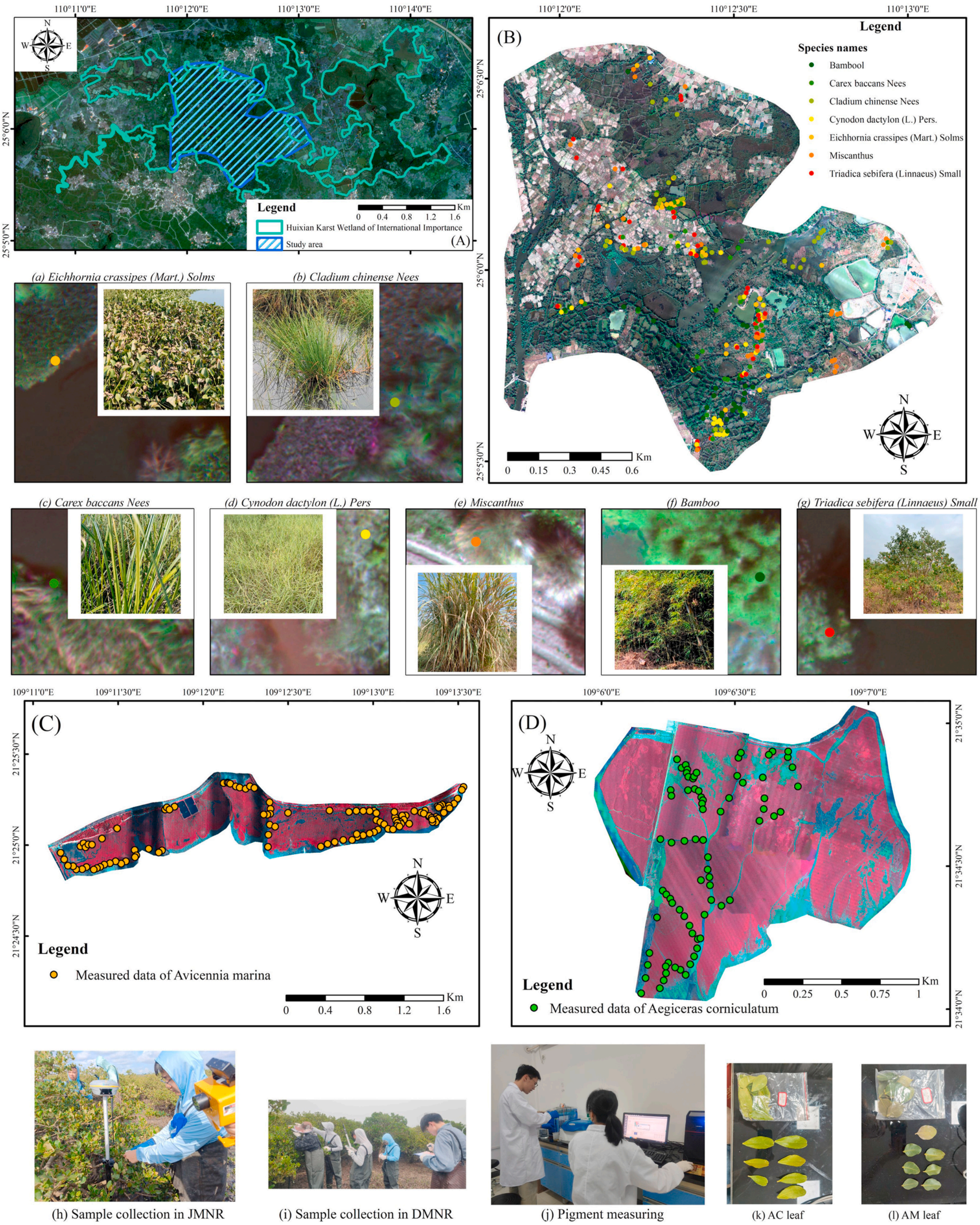


Fig. 1. Overview of the study area and the photos of typical vegetation species. (A) is the Huixian Karst Wetland of International Importance (HKWI) and its study areas and the background is true-color image from Jinlin-1 satellite; (B) is the true-color UAV image of study area of karst wetland; (C) and (D) are the Jinhaiwan Mangrove Natural Reserve (JMNR) and the Dangjiang Mangrove Natural Reserve (DMNR) areas, respectively, and the background is false-color image (R: NIR; G: Red; B: Green) from UAV multispectral sensor.

data, and the biophysical and biochemical parameters. Then, the LUT of the optimal solution range was used to obtain the matched spectrum with the field-based *LAI*, *Cw*, and *Cab* as the constrained conditions (Equation (1)). The LUT was obtained by PROSAIL-PRO forward simulation. The specific process of CPSM is shown in steps (1) ~ (4) below.

$$S = f\left(LUT, B_{Blue}, B_{Green}, B_{Red}, B_{Rededge}, B_{Nir}, Cab, Cw, LAI\right) \quad (1)$$

where S is the results of spectra matching; LUT is the look-up table of each vegetation species using PROSAIL-PRO forward simulation; B_{Blue}, B_{Green}, B_{Red}, $B_{Rededge}$, and B_{Nir} are the original blue, green, red, red edge and near infrared bands of P4M UAV multispectral images, respectively; and *Cab*, *Cw*, and *LAI* are the field-based chlorophyll content, equivalent water thickness, and leaf area index, respectively.

(1) **Generating the LUT of each vegetation species.** We combined PROSPECT-PRO (the latest version) with 4 SAIL to create the LUT of each species, which contained 14,000 simulated spectra from 400 nm to 2500 nm and corresponding biophysical and biochemical parameters. Seven parameter ranges of the PROSPECT-PRO and 4 SAIL models for each species were set according to the measured data (Table S2 in Supplementary Materials 1), and the settings and calculations (Equations (2) and (3)) of other parameters were referred to the studies of Yeoh and Wee [37], Candiani et al. [38], Raya-Sereno et al. [39]. The PROSAIL-PRO model parameters of all vegetation species are shown in Table S3 of Supplementary Materials 1. The CNC and CPC could be obtained by multiplying the protein and phosphorus contents by *LAI*, respectively (Equation (4)).

$$protein = \frac{TN(g/kg)/1000}{LA/W_d} \times 4.43 \tag{2}$$

$$\mathrm{Cab} = 8.24 \times e^{0.0324 \times LCC} \tag{3}$$

$$CNC = \mathrm{protein} \times \mathrm{LAI}, \mathrm{CPC} = \mathrm{TP} \times \mathrm{LAI} \tag{4}$$

where *protein*, *TN(g/kg)*, *Cab*, and *TP* are the leaf protein content, total nitrogen content, chlorophyll content, and total phosphorus content, respectively; *LAI* is the leaf area index; and W_d and *LA* are the leaf dry weight and leaf area, respectively. LCC is the leaf chlorophyll content measured using SPAD-502 Plus.

(2) **Simulating UAV multispectral data.** The spectral data were selected from the LUT of each species according to the bandwidth range of five P4M multispectral bands. The average values of the spectral data were calculated as the simulated spectral reflectance data (*LUT_P4M*) of P4M UAV sensor (Equation (5)) on the basis of its center wavelengths of five multispectral bands (Blue: 450 nm; Green: 560 nm; Red: 650 nm; Red edge: 730 nm; NIR: 840 nm).

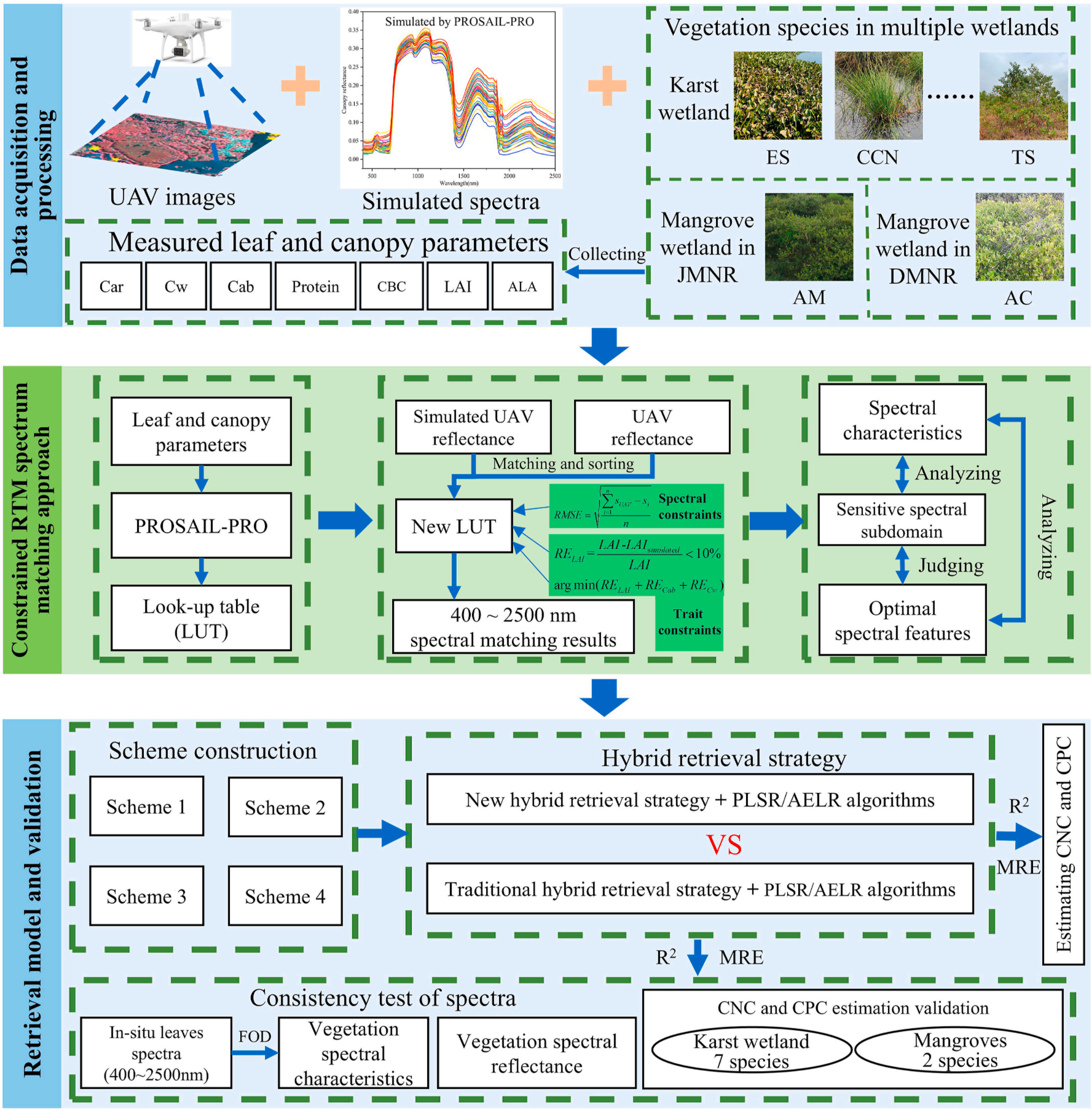


Fig. 2. The technical flow of this study.

$$B_{\mathrm{c}} = \frac{\sum_{a=c-b}^{c+b} B_a}{2b}, s_i = \left(B_{450} \quad \cdots \quad B_c \right), LUT\_P4M = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \tag{5}$$

where, B_c is the simulated reflectance of P4M, B_a is the simulated reflectance at wavelength a, c is the central wavelength ($c = 450$, 560, 650, 730, 840), b is half of the bandwidth at c band, and s_i is the ith simulated spectrum of P4M ($i = 1, 2, \ldots, n$).

(3) **Matching simulated and original spectral reflectance of UAV sensor.** The $RMSE_{P4M}$ between *LUT_P4M*'s spectral reflectance and the original P4M spectral reflectance was calculated according to Equation (6). *LUT_P4M* and the LUT were sorted based on *RMSE* in ascending order to retain their consistent indexes. The optimal solution dataset of canopy reflectance, $LUT_{20\ \%}$ (Equation (8)), was extracted from LUT according to the *RMSE*'s indexes of the first 20 % of *LUT_P4M* spectral data (Equation (7)).

$$RMSE_{P4M} = \left\{ A_{s_i} \middle| A_{s_i} = \sqrt{\frac{\sum_{i=1}^{n} \left(s_u - s_i\right)^2}{n}} \right\}, A_{s_i} < A_{s_{i+1}} \tag{6}$$

$$LUT\_P4M_{20\%} = f(t, LUT\_P4M) = \begin{pmatrix} s_{(1,1)} & \cdots & s_{(1,h)} \\ \vdots & \ddots & \vdots \\ s_{(t,1)} & \cdots & s_{(t,h)} \end{pmatrix}, t = \{t \in Z | 0 < t < (n \times 20\%)\} \tag{7}$$

$$P = \begin{pmatrix} P_1 \\ \vdots \\ P_u \end{pmatrix} = \min(CRE_x), S = f(LUT_{20\%}, P) = \begin{pmatrix} S_{(1,400)} & \cdots & S_{(1,j)} \\ \vdots & \ddots & \vdots \\ S_{(u,400)} & \cdots & S_{(u,j)} \end{pmatrix} \tag{10}$$

where CRE_x is the set of cumulative relative errors of the xth sample point; E_t is the cumulative relative error of the reflectance of the tth simulated spectrum; $RE_{LAI_t}, RE_{Cw_t}, RE_{Cab_t}$ are the relative errors between the tth simulated *LAI* and the uth measured *LAI*, between the tth simulated *Cw* and the uth measured *Cw*, and between the tth simulated *Cab* and the uth measured *Cab*, respectively; T_m and Y_m are the simulated and measured vegetation physiological traits (*LAI*, *Cw*, *Cab*), respectively; P is the index of the final matched spectrum data; S is the result of matched spectrum; and S_{uj} is the canopy spectral reflectance of the uth sample at j band.

3.2. Spectral characteristic extraction and retrieval scheme construction

In this study, we extracted the position and width of the red edge, the position of the blue edge and the yellow edge, the chlorophyll absorption valley, and the water absorption valley from the CPSM's spectra and the measured field-based spectral data, respectively, and to verify their consistency. Meanwhile, the correlation coefficient was used to extract the sensitive spectral subdomain from the CPSM's spectra. Furthermore, we used the fractional order derivative (FOD) derived by Grumwald-Letnikov and spectral indexes (SIs) to mine the spectral features sensitive to CNC and CPC. The extraction method of spectral characteristics and the specific formula of FOD are shown in Supplementary Materials 2.

$$LUT_{20\%} = f(t, LUT) = (T, S) = \begin{pmatrix} T_{(1,1)} & \cdots & T_{(1,v)} & S_{(1,400)} & \cdots & S_{(1,j)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ T_{(t,1)} & \cdots & T_{(t,v)} & S_{(t,400)} & \cdots & S_{(t,j)} \end{pmatrix}, j = 400, 401, \cdots, 2500 \tag{8}$$

where $RMSE_{P4M}$ is the root mean square error between the simulated P4M spectra reflectance and P4M reflectance; $LUTs_{20\ \%}$ is the top 20 % of LUT data; A_{s_i} is the root mean square error between the ith simulated P4M spectrum and the reflectance spectrum of P4M; s_u is the P4M reflectance spectrum of the uth sample; t is the index of the simulated P4M spectrum sorted according to RMSE.; h is the number of bands of P4M ($h = 1, 2, \ldots, 5$); T and S are the simulated vegetation physiological traits and simulated spectra for spectral matching, respectively. $v = 1, 2, 3$ are the *LAI*, *Cw*, and *Cab* used as constraints for spectral matching, respectively.

(4) **Obtaining the matched spectra based on vegetation parameter constraints.** We calculated the relative error (*RE*) between the measured *LAI*, *Cw*, *Cab* and the simulated *LAI*, *Cw*, *Cab* within $\mathrm{LUT}_{20\ \%}$, respectively. Since the complexity of the canopy structure greatly affects the relationship between canopy reflectance and vegetation physiological traits [40], the *RE* between the measured *LAI* and the simulated *LAI* was limited to less than 10 %. Then, the cumulative relative error (*CRE*) was obtained by adding the *RE* of *LAI*, the *RE* of *Cw*, and the *RE* of *Cab*. Finally, the canopy reflectance spectral curve with the smallest *CRE* (Equation (9)) was selected in $LUT_{20\ \%}$ as the result of matched spectrum of every UAV imagery pixel (Equation (10)).

$$CRE_x = \begin{pmatrix} E_1 \\ \vdots \\ E_t \end{pmatrix}, E_t = RE_{LAI_t} + RE_{Cw_t} + RE_{Cab_t} = \sum_{m=1}^{v} \frac{|T_m - Y_m|}{Y_m}, x = 1, 2, \cdots, u, RE_{LAI_t} < 10\% \tag{9}$$

We constructed four estimation schemes (Table S5 in Supplementary Materials 1) with spectral variables, including the original UAV spectral bands, 14 spectral indexes (Table S4 in Supplementary Materials 1), and FOD-based spectra. The difference spectral index (DSI), the ratio spectral index (RSI), the normalized difference spectral index (NDSI), and the nonlinear spectral Index (NLI) were calculated using the FOD-based spectra with the greatest correlation with CNC and CPC. The Pearson's correlation coefficient and the Recursive Feature Elimination (RFE) algorithm were used to perform feature optimization and multicollinearity elimination for each feature type.

3.3. Construction of retrieval models using novel hybrid retrieval strategy

Because the PROSAIL-PRO model considers a large number of biophysical and biochemical properties of vegetation in the leaf and canopy structure, combining the PROSAIL-PRO model with ML/DL algorithms provides great potential in the field of estimating vegetation physiological traits. Therefore, this study combined the PROSAIL-PRO model with partial least squares regression (PLSR) and adaptive ensemble learning regression (AELR) algorithms to develop a novel hybrid retrieval strategy (Fig. S3 in Supplementary Materials 1) to estimate CNC and CPC. The ensemble strategy of AELR and its main process are shown in Section 4 of Supplementary Materials 2.

The AELR model was constructed using five base models. The weight coefficients (Table S1 in Supplementary Materials 2) of each base model were automatically assigned based on CRE, derived from their training results. These base models included XGBoost (XGB), GradientBoost (GB), Random Forest (RF), Decision Tree (DT), and CatBoost (CB). Through comparison with traditional hybrid retrieval strategies [30], it can be seen that our hybrid retrieval strategy can take the canopy spectral reflectance of 400 ~ 2500 nm generated by CPSM and its derived spectral features as independent variables, and the measured CNC (and CPC) as dependent variable to construct estimation models. However, the independent variable of the traditional hybrid retrieval strategy is only the simulated spectral data of inherent bands of the multispectral sensor, and the dependent variable is the simulated biophysical or biochemical parameters. The specific process of the new hybrid retrieval strategy is as follows: (1) CPSM is used to obtain 400 ~ 2500 nm matched spectra data; (2) the original spectral dataset was created by combining the measured CNC and CPC data with their one-to-one corresponding matched spectra data; (3) the original spectral dataset was used to construct four estimation schemes; (4) the correlation coefficient and the RFE algorithm were used to perform feature selection from each estimation schemes to obtain its final dataset for estimating CNC and CPC. 70 % of the data in the final dataset were used for retrieval model training and 30 % were used for accuracy validation. The concrete paradigms of the traditional hybrid retrieval strategy and our hybrid retrieval strategy are shown in Equations (11) and (12), respectively.

$$PROSAIL \overset{Output}{\rightarrow} x_{simulation}, y_{simualtion} \underset{Train}{\overset{Input}{\rightarrow}} PLSR \Big/ AELR \underset{Input}{\overset{Output}{\rightleftarrows}} \begin{matrix} prediction \\ R_{rs} \end{matrix} \overset{Validation}{\longleftarrow} y_{measured} \tag{11}$$

$$PROSAIL \overset{Output}{\rightarrow} x_{simulation}, y_{simualtion} \underset{CPSM}{\overset{Output}{\rightleftarrows}} \begin{matrix} S_{n\times m} \\ R_{rs}, LAI, Cw, Cab \end{matrix} \underset{Split}{\overset{Combine}{\rightarrow}} \begin{matrix} S_{70\%(n\times m)}; R_{rs}, y_{measured_{70\% n\times 1}} \\ S_{30\%(n\times m)}; R_{rs}, y_{measured_{30\% n\times 1}} \end{matrix} \underset{Validation}{\overset{Train}{\rightleftarrows}} PLSR \Big/ AELR \tag{12}$$

where $x_{simulation}$ and $y_{simulation}$ are the simulated spectral reflectance and biophysical (or biochemical) parameters obtained from PROSAIL-PRO, respectively; PLSR and AELR are the partial least squares regression and adaptive ensemble learning regression algorithms; R_{rs} is the reflectance of the original multispectral images; $y_{measured}$ is the measured CNC or CPC; $S_{n\times m}$ represents results of spectra matching; n and m are the number of field samples and spectral bands, respectively.

3.4. Evaluation metric

The parameter constraint accuracy of CPSM, the consistency between the CPSM's spectral reflectance and UAV spectral reflectance, the consistency of spectral morphology and characteristics between the CPSM's spectra and measured leaves spectra were quantitatively evaluated using the coefficient of determination (R^2), mean relative error (MRE), root mean square error (RMSE), relative error (RE), and cumulative relative error (CRE). Furthermore, we use R^2, MRE and RMSE to evaluate the performance of CNC and CPC estimation in multiple wetlands, including evaluating the accuracy differences before and after spectra matching, between our retrieval strategy and traditional strategy, and between different algorithms. In addition, RE and CRE were used to assist in spectra matching and determine the weight of base models in the AELR algorithm. The evaluation metrics used in this study are shown in Supplementary Materials 2.

4. Results

4.1. Consistency verification and spectral sensitivity analysis of physiological traits

4.1.1. Consistency verification of spectral reflectance

We verified the precision of CPSM-based spectral matching results through spectral reflectance's consistency and constraint accuracies of physiological traits (*LAI*, *Cab*, and *Cw*). The CPSM-based spectral reflectance of karst wetland vegetation and mangroves matched well with the UAV multispectral reflectance (Fig. 3e) and had a good consistency (Fig. 3a) with the R^2 reaching 0.82 and 0.86, RMSE reaching 0.1091 and 0.0729, respectively. The RE and CRE (between field-based vegetation parameters and simulated vegetation parameters) of the spectral matching results for most of sample pixels were below 5 % and 10 %, respectively (Fig. 3f, g, h). Only 8 % of the pixel spectral matching results had CRE above 10 %. All of these results confirm that the CPSM-based spectral matching results can provide reliable spectral data for estimating CNC and CPC.

4.1.2. Consistency verification of vegetation spectral characteristics

We computed the first-order derivative of the measured spectral data from 350 to 2500 nm using the ASD FieldSpec 4 Hi-Res Spectroradiometer (Analytic Spectra Devices, Inc.) and then validated the consistency of ten key vegetation spectral characteristics (the positions of the red edge, blue edge, and yellow edge, as well as red edge width, chlorophyll absorption valley, and water absorption valley) between the CPSM spectra and the measured spectra (Fig. 3i, j, k). Our analysis revealed that most of the vegetation spectral features in the CPSM spectrum of each species closely matched those of the measured leaf spectra (Fig. 3i), and the CPSM spectra showed strong agreement with the measured spectra ($R^2 = 0.99$, MRE = 4.58 %) (Fig. 3j). Furthermore, the T-test result indicated there was not significant difference between the measured and CPSM spectral characteristics of nine vegetation species (*P*-values>0.05) (Fig. 3k). These results collectively confirm that the CPSM spectra align well with the measured spectra in terms of both spectral shape and vegetation spectral features.

4.1.3. Spectral sensitivity analysis of vegetation physiological traits

We used the improved Sobol's sensitivity index [41] to explore the spectral characteristics of physiological traits of the species in karst wetlands and mangrove wetlands (Fig. 4 and Table S6 in Supplemental Materials 1). We found that *LAI* was most sensitive to the canopy reflectance of all species at 400 ~ 2500 nm, with average contribution of 22.81 % ~ 69.44 % (Table S6 in Supplemental Materials 1), and it significantly affected the canopy reflectance at 400 ~ 684 nm (Fig. 4h).

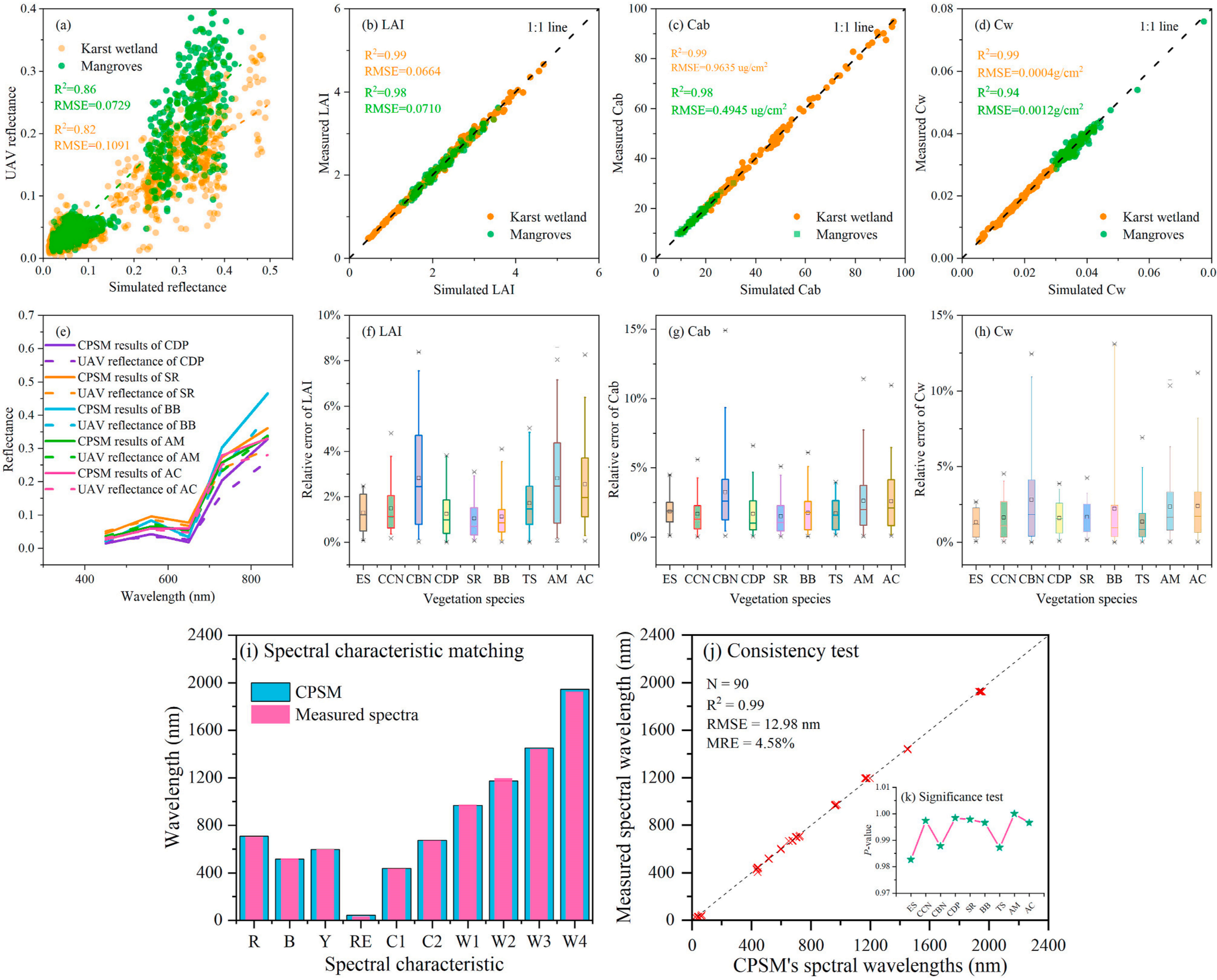


Fig. 3. Accuracy validation of spectra matching results. (a) Is the consistency of spectral reflectance between spectral reflectance of CPSM and the UAV spectral reflectance. (b) ~ (d) are the accuracies of constrained parameters during spectra matching. (e) are the matching results of the spectral reflectance of UAV images with the CPSM's reflectance. (f) ~ (h) are the relative errors between measured vegetation parameters and simulated parameters; (i) ~ (k) are the verification of spectral characteristic similarity between the CPSM spectra and the measured leaf spectra; R, B, and Y represent the position of red edge, blue edge and yellow edge, respectively; RE is the red edge width; C1 and C2 are the chlorophyll absorbing valley 1 and 2, respectively; W1, W2, W3, and W4 are the water-absorbing valley 1, 2, 3, and 4, respectively.

Cab, *Car* and *Ant* are sensitive to canopy reflectance only at visible light (400 ~ 760 nm) (Fig. 4b, c, d), but their sensitivity to canopy reflectance was very weak (<1 %) after 768 nm, 552 nm and 614 nm. Furthermore, *Protein* was least sensitive to the canopy reflectance of each species, and only response to the reflectance of 1433 ~ 2500 nm (Fig. 4f), with an average contribution of 0.003 % ~ 0.20 %. This might be the reason why the retrieval accuracy of CNC using only UAV multispectral images is always relatively low, i.e., the response spectral range between canopy reflectance and *Protein* content did not fall into the band range of UAV multispectral images (434 ~ 866 nm). In addition, the vegetation reflectance was dominated by *Cw* after approximately 1300 nm, with sensitivity of 0.42 % ~ 6.41 %. Although the response trends of biophysical and biochemical parameters to the canopy reflectance of different species were almost same, the response degree were significantly different, which might result in differences in the sensitivity spectrum of CNC and CPC for each vegetation species. Therefore, we selected the three physiological traits (*Cab*, *Cw*, and *LAI*) with the greatest influence in the visible, near-infrared, and short-wave infrared ranges as the constraints condition for spectral matching.

4.2. Sensitive spectra and optimal retrieval features of CNC and CPC

4.2.1. Sensitive spectral subdomains of CNC and CPC

We plotted the correlation between spectral features and dependent variables of each species (Fig. 5). The highest correlation between the original bands of UAV images and CNC, CPC reached −0.56 and 0.52, respectively. We found that there were significant differences in sensitive bands of UAV images among different species, and even some correlations were extremely low, such as CCN's and AC's CNC, BB's and AM's CPC. Compared with the original bands of UAV images, we confirmed that the sensitive spectral subdomains of CNC and CPC among most of the species were covered by the CPSM-based spectra matching data, and their correlation was significantly improved by 0.11 ~ 0.41 and 0.14 ~ 0.64, respectively. It is worth noting that most of the canopy spectral bands were negatively correlated with CNC and CPC of each species, and their sensitive spectral subdomains are mainly in the range of 1380 ~ 2400 nm, and were rarely concentrated in the visible bands (540 ~ 730 nm). This result was consistent with the response ranges of protein content, CNC, and CPC to canopy spectral reflectance (Figs. 4 and 5).

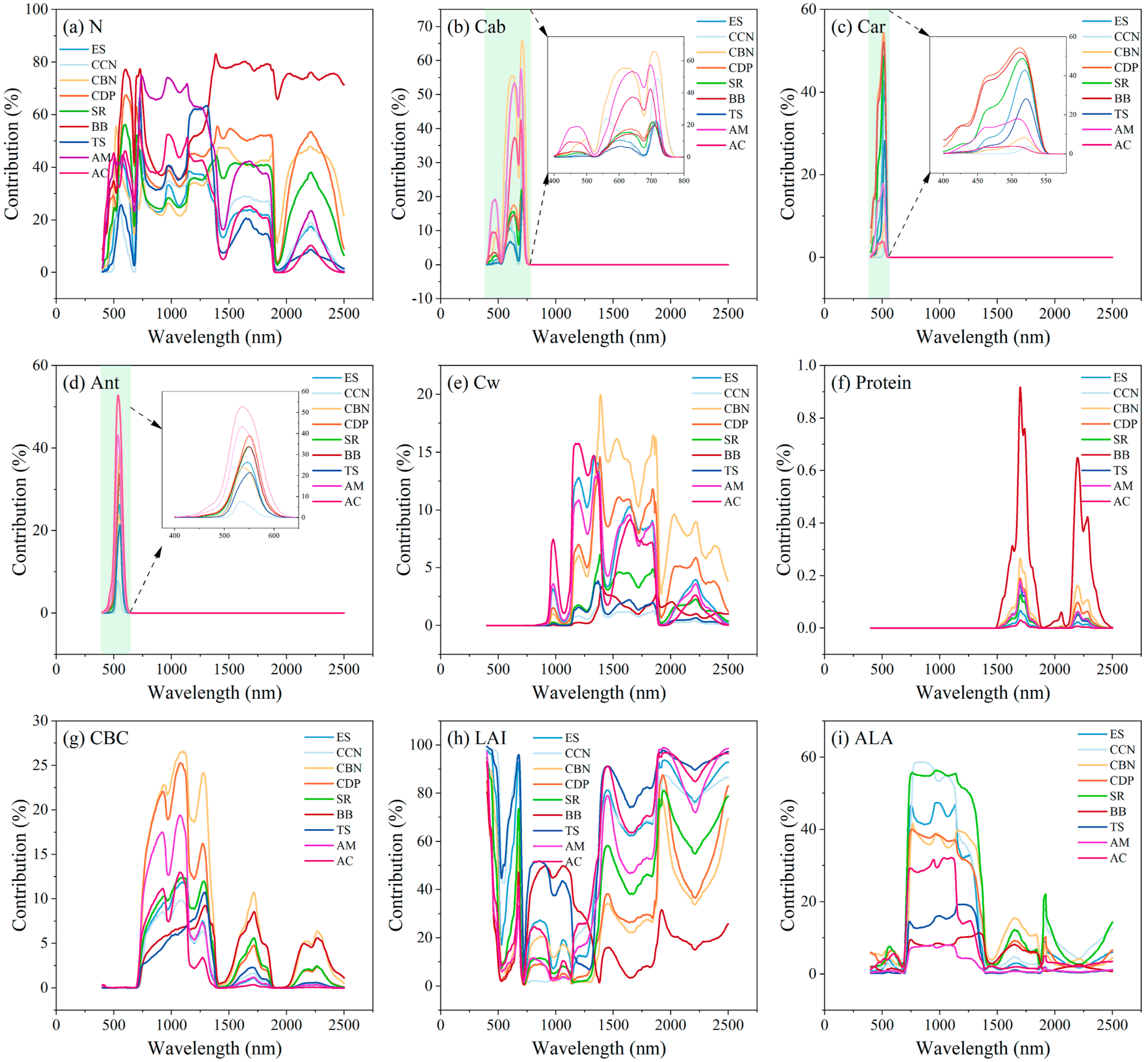


Fig. 4. Global sensitivity of biophysical and biochemical parameters of vegetation species to canopy spectral reflectance at 400 ∼ 2500 nm. *N*: leaf structure parameters; *Cab*: leaf chlorophyll content; *Car*: carotenoids; *Ant*: anthocyanin; *Cw*: equivalent water thickness; *CBC*: carbon-based constituents; *LAI*: leaf area index; *ALA*: Average leaf angle.

Besides, the correlation between the matched spectra obtained by CPSM and CNC of ES and TS, as well as the correlation between the matched spectra and CPC of BB and TS, reached over 0.8 and 0.7, respectively. The above results have confirmed that the CPSM has great potential to obtain the spectral features of estimating vegetation physiological trait parameters in both karst wetland and mangroves.

Through combination of the correlation heat map (Fig. 5) and the spectral characteristics (Fig. 3i) of the whole species, we found that a few spectral subdomains were sensitive to CNC and CPC near the position of red edge (701 ∼ 725 nm) and most of sensitive subdomains were distributed near the water absorption valley (1170 nm, 1450 nm, and 1940 nm). Meanwhile, this study also found that the sensitivity of spectral reflectance to CNC and CPC would change with the spectral characteristics of species: compared with other species, water absorption valley 2 of ES shifted towards longer wavelengths by 26 ∼34 nm, resulting in a shift of the sensitive spectral subdomains of CNC and CPC towards shorter wavelengths. The appearance of sensitive spectral information in visible light bands and near the water absorption valley was due to the presence of proteins in photosynthetic pigments and the dry matter content, respectively. Above results indicated that the analysis of the spectral characteristics for species was the basis for understanding the spectral sensitivity of CNC and CPC.

We summarized the sensitive spectral subdomain (Table 1) of each species extracted from CPSM data and their corresponding spectral bands of UAV images based on the correlation thresholds. We found that only a few of the bands (less than three) in the sensitive spectral subdomains covered the bandwidth of UAV multispectral images, which indicated that it is possible to estimate CNC and CPC of some vegetation species using only UAV multispectral images (Table 2). For example, the sensitive spectral subdomains of CPC of SR distributed in 401 ∼ 490 nm,

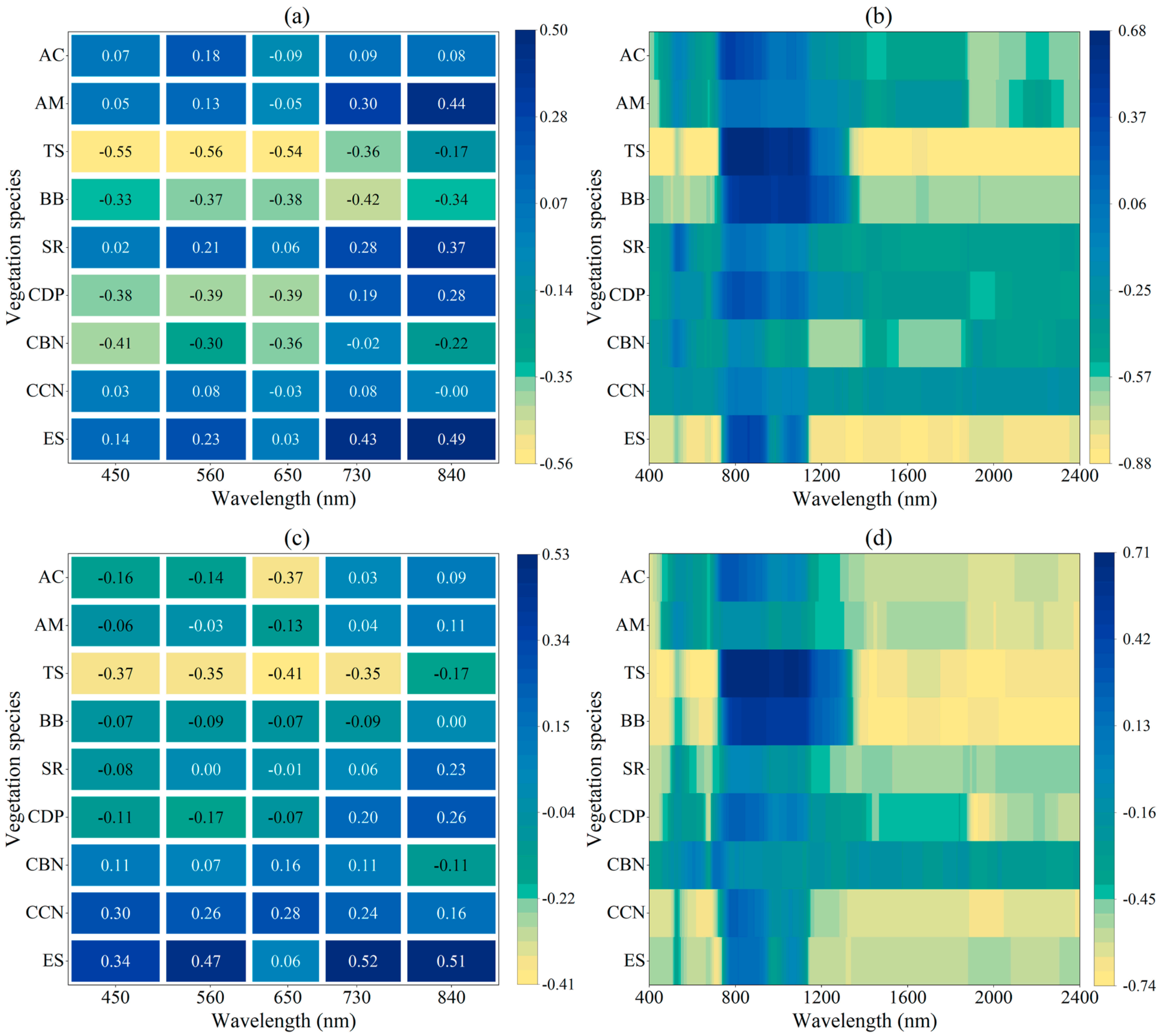


Fig. 5. Pearson's correlation (r) between spectral features and dependent variables (CNC and CPC). (a) and (c) are the correlation between the original bands of P4M UAV multispectral images and CNC, CPC, respectively. (b) and (d) are the correlation between 401 ~ 2400 nm spectrum data obtained from CPSM and CNC, CPC, respectively.

and 658 ~ 687 nm, which fell into the blue and red bands of the UAV multispectral images, respectively. However, UAV multispectral images cannot estimate the CNC and CPC of all species. For example, the sensitive spectral subdomains of CPC for BB were not covered by UAV multispectral images.

4.2.2. *Optimal spectral features of CNC and CPC for each species*

We used the fractional order derivative (FOD) method and spectral indexes to mine the derivative features with the higher correlation with CNC and CPC from the CPSM-based spectral data and extracted their optimal spectral index features for each species (Fig. 6 and Table S9 in Supplemental Materials 1). It was obviously visible from Fig. 6j, k, l that the absolute correlation coefficients (|r|) between CNC/CPC and FOD spectral features and spectral indexes (SIs) were both significantly higher than the original CPSM-based spectrum, and the correlation coefficient increased by 0.04 ~ 0.28 and 0.06 ~ 0.37, respectively. The increased correlations indicate that the FOD method could extract sensitive spectral features from the CPSM-based spectral data, and thus improve estimation accuracy of CNC and CPC. The absolute correlation coefficients of SIs were 0.15 higher than FOD spectral features. In addition, we found that only some the optimal FOD features fell into the sensitive subdomains of CNC and CPC (Fig. 6a ~ i), and that the sensitive spectral subdomains become more discrete with the increase of the order of FOD spectrum, and their correlation would gradually improve. Although the spectral bands with the highest correlation with CNC and CPC were quite different, the FOD spectral bands with the highest sensitivity were almost all obtained in the derivative transformation of order 1.0 to 2.0 (Tables S10, S11 and S12 in Supplemental Materials 1).

The FOD features, which were finally inputted into the estimated model of CNC and CPC for each species, were located near the position of the red edge and water absorption valley (Table S7 in Supplemental Materials 1 and Fig. 6a ~ i), demonstrating that the sensitive spectra of CNC and CPC obscured by the water content and pigments could be extracted by FOD method, which confirmed the feasibility of CPSM in identifying the sensitive bands of CNC and CPC in different species. Although the spectral indexes calculated using the discrete FOD-based sensitive spectral bands were highly sensitive in estimating CNC and CPC of each species, we found that the low-correlation spectral bands through calculating spectral indexes could also provide spectral features sensitive to CNC and CPC (Table S9 in Supplemental Materials 1). Therefore, the spectral bands with low correlation should not be discarded when calculating the spectral indexes.

Table 1
Summary of UAV optimal features and sensitive spectral subdomain (SSS) of CNC and CPC. |r| is the absolute correlation coefficient.

Traits	Vegetation species	CPSM-based SSS (nm)	Original UAV bands ⊆ SSS and its derived features		The number of optimal derived features ⊆ SSS
			Original bands (nm)	Optimal derived features	
CNC	ES	574 ~ 655 688 ~ 723 1154 ~ 1883 2068 ~ 2299 ($\lvert r\rvert > 0.78$)	Red edge (714 ~ 723)	RDSI$_{(5,1,4,1)}$, RDSI$_{(5,2,5,4)}$, RDSI$_{(5,4,3,1)}$, RDSI$_{(5,4,2,1)}$, RDSI$_{(4,3,2,1)}$, NLI$_{(5,1,1,3)}$ NLI$_{(5,3,4,2)}$	6
	CCN	401 ~ 506 585~ 694 1147 ~ 1265 1283 ~ 1441 1463 ~ 1670 1792 ~ 1873 ($\lvert r\rvert > 0.3$)	Blue (434 ~ 466) Red (634 ~ 666)	DSI$_{(3,2)}$, FRSI-2$_{(3,2,5,4)}$, INDSI$_{(5,1,4,3)}$, INDSI$_{(5,2,4,3)}$, INDSI$_{(5,3,4,1)}$, INDSI$_{(5,4,4,3)}$, NLI$_{(2,4,5,3)}$	7
	CBN	1144 ~ 1381 1633 ~ 1662 ($\lvert r\rvert > 0.6$)	None	FRSI-2$_{(5,2,5,1)}$, FRSI-2$_{(5,3,5,1)}$, TRSI-1$_{(3,2,1)}$, NLI$_{(2,4,4,5)}$, NLI$_{(3,1,1,4)}$, NLI$_{(3,1,5,5)}$, NLI$_{(3,5,1,2)}$, NLI$_{(3,5,5,5)}$, NLI$_{(5,3,5,2)}$, NLI$_{(5,5,1,1)}$,	0
	CDP	401 ~ 460 666 ~ 684 1416 ~ 1492 1870 ~ 2400 ($\lvert r\rvert > 0.4$)	Blue (434 ~ 460)	FRSI-2$_{(5,1,2,1)}$, FRSI-2$_{(2,1,5,2)}$, INDSI$_{(4,3,2,1)}$, RDSI$_{(2,1,5,3)}$	4
	SR	401 ~ 453 660 ~ 687 1307 ~ 1399 1501 ~ 1872 1892 ~ 1998 ($\lvert r\rvert > 0.42$)	Blue (434 ~ 453) Red (660 ~ 666)	FRSI-2$_{(3,2,3,1)}$, INDSI$_{(2,1,2,1)}$, RDSI$_{(2,1,2,1)}$, NLI$_{(1,2,2,1)}$, NLI$_{(2,1,4,2)}$, NLI$_{(3,1,1,2)}$, NLI$_{(4,3,4,2)}$, NLI$_{(4,3,5,2)}$, NLI$_{(5,1,1,2)}$,	9
	BB	492 ~ 518 549 ~ 603 1465 ~ 1570 1776 ~ 1877 2042 ~ 2400 ($\lvert r\rvert > 0.65$)	Green (549 ~ 576)	RDSI$_{(4,1,5,1)}$, NLI$_{(1,5,5,3)}$, NLI$_{(2,5,5,2)}$, NLI$_{(3,1,4,3)}$, NLI$_{(3,5,4,3)}$, NLI$_{(4,2,3,3)}$, NLI$_{(5,5,4,1)}$,	2
	TS	401 ~ 511 565 ~ 713 1378 ~ 2400 ($\lvert r\rvert > 0.8$)	Blue (434 ~ 466) Green (565 ~ 576) Red (634 ~ 666)	Blue, TRSI-1$_{(4,2,3)}$, TRSI-1$_{(4,3,2)}$, TRSI-1$_{(4,3,4)}$, FRSI-1$_{(5,2,5,1)}$, FRSI-1$_{(5,2,5,4)}$, FRSI-1$_{(3,2,3,1)}$, INDSI$_{(5,2,4,3)}$, INDSI$_{(5,4,2,1)}$, RDSI$_{(4,1,5,1)}$, RDSI$_{(2,1,5,4)}$, NLI$_{(1,4,2,4)}$, NLI$_{(2,1,2,3)}$,	12
	AM	401 ~ 449 1883~ 2092 2300 ~ 2400 ($\lvert r\rvert > 0.55$)	Blue (434 ~ 449)	RDSI$_{(5,4,3,1)}$, RDSI$_{(5,4,3,2)}$, RDSI$_{(5,4,2,1)}$, NLI$_{(5,3,3,3)}$	2
	AC	401 ~ 429 1429 ~ 1469 1876 ~ 2197 2226 ~ 2400 ($\lvert r\rvert > 0.55$)	None	NLI$_{(1,3,2,3)}$, NLI$_{(2,3,1,3)}$, NLI$_{(2,3,2,3)}$,	0
CPC	ES	574 ~ 664 685~ 733 1146 ~ 1883 2063 ~ 2339 ($\lvert r\rvert > 0.54$)	Green (574 ~ 576) Red (634 ~ 666) Red edge (714 ~ 733)	Green, DSI$_{(2,1)}$, RDSI$_{(5,2,3,2)}$, RDSI$_{(5,2,2,1)}$, RDSI$_{(4,3,5,1)}$, RDSI$_{(3,2,2,1)}$, RDSI$_{(2,1,2,1)}$ NLI$_{(2,1,1,4)}$	7
	CCN	401 ~ 512 553 ~ 710 1174 ~ 2400 ($\lvert r\rvert > 0.50$)	Blue (434 ~ 466) Green (553 ~ 576) Red (634 ~ 666)	Green, RDSI$_{(4,2,3,1)}$, RDSI$_{(3,1,3,2)}$, NLI$_{(1,2,1,2)}$	4
	CBN	401 ~ 451 1898 ~1970 2244 ~ 2373 ($\lvert r\rvert > 0.3$)	Blue (434 ~ 451)	TRSI-2$_{(3,2,4)}$, FRSI-2$_{(3,1,4,2)}$, FRSI-1$_{(4,3,4,1)}$, RDSI$_{(3,2,3,2)}$, NLI$_{(2,2,4,5)}$, NLI$_{(5,5,4,1)}$	2
	CDP	401 ~ 543 665 ~ 684 1883 ~ 2173 2245 ~ 2400 ($\lvert r\rvert > 0.5$)	Blue (434 ~ 466) Red (634 ~ 666)	FRSI-2$_{(5,1,2,1)}$, FRSI-2$_{(5,4,2,1)}$, NLI$_{(4,3,3,5)}$, NLI$_{(4,3,5,3)}$	4
	SR	401 ~ 490 658 ~ 687 1321 ~ 1384 1541 ~ 1853 1892 ~ 1899 1921 ~ 1995 ($\lvert r\rvert > 0.5$)	Blue (434 ~ 466) Red (658 ~ 666)	TRSI-2$_{(5,4,1)}$, TRSI-1$_{(5,4,1)}$, FRSI-1$_{(5,3,4,1)}$, NLI$_{(2,4,5,4)}$, NLI$_{(3,1,5,2)}$, NLI$_{(5,2,3,3)}$, NLI$_{(5,4,5,3)}$, NLI$_{(5,5,1,1)}$,	7
	BB	1476 ~ 1573 1780 ~ 1872	None	TRSI-2$_{(2,1,5)}$, FRSI-2$_{(5,4,3,2)}$, FRSI-2$_{(4,2,5,4)}$, INDSI$_{(5,2,5,2)}$, INDSI$_{(5,4,4,2)}$, RDSI$_{(4,2,5,2)}$, NLI$_{(5,5,4,1)}$	0

(*continued on next page*)

Table 1 (*continued*)

Traits	Vegetation species	CPSM-based SSS (nm)	Original UAV bands $\subseteq$ SSS and its derived features		The number of optimal derived features $\subseteq$ SSS
			Original bands (nm)	Optimal derived features	
		2046 ∼ 2400 ($\lvert r\rvert > 0.70$)			
	TS	612 ∼ 662 686 ∼ 697 774 ∼ 918 ($\lvert r\rvert > 0.7$)	Red (634 ∼ 662) NIR (814 ∼ 866)	TRSI-1$_{(5,2,4)}$, INDSI$_{(5,2,4,1)}$, NLI$_{(1,5,4,1)}$, NLI$_{(4,5,3,4)}$ NLI$_{(4,5,4,2)}$, NLI$_{(4,5,4,5)}$, NLI$_{(5,5,4,1)}$, NLI$_{(5,5,4,2)}$, NLI$_{(5,5,4,4)}$,	9
	AM	401 ∼ 439 1413 ∼ 1496 1871 ∼ 2400 ($\lvert r\rvert > 0.55$)	Blue (434 ∼ 439)	NLI$_{(1,3,2,1)}$, NLI$_{(2,3,1,1)}$, NLI$_{(2,3,2,1)}$	3
	AC	401 ∼ 431 1405 ∼ 1522 1665 ∼ 1669 1675 ∼ 1812 1855 ∼ 2400 ($\lvert r\rvert > 0.55$)	None	NLI$_{(1,2,2,3)}$, NLI$_{(2,2,2,3)}$, NLI$_{(2,2,3,3)}$, NLI$_{(3,2,2,3)}$, NLI$_{(3,2,3,3)}$,	0

4.3. Evaluating performance of CNC and CPC based on new hybrid retrieval strategy

4.3.1. Our hybrid retrieval strategy VS traditional hybrid retrieval strategy

We evaluated the retrieval performance of our hybrid retrieval strategy in estimating CNC of all vegetation species by comparing its estimated results with the traditional hybrid retrieval strategy (Fig. 7). We found that the estimated accuracy of CNC based on scheme 4 (traditional hybrid retrieval strategy) was significantly lower than that based on scheme 2 (P4M + FOD) and scheme 3 (P4M + FOD + Indexes).

Table 2

Summary of the estimated results of CNC and CPC for seven vegetation species under different schemes.

Vegetation species	Algorithms	Metrics	CNC			CPC		
			Non-CPSM	CPSM		Non-CPSM	CPSM	
			Scheme 1	Scheme 2	Scheme 3	Scheme 1	Scheme 2	Scheme 3
ES	PLSR	R^2	0.02	0.88	0.90	0.09	0.09	0.86
		RMSE	0.000640	0.000211	0.000179	1.21	1.21	0.47
		MRE	44.29 %	19.53 %	15.06 %	29.12 %	29.12 %	17.62 %
	AELR	R^2	0.12	0.91	0.94	0.70	0.87	0.91
		RMSE	0.000534	0.000170	0.000151	0.90	0.57	0.38
		MRE	40.10 %	15.88 %	13.07 %	19.64 %	12.92 %	12.63 %
CCN	PLSR	R^2	0.51	0.67	0.70	0.01	0.34	0.69
		RMSE	0.000305	0.000262	0.000237	0.56	0.48	0.33
		MRE	26.56 %	24.91 %	18.17 %	39.81 %	34.10 %	27.65 %
	AELR	R^2	0.50	0.80	0.76	0.02	0.86	0.91
		RMSE	0.000314	0.000218	0.000212	0.63	0.25	0.21
		MRE	26.39 %	17.32 %	15.96 %	34.11 %	17.78 %	17.27 %
CBN	PLSR	R^2	0.71	0.73	0.73	0.04	0.04	0.46
		RMSE	0.000242	0.000201	0.000204	1.83	1.83	1.26
		MRE	19.17 %	14.00 %	14.12 %	41.57 %	41.57 %	28.85 %
	AELR	R^2	0.77	0.77	0.83	0.02	0.63	0.72
		RMSE	0.000205	0.000196	0.000184	1.74	1.02	0.90
		MRE	14.45 %	13.54 %	11.24 %	31.44 %	26.25 %	15.83 %
CDP	PLSR	R^2	0.21	0.45	0.49	0.24	0.50	0.70
		RMSE	0.000440	0.000370	0.000347	2.09	1.65	1.49
		MRE	22.53 %	22.83 %	17.52 %	37.91 %	31.02 %	26.66 %
	AELR	R^2	0.54	0.60	0.71	0.64	0.74	0.77
		RMSE	0.000331	0.000328	0.000281	1.46	1.22	1.18
		MRE	21.18 %	18.35 %	17.53 %	23.92 %	20.22 %	19.14 %
SR	PLSR	R^2	0.57	0.70	0.88	0.61	0.48	0.76
		RMSE	0.000210	0.000199	0.000105	0.66	0.63	0.57
		MRE	18.42 %	17.58 %	14.70 %	20.47 %	18.75 %	17.26 %
	AELR	R^2	0.85	0.88	0.93	0.71	0.79	0.89
		RMSE	0.000146	0.000093	0.000080	0.48	0.39	0.32
		MRE	14.42 %	12.91 %	9.84 %	15.45 %	12.73 %	10.21 %
BB	PLSR	R^2	0.10	0.80	0.88	0.07	0.20	0.77
		RMSE	0.000535	0.000254	0.000226	1.49	1.48	0.72
		MRE	52.89 %	21.62 %	18.47 %	50.20 %	35.00 %	18.17 %
	AELR	R^2	0.20	0.93	0.96	0.33	0.83	0.98
		RMSE	0.000512	0.000204	0.000137	1.17	0.61	0.36
		MRE	43.87 %	13.67 %	11.02 %	30.93 %	13.61 %	5.91 %
TS	PLSR	R^2	0.32	0.80	0.88	0.21	0.63	0.89
		RMSE	0.000235	0.000164	0.000107	0.97	0.97	0.58
		MRE	44.08 %	30.21 %	16.57 %	51.23 %	42.33 %	40.14 %
	AELR	R^2	0.54	0.93	0.94	0.29	0.75	0.97
		RMSE	0.000191	0.000079	0.000078	0.93	0.58	0.25
		MRE	24.11 %	8.74 %	7.61 %	33.62 %	21.51 %	9.58 %

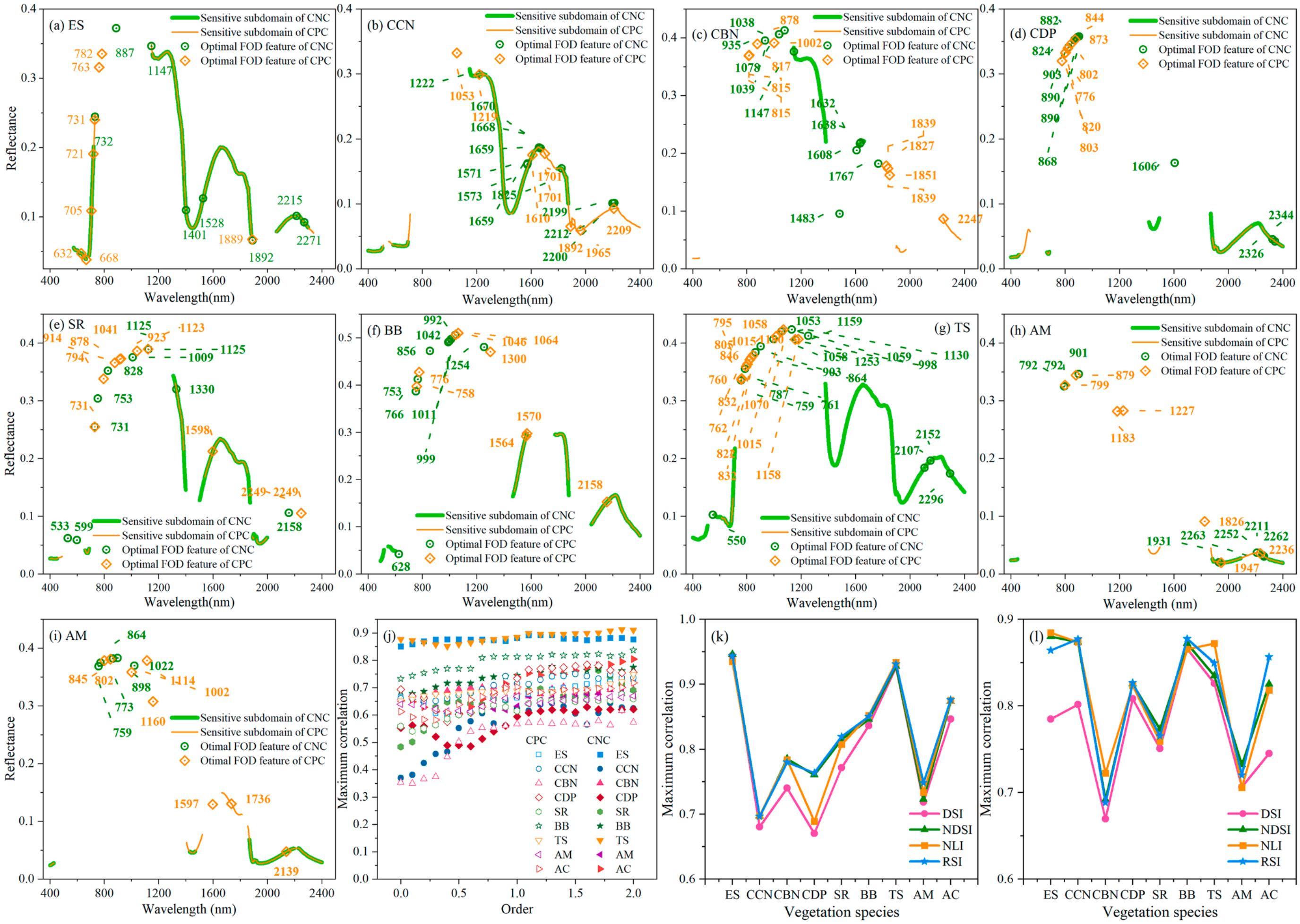


Fig. 6. Response relationship between CNC, CPC and spectral features. (a)–(i) are sensitive spectral subdomains and optimal FOD features of vegetation species; (j) is the highest correlation between CNC, CPC and FOD spectrum, respectively, where the order = 0 represents the original spectrum. (k) and (l) are the highest correlation between spectral indexes and CNC, CPC, respectively.

Compared with scheme 4, the RMSE and MRE of scheme 2 and scheme 3 decreased by 0.000200 ~ 0.000248 g/cm^2 and 8.64 % ~ 20.24 %, respectively, demonstrating that our new hybrid retrieval strategy (scheme 2 and scheme 3) was significantly superior to the traditional hybrid retrieval strategy (scheme 4). In addition, we found that the estimated accuracy of CNC using only UAV multispectral images (scheme 1) was higher than that of the traditional hybrid retrieval strategy (scheme 4). This might be because the spectral data simulated by PROSAIL-PRO model had significant differences in the original P4M spectral data, resulting in the low generalization ability of retrieval model constructed by scheme 4. The above results indicate that our hybrid retrieval strategy has great advantages for estimating CNC of vegetation in karst wetland. These results also confirm that the new hybrid retrieval strategy proposed in this study is obviously superior to the traditional hybrid retrieval strategy.

4.3.2. Retrieving CNC and CPC of vegetation species using different schemes

It can be seen from Fig. 7 that the estimation model of CNC constructed by all samples did not achieve satisfactory results. Therefore, we estimated the CNC and CPC of seven vegetation species in karst wetland and presented the estimation results and accuracy metrics (Table 2) under three schemes. We found that the estimated accuracy of CNC at the species scale was better than that of all vegetation, and the MRE decreased by 8.24 % ~ 18.68 %. Meanwhile, the hybrid model successfully achieved CPC estimation for the first time, and the R^2 and MRE were distributed at 0.46 ~ 0.98 and 5.91 % ~ 40.14 %, respectively. In scheme 1 (only P4M multispectral data), the R^2 and MRE of estimating CNC were concentrated at 0.02 ~ 0.85 and 14.42 % ~ 52.89 %, respectively, where the estimation accuracy of SR was the highest (R^2 = 0.85, RMSE = 0.000146 g/cm^2, MRE = 14.42 %). Although the CNC of CBN and SR and the CPC of ES and SR both obtain satisfactory estimated accuracy in scheme 1 (R^2 = 0.70 ~ 0.85, MRE = 14.42 % ~ 19.64 %), the estimated accuracies of CNC and CPC for other species were still low (R^2 = 0.01 ~ 0.64, MRE = 21.18 % ~ 52.89 %), indicating that the estimation of CPC of all vegetation species in karst wetland cannot be realized by using only UAV multispectral images.

The estimated accuracies of CNC and CPC for the seven species in Scheme 2 (P4M + FOD) were almost all higher than those in scheme 1 (P4M), with the MRE decreasing by 0.84 % ~ 31.27 % (Table 3). It is noteworthy that CNC and CPC obtained the highest estimation accuracy (R^2 = 0.46 ~ 0.98, MRE = 5.91 % ~ 40.14 %) in scheme 3 (P4M + FOD + Indexes), with MRE 0.29 % ~ 16.83 % lower than that in Scheme 2 (P4M + FOD), which further confirmed the feasibility of combining CPSM with the new hybrid retrieval strategy to retrieve CNC and CPC of different species and indicated that the spectral indexes could provide effective spectral information that was different from FOD spectral features for further improving estimated accuracies of CNC and CPC. The above results confirm that combining CPSM with new hybrid retrieval strategy could achieve the high-precision retrieval of CNC and CPC for vegetation species and indirectly prove that accurate analysis of the spectral characteristics of species is the prerequisite for identifying

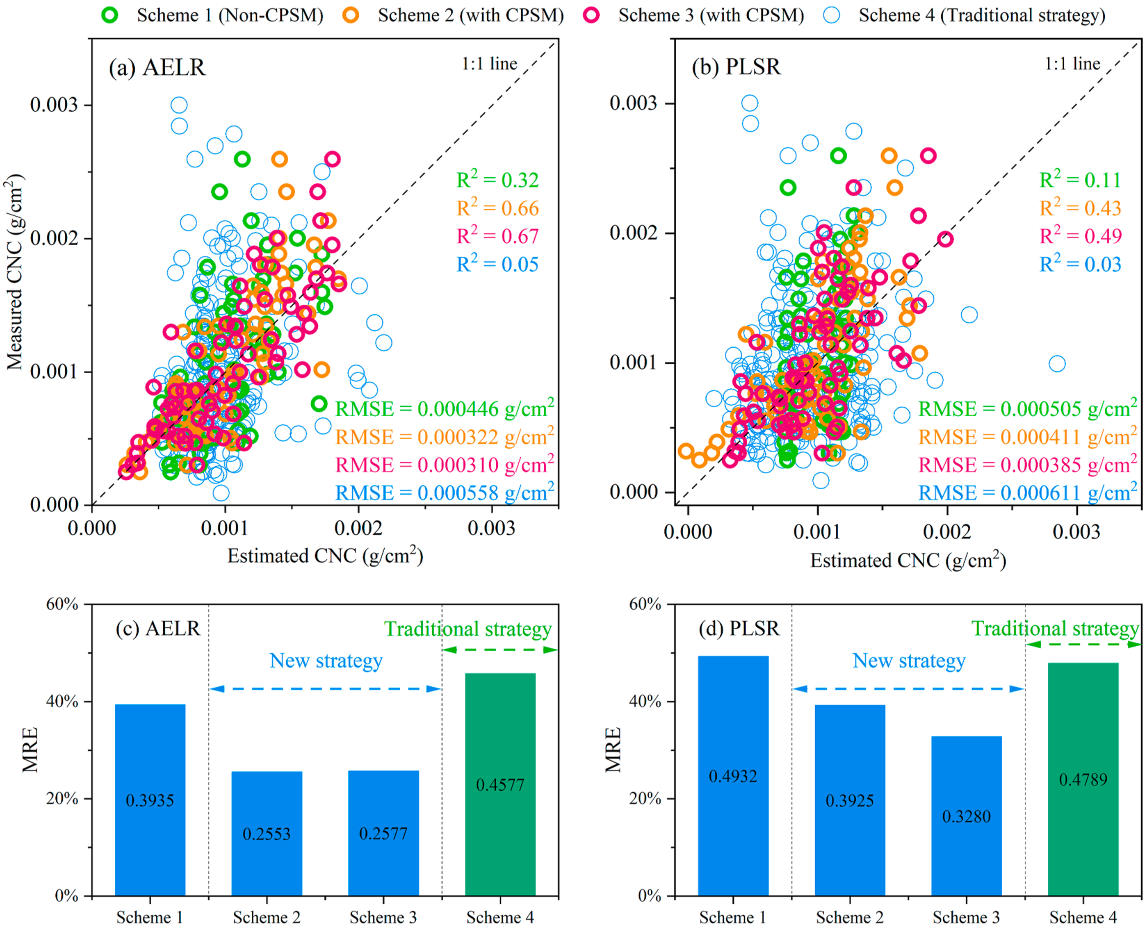


Fig. 7. Comparison of retrieval results between new and traditional retrieval strategies. (a) and (b) were retrieval results using AELR, PLSR algorithms, respectively. (c) and (d) were MRE of retrieval results using AELR and PLSR algorithms, respectively; scheme 1 only included UAV multispectral image features. Scheme 2 and scheme 3 were used for new hybrid retrieval strategy; Scheme 4 was used by traditional hybrid retrieval strategy.

Table 3
Estimated results of CNC and CPC by combining the PLSR, ML, and AELR algorithms with new hybrid retrieval strategy.

Vegetation species	Algorithms	CNC				CPC			
		Scheme 2		Scheme 3		Scheme 2		Scheme 3	
		R^2	MRE	R^2	MRE	R^2	MRE	R^2	MRE
ES	PLSR	0.88	19.53 %	0.90	15.06 %	0.09	29.12 %	0.86	17.62 %
	ML	0.89	20.89 %	0.96	11.68 %	0.78	12.96 %	0.88	16.05 %
	AELR	0.91	15.88 %	0.94	13.07 %	0.87	12.92 %	0.91	12.63 %
CCN	PLSR	0.67	24.91 %	0.70	18.17 %	0.34	34.10 %	0.69	27.65 %
	ML	0.72	17.55 %	0.61	19.24 %	0.78	17.90 %	0.85	20.65 %
	AELR	0.80	17.32 %	0.76	15.96 %	0.86	17.78 %	0.91	17.27 %
CBN	PLSR	0.73	14.00 %	0.73	14.12 %	0.04	41.57 %	0.46	28.85 %
	ML	0.73	13.70 %	0.80	14.04 %	0.48	28.52 %	0.62	22.15 %
	AELR	0.77	13.54 %	0.83	11.24 %	0.63	26.25 %	0.72	15.83 %
CDP	PLSR	0.45	22.83 %	0.49	17.52 %	0.50	31.02 %	0.70	26.66 %
	ML	0.47	21.71 %	0.77	16.38 %	0.70	19.65 %	0.75	17.44 %
	AELR	0.60	18.35 %	0.71	17.53 %	0.74	20.22 %	0.77	19.14 %
SR	PLSR	0.70	17.58 %	0.88	14.70 %	0.48	18.75 %	0.76	17.26 %
	ML	0.90	12.99 %	0.94	12.41 %	0.76	13.27 %	0.85	11.81 %
	AELR	0.88	12.91 %	0.93	9.84 %	0.79	12.73 %	0.89	10.21 %
BB	PLSR	0.80	21.62 %	0.88	18.47 %	0.20	35.00 %	0.77	18.17 %
	ML	0.88	15.00 %	0.87	12.68 %	0.68	17.12 %	0.91	12.46 %
	AELR	0.93	13.67 %	0.96	11.02 %	0.83	13.61 %	0.98	5.91 %
TS	PLSR	0.80	30.21 %	0.88	16.57 %	0.63	42.33 %	0.89	40.14 %
	ML	0.93	11.18 %	0.95	10.99 %	0.57	30.49 %	0.95	12.80 %
	AELR	0.88	19.53 %	0.90	15.06 %	0.09	29.12 %	0.86	17.62 %

sensitive spectral features to achieve CNC and CPC estimation.

We found that the PLSR algorithm obtained the same accuracy when estimating the CNC and CPC of some vegetation species under Scheme 1 (P4M) and scheme 2 (P4M + FOD), such as when estimating the CPC for

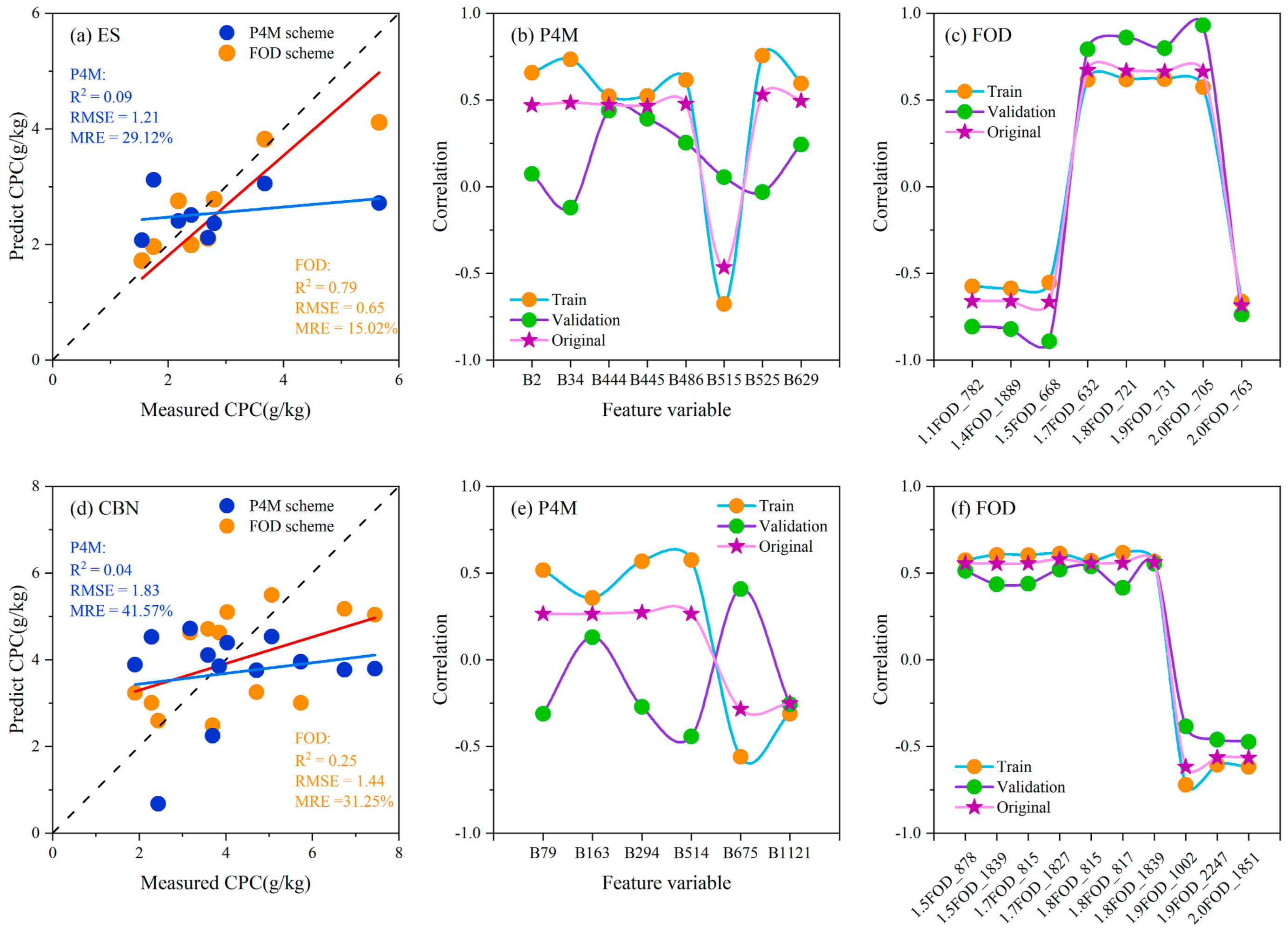


Fig. 8. Comparison of CPC estimation results and the Pearson's correlation (between CPC and spectral variables) for ES and CBN under P4M and FOD scheme. (a) and (d) were the estimation results of CPC for ES and CBN under the P4M and FOD schemes, respectively. (b), (c), (e) and (f) are the Pearson's correlation between CPC and the spectral variables in the original, training and validation datasets under the P4M and FOD schemes.

ES and CBN (Table 2). In order to clarify the reason why the PLSR algorithm produced the same result after adding FOD spectral features, we separated scheme 2 and compared its performance in retrieving CPC of ES and CBN using only P4M or FOD spectral features (Fig. 8). It is worth noting that the estimated accuracies of CPC using only FOD spectral features were higher than when using only P4M features (Fig. 8a, d) and P4M + FOD features (Table 2), and its RMSE and MRE decreased by 0.56 g/kg and 0.39 g/kg, 14.10 % and 10.32 %, respectively. On the one hand, it might be that the correlation between the original spectral features of P4M and CPC was too low ($-0.47 < r < 0.53$), especially in CBN ($-0.29 < r < 0.27$), which affected the performance of FOD features in the PLSR algorithm. On the other hand, the difference in correlation (between spectral features and CPC) between the training dataset and the validation dataset was too large, resulting in the high heterogeneity of spectral information (Fig. 8b, e). For example, some features were positively correlated with CPC in the training dataset but were negatively correlated with CPC in the validation dataset. The above results also illustrate that it is more important to pay attention to the effect of feature differentiation on the estimated accuracies when constructing the model.

4.3.3. Evaluating robustness and generalization of CNC and CPC models

A leave-one-out cross-validation (LOOCV) approach was employed to assess the robustness and generalization capacity of estimation models developed with limited sample sizes (Fig. 9). Two models exhibited notable predictive accuracy across small-sample datasets ($R^2 = 0.60 \sim 0.95$, MRE = 8 % ~ 24 %), indicating that even constrained environmental sampling can yield reliable CNC and CPC estimation frameworks. Besides, for each species, 75 % of the estimated CNC and CPC exhibited relative errors below 30 %, with only a minimal proportion exceeding 40 %, underscoring the methodological consistency of the small-sample models. These results validate the reliability and generalizability of the proposed models under data-limited conditions.

4.3.4. Retrieval performance of different algorithms

We summarized the estimated results of PLSR, ML, and AELR algorithms in scheme 2 and scheme 3 (Table 3). Detailed information about the accuracy evaluation of all ML algorithms' estimation results is shown in Fig. S4 and Table S8 of Supplementary Materials 1. We found that the AELR algorithm proposed in this study almost obtained the highest accuracies for estimating CNC ($R^2 = 0.60 \sim 0.96$, RMSE = 0.000078 ~ 0.000328 g/cm^2, MRE = 7.61 % ~ 18.35 %) and CPC ($R^2 = 0.72 \sim 0.98$, RMSE = 0.21 ~ 1.18 g/kg, MRE = 10.21 % ~ 19.94 %), and its MRE was 1.60 % ~ 30.56 % lower than PLSR and ML algorithms, indicating that AELR algorithm has a great advantages in estimating CNC and CPC of vegetation species. In general, in scheme 2 and scheme 3, the estimation accuracies of CNC and CPC for the seven species based on ML algorithm were almost higher than those of using the PLSR algorithm, and the MRE decreased by 0.08 % ~ 27.34 %. However, ML algorithm was more prone to overfitting (Fig. S4h, p and the models

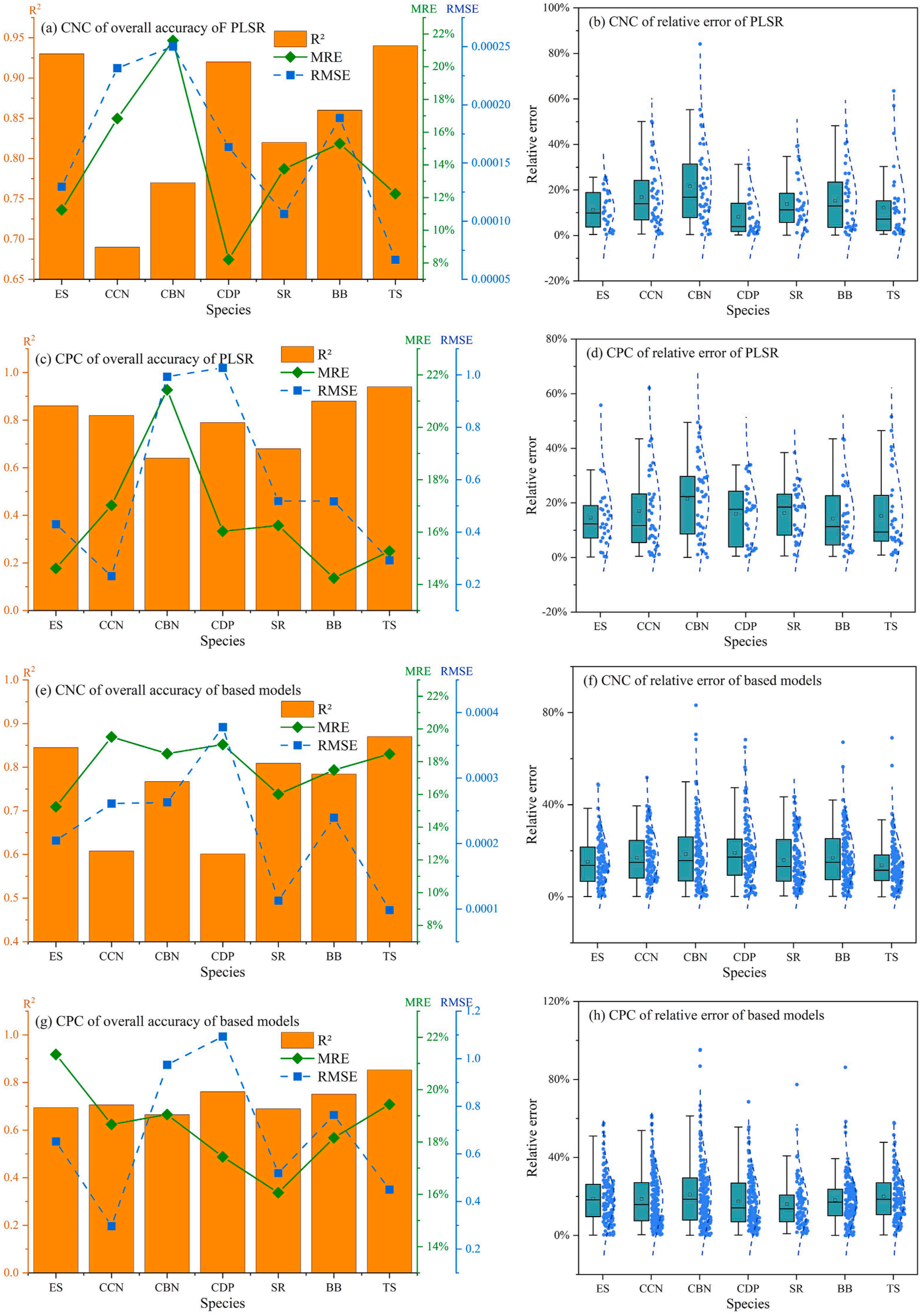


Fig. 9. The robustness and generalization of the CNC and CPC estimated models. (a) (c), (e), (g) are the overall estimated accuracies of CNC and CPC using PLSR model and based models of AELR, respectively; (b), (d), (f), (h) are the relative errors of every sample of CNC and CPC, respectively.

corresponding to the bold font in Table S8 of Supplementary Materials 1).

In this study, we further evaluated the retrieval performance of ML algorithms in estimating CNC and CPC of the seven species based on scheme 3. The optimal ML algorithms for estimating CNC of ES, CCN, CBN, CDP, SR, BB, and TS were Decision Tree, CatBoost, GradientBoost, XGBoost, Decision Tree, GradientBoost, and Random Forest, respectively. Their MRE of CNC reduced by 0.25 % ~ 19.22 %. The GradientBoost algorithm obtained the highest accuracies ($R^2 = 0.85 \sim 0.95$, RMSE = 0.33 ~ 0.58 g/kg, MRE = 11.83 % ~ 12.80 %) for estimating CPC of SR, BB, and TS. The MRE of CPC estimation for CBN and CDP using the XGB algorithm was 5.21 % ~ 8.63 % lower than that of the others ML algorithms. The MRE of CPC estimation of SR, BB, TS using the GradientBoost algorithm was at most 12.60 % lower than that of the others ML algorithms. In summary, the AELR algorithm outperformed the ML and PLSR algorithms in estimating CNC and CPC of the seven vegetation species in karst wetland.

4.4. Transfer performance of CPSM and hybrid retrieval strategy for estimating CNC and CPC of mangroves

We utilized CPSM and our hybrid strategy to estimate CNC and CPC of mangrove species to evaluate their performance for different wetland schemes (Fig. 10). The CNC and CPC of *Aegiceras corniculatum* (*AC*) and *Avicennia marina* (*AM*) have obtained high precision results ($R^2 = 0.77 \sim 0.89$, MRE = 9.65 % ~ 16.87 %), regardless of whether we used PLSR or AELR algorithms. These results proved that combining CPSM with hybrid retrieval strategy is feasible for CNC and CPC estimation of mangroves. We found that the AELR algorithm performed better than PLSR algorithm for estimating CNC and CPC, which was consistent with the conclusion in section 4.3.4, with a decreasing in MRE ranging from 0.23 % to 4.43 %. The estimated CNC and CPC values of *AC* and *AM* were almost distributed in the error ellipse of the 95 % confidence interval, and they concentrated in the ranges of 0.000825 ~ 0.00502 g/cm^2 and 1.27 ~ 7.77 g/kg, respectively, which were consistent with the measured values (0.000642 ~ 0.00473 g/cm^2 and 0.91 ~ 9.04 g/kg).

5. Discussion

5.1. Effectiveness of spectral sensitivity enhancement using FOD

The spectral response of nitrogen and phosphorus presents greater challenges compared to chlorophyll or water content due to their subtle absorption features, which may be obscured by overlapping spectral signatures from co-occurring biochemical constituents [42]. This study addresses this limitation by employing FOD analysis to isolate diagnostically sensitive spectral regions for CNC and CPC in the visible light region dominated by pigments and the short-wave near infrared region dominated by water (Fig. 6 and Table S7 in Supplemental Materials 1), especially peak and valley, indicating that FOD method can eliminate the spectral overlapping caused by the existence of similar molecules or ions between chlorophyll and nitrogen. Meanwhile, our study proved that FOD-processed CPSM spectra and their derived vegetation indices exhibited significantly stronger correlations with CNC and CPC than the original spectral data (Fig. S5 in Supplemental Materials 1). These findings remained consistent across species-level analyses (Fig. 6i, k, l), aligning with recent evidence that FOD techniques amplify spectral sensitivity to vegetation biochemical traits [43]. Notably, FOD's capacity to normalize structural variability, such as interspecific differences in canopy architecture and bark-mediated photon scattering [44], proved critical for robust nutrient estimation. By emphasizing absorption feature morphology over absolute reflectance, FOD minimized confounding effects from physical canopy properties. This methodological approach accounted for 79 % of observed CPC variability in validation trials (Fig. 8a). Furthermore, species-scale CNC and CPC prediction accuracies (Fig. 7 and Table 2) quantitatively confirmed FOD's efficacy in compensating for biophysical interference, thereby advancing the precision of hyperspectral nutrient retrieval.

5.2. Spectral response to nitrogen and phosphorous

The optimal sensitive spectral features of CNC and CPC identified in this study predominantly localize to the short-wave near infrared region and a small part in the visible light region (Fig. 6 and Table S7 in Supplemental Materials 1). This distribution aligns with nitrogen's biochemical role in chlorophyll and protein synthesis, wherein photosynthetic pigments primarily govern visible-light reflectance through their absorption properties [45]. Furthermore, prior research has established that about 12 minor absorption features of protein and nitrogen caused by the N-H bond stretches have been identified [46], such as 1020 nm, 1510 nm, 1690 nm, 1940 nm, 1980 nm, 2180 nm, etc. This is consistent with our findings of the optimal spectral features of CNC in short-wave near infrared region. In contrast, phosphorus detection remains largely confined to empirical reflectance-correlation analyses [44], with limited mechanistic exploration of spectral responses compared to nitrogen, cellulose, or water. Previous studies have proved that the initial effect of phosphorous stress on crop was an increase in the number of smaller cells per unit of leaf area compared with a non-stressed plant (Jacob and Lawlor, 1991), which amplifies the near-infrared reflectance due to altered internal leaf scattering [47]. This mechanism likely explains the prominence of CPC-sensitive features at 800~1300 nm in our results (Fig. 6). Besides, phosphorus stress elevates anthocyanin synthesis [48], which exhibits strong absorption in green wavelengths (550 nm) and reflectance in blue/red regions [49], consistent with the CPC feature distribution near green reflectance peaks and red absorption troughs (Fig. 6e and f). While the physiological basis of photosynthetic pigment-reflectance relationships is well-studied, advancing the spectral interpretation of phosphorus, potassium, sulfur, and barium dynamics remains critical for robust biophysical trait modeling.

5.3. Advantage and limitation of our CPSM

The constrained PROSAIL-PRO spectra matching (CPSM) approach and the hybrid retrieval strategy were combined to achieve the estimation of vegetation CNC and CPC using multispectral images, which greatly reduces the high cost of purchasing hyperspectral sensors for monitoring vegetation health and provides a new idea for accurately monitoring other vegetation physiological trait parameters. The greatest advantage of CPSM is that it extends the UAV multispectral reflectance to 400 ~ 2500 nm, and its continuous narrow bands amplifies the sensitive features of vegetation trait parameters. Meanwhile, CPSM utilizes other traits for constraints to perform an inversion of the observed reflectance, which project the observed multispectral images into the look-up table space and ensures a consistent response relationship between spectral reflectance and target trait. However, there are still some limitations. On the one hand, this method requires vegetation physiological trait parameters as prerequisites to match simulated spectral curves at 400 ~ 2500 nm. These constrained parameters (*LAI*, *Cab*, *Cw*, etc.) have been successfully estimated through multispectral images [50,51]. Currently, *LAI* products can be obtained quickly through airborne LiDAR sensors [52]. Furthermore, Fang et al. [53] also summarized the mainstream *LAI* remote sensing products. On the other hand, for ultra-high-spatial-resolution images, the mapping of the spatial distribution of vegetation physiological traits will spend a lot of time on the spectra matching. Therefore, this method may be more suitable for retrieving vegetation physiological traits using moderate-to-high-resolution multispectral satellite images such as Landsat-8/9 OLI, Sentinel-2 A/B, SPOT-5/6, ZiYuan-3, and so on.

5.4. Advantage and limitation of hybrid retrieval model

The radiative transfer model was essentially limited by the ill-posed inversion problem [54], that is, different input parameters might generate similar spectra, resulting in the estimated accuracy of vegetation physiological traits parameters not meeting the requirements. Therefore, most current studies have coupled the radiative transfer model with machine learning (ML) or deep learning (DL) algorithms to form a hybrid model for estimating vegetation physiological traits. This hybrid model strategy takes advantage of simulated data with physical mechanisms generated from radiative transfer model and the data-driven ML or DL algorithms. It not only expands the training sample but also makes the model learn additional physical laws of the radiative transfer model from the simulated data. The primary distinction between our hybrid framework and conventional approaches lies in

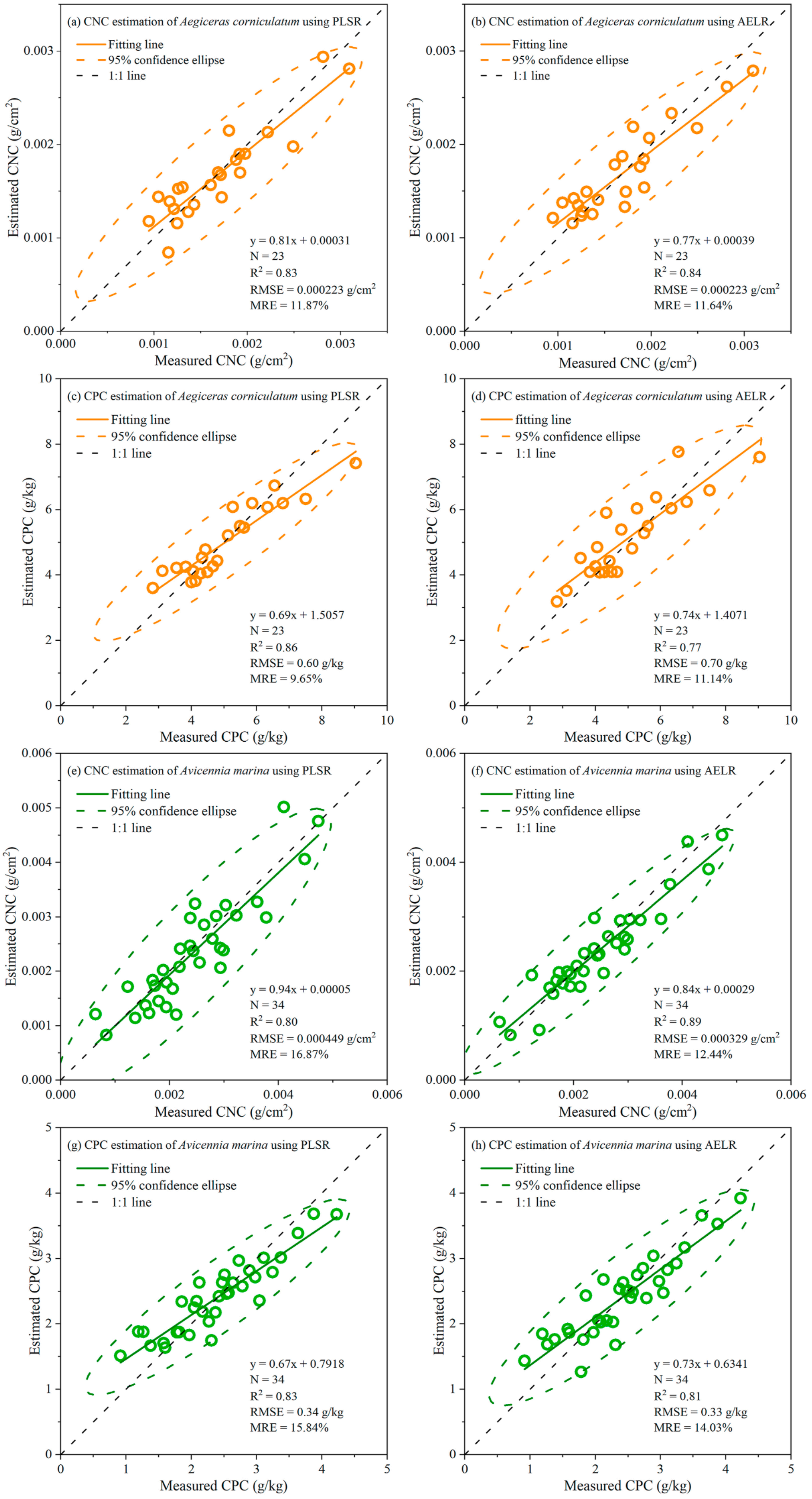


Fig. 10. Estimation results of mangrove CNC and CPC in scheme 3 using PLSR and ALER algorithms.

the integration of simulated spectral data with trait measurements during model training. This synthesis enables simulated spectra to break through the limitation of the virtual scene, thereby facilitating the identification of complex nonlinear relationships between spectral signatures and *in situ* vegetation biochemical properties. For the traditional hybrid retrieval strategy, the spectral response function of different remote sensing sensors could achieve the project of the simulated spectrum. However, the consistency between the remote sensing sensor and the simulated reflectance data could not be guaranteed. Furthermore, due to the actual complex growth environment, there are great differences between the measured vegetation physiological traits and the data simulated by the radiative transfer models under an ideal environment, resulting in the low generalization ability of hybrid model. Therefore, Gaussian noise was introduced during the simulation phase to emulate the variability in reflectance inherent to satellite imaging conditions, while concurrently enhancing the hybrid model's generalizability for functional trait parameter estimation [55]. Moreover, hybrid model efficacy depends on targeted optimization of radiative transfer model parameters governing target-specific sensitive spectral bands, guided by domain-specific prior knowledge [56]. For example, the spectral ranges that affect CNC estimation were at the 670 ~ 890 nm and 900 ~ 2400 nm. It is necessary to consider optimizing the parameters that have a great influence on the canopy reflectance in these spectral ranges, such as *N*, *Cw*, and *LAI* (Fig. S6 in Supplemental Materials 1). Then, the parameters that have no or little effect on the canopy reflectance in the sensitive spectral ranges, such as *Car*, *Ant*, solar zenith angle (*SZA*), and observer zenith angle (*OZA*), can be fixed.

In addition, the radiative transfer models and traditional hybrid retrieval strategy cannot directly estimate the content of phosphorus, potassium, and carbon in vegetation. This is because radiative transfer models do not yet take into account these physiological traits. However, the new hybrid retrieval strategy proposed in this study can provide a possibility for the estimation of vegetation physiological traits such as phosphorus, potassium and carbon contents. This strategy uses measured vegetation physiological trait parameters, and multispectral image reflectance to match simulated spectra, eliminating the process of optimizing the parameters of radiation transfer model. These parameters used as constraint can be existing estimation products. This strategy also conquers the limitations of the traditional hybrid retrieval strategy, which can only use the simulated spectra of the inherent band of the sensor, i.e., this strategy can use the reflectance spectral curve of 400 ~ 2500 nm with a spectral resolution of 1 nm. Moreover, this strategy not only solves the problem of low estimation accuracy caused by data inconsistency in the retrieval process, but also expands the spectral information of vegetation for multispectral sensors, so that multispectral images can be widely used in the estimation of vegetation physiological traits.

6. Conclusion

In this study, we proposed a novel PROSAIL-PRO spectra matching (CPSM) approach and a new hybrid retrieval strategy to realize the high-precision estimation of the canopy nitrogen content (CNC) and canopy phosphorus content (CPC) for vegetation species in multiple wetlands in China. We confirmed that our CPSM could extend the spectral reflectance of UAV multispectral images to 400 ~ 2500 nm, maintain good spectral consistency with the original bands both in karst wetland and mangroves (Karst wetland: $R^2 = 0.82$, RMSE = 0.1091; Mangroves: $R^2 = 0.86$, RMSE = 0.0729), and also delineate similar spectral properties at the red, blue, and yellow edge position, chlorophyll absorbing valleys (about 435 nm and 670 nm), and water-absorbing valleys (about 970 nm and 1450 nm). This study firstly analyzed the spectral characteristics of seven typical vegetation species in karst wetland, and demonstrated that the optimal estimation features of CNC and CPC always be found near red edge position and water-absorbing valley in the vegetation spectrum. Our findings can provide a spectral basis for photosynthesis, pigment formation, phototropism, and the morphogenesis of vegetation, and can provide scientific references for the protection, restoration, and sustainable development of the wetland ecological environment. Besides, our results proved that our hybrid retrieval strategy outperformed the traditional hybrid retrieval strategy, which achieved the high-precision estimation of CNC and CPC for seven species in karst wetland using AELR algorithm ($R^2 = 0.71 \sim 0.98$, MRE = 5.91 % ~ 19.14 %). Finally, combining CPSM with our hybrid retrieval strategy was successfully applied to estimate CNC and CPC of two mangrove species ($R^2 = 0.77 \sim 0.89$, MRE = 9.65 % ~ 16.87 %), confirming the ability and applicability of transferring these methods to other wetlands, which further indicated that combining CPSM and our strategy can provide a new idea for precisely estimating vegetation physiological traits using multispectral sensors.

This study offers a reliable method and a scientific basis for building a robust estimation model of vegetation nutrient components through multispectral images in large-scale wetlands. In future studies, we will continue to verify the feasibility of combining CPSM and our strategy for other vegetation physiological traits in different wetland types.

CRediT authorship contribution statement

Zhinan Lao: Conceptualization, Data curation, Methodology, Software, Writing – original draft, review & editing. Bolin Fu: Conceptualization, Funding acquisition, Writing – review & editing. Weiwei Sun: Funding acquisition, Writing - review & editing. Yeqiao Wang: Writing – review & editing. Yuyu Zhou: Writing – review & editing. Hongchang He: Supervision, Funding acquisition. Tengfang Deng: Supervision. Ertao Gao: Supervision.

Data availability

The example of CPSM code and data can be accessed via https://github.com/RSFubolin/Python_CPSM.git.

Declaration of competing interest

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

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant number 42371341), the Natural Science Foundation of Guangxi Zhuang Autonomous Region (CN) (Grant number 2025GXNSFFA069008; 2024GXNSFAA010351), Key Laboratory of Tropical Marine Ecosystem and Bioresource Ministry of Natural Resources (Grant number 2023ZD02), Zhejiang Province "Pioneering Soldier" and "Leading Goose" R&D Project (Grant number 2023C01027).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plaphe.2025.100059.

References

[1] R.P. Powers, G.J. Hay, G. Chen, How wetland type and area differ through scale: a GEOBIA case study in Alberta's Boreal Plains, Rem. Sens. Environ. 117 (2012) 135–145, https://doi.org/10.1016/j.rse.2011.07.009.

[2] H. Wen, E. Liu, S. Yan, J. Chang, N. Xiao, J. Zhou, Conserving karst cavefish diversity in Southwest China, Biol. Conserv. 273 (2022) 109680, https://doi.org/10.1016/j.biocon.2022.109680.

[3] R.C. Gardner, C. Finlayson, Global Wetland Outlook: State of the World's Wetlands and Their Services to People, vols. 2020–5, Ramsar Convention Secretariat, 2018.

[4] B. Fu, F. Lan, H. Yao, J. Qin, H. He, L. Liu, L. Huang, D. Fan, E. Gao, Spatio-temporal monitoring of marsh vegetation phenology and its response to hydro-

meteorological factors using CCDC algorithm with optical and SAR images: in case of Honghe National Nature Reserve, China, Sci. Total Environ. 843 (2022) 156990, https://doi.org/10.1016/j.scitotenv.2022.156990.

[5] K.J. Niklas, T. Owens, P.B. Reich, E.D. Cobb, Nitrogen/phosphorus leaf stoichiometry and the scaling of plant growth, Ecol. Lett. 8 (2005) 636–642, https://doi.org/10.1111/j.1461-0248.2005.00759.x.

[6] S. Güsewell, U. Bollens, P. Ryser, F. Klötzli, Contrasting effects of nitrogen, phosphorus and water regime on first- and second-year growth of 16 wetland plant species, Funct. Ecol. 17 (2003) 754–765, https://doi.org/10.1111/j.1365-2435.2003.00784.x.

[7] W. Koerselman, A.F.M. Meuleman, The vegetation N:P ratio: a new tool to detect the nature of nutrient limitation, J. Appl. Ecol. 33 (1996) 1441, https://doi.org/10.2307/2404783.

[8] P.B. Reich, Key canopy traits drive forest productivity, Proc. Biol. Sci. 279 (2012) 2128–2134, https://doi.org/10.1098/rspb.2011.2270.

[9] Z. Wang, A.K. Skidmore, R. Darvishzadeh, T. Wang, Mapping forest canopy nitrogen content by inversion of coupled leaf-canopy radiative transfer models from airborne hyperspectral imagery, Agric. For. Meteorol. 253–254 (2018) 247–260, https://doi.org/10.1016/j.agrformet.2018.02.010.

[10] M.E. Martin, L.C. Plourde, S.V. Ollinger, M.-L. Smith, B.E. McNeil, A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems, Rem. Sens. Environ. 112 (2008) 3511–3519, https://doi.org/10.1016/j.rse.2008.04.008.

[11] M.A. Cho, J. van Aardt, R. Main, B. Majeke, Evaluating variations of physiology-based hyperspectral features along a soil water gradient in a Eucalyptus grandis plantation, Int. J. Rem. Sens. 31 (2010) 3143–3159, https://doi.org/10.1080/01431160903154390.

[12] A. Ramoelo, A.K. Skidmore, M. Schlerf, R. Mathieu, I.M.A. Heitkönig, Water-removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations, ISPRS J. Photogrammetry Remote Sens. 66 (2011) 408–417, https://doi.org/10.1016/j.isprsjprs.2011.01.008.

[13] M.A. Cho, Hyperspectral remote sensing of biochemical and biophysical parameters: the derivative red-edge "double-peak feature". A Nuisance or an Opportunity?, Wageningen University and Research, 2007.

[14] J. He, Q. Yuan, J. Li, Y. Xiao, D. Liu, H. Shen, L. Zhang, Spectral super-resolution meets deep learning: achievements and challenges, Inf. Fusion 97 (2023) 101812, https://doi.org/10.1016/j.inffus.2023.101812.

[15] W. Verhoef, Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model, Rem. Sens. Environ. 16 (1984) 125–141, https://doi.org/10.1016/0034-4257(84)90057-9.

[16] X. Zou, R. Hernández-Clemente, P. Tammeorg, C. Lizarazo Torres, F.L. Stoddard, P. Mäkelä, P. Pellikka, M. Mõttus, Retrieval of leaf chlorophyll content in field crops using narrow-band indices: effects of leaf area index and leaf mean tilt angle, Int. J. Rem. Sens. 36 (2015) 6031–6055, https://doi.org/10.1080/01431161.2015.1110262.

[17] R.F. Kokaly, G.P. Asner, S.V. Ollinger, M.E. Martin, C.A. Wessman, Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies, Rem. Sens. Environ. 113 (2009) S78–S91, https://doi.org/10.1016/j.rse.2008.10.018.

[18] R. Kokaly, Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression, Rem. Sens. Environ. 67 (1999) 267–287, https://doi.org/10.1016/s0034-4257(98)00084-4.

[19] K.F. Palmer, D. Williams, Optical properties of water in the near infrared, J. Opt. Soc. Am. 64 (1974) 1107–1110, https://doi.org/10.1364/JOSA.64.001107.

[20] A.R. Huete, A soil-adjusted vegetation index (SAVI), Rem. Sens. Environ. 25 (1988) 295–309, https://doi.org/10.1016/0034-4257(88)90106-x.

[21] B. Fu, S. Li, Z. Lao, B. Yuan, Y. Liang, W. He, W. Sun, H. He, Multi-sensor and multi-platform retrieval of water chlorophyll a concentration in karst wetlands using transfer learning frameworks with ASD, UAV, and Planet CubeSate reflectance data, Sci. Total Environ. 901 (2023) 165963, https://doi.org/10.1016/j.scitotenv.2023.165963.

[22] B.-B. Guo, Y.-L. Feng, C. Ma, J. Zhang, X. Song, M.-Y. Wang, D.-H. Sheng, W. Feng, N. Jiao, Suitability of different multivariate analysis methods for monitoring leaf N accumulation in winter wheat using in situ hyperspectral data, Comput. Electron. Agric. 198 (2022) 107115, https://doi.org/10.1016/j.compag.2022.107115.

[23] R. Zhao, L. An, W. Tang, D. Gao, L. Qiao, M. Li, H. Sun, J. Qiao, Deep learning assisted continuous wavelet transform-based spectrogram for the detection of chlorophyll content in potato leaves, Comput. Electron. Agric. 195 (2022) 106802, https://doi.org/10.1016/j.compag.2022.106802.

[24] J. Jin, Q. Wang, Hyperspectral indices developed from the low order fractional derivative spectra can capture leaf dry matter content across a variety of species better, Agric. For. Meteorol. 322 (2022) 109007, https://doi.org/10.1016/j.agrformet.2022.109007.

[25] S.S. Kharintsev, M.K. Salakhov, A simple method to extract spectral parameters using fractional derivative spectrometry, Spectrochim. Acta Mol. Biomol. Spectrosc. 60 (2004) 2125–2133, https://doi.org/10.1016/j.saa.2003.11.013.

[26] N.C. Coops, M.-L. Smith, M.E. Martin, S.V. Ollinger, Prediction of eucalypt foliage nitrogen content from satellite-derived hyperspectral data, IEEE Trans. Geosci. Rem. Sens. 41 (2003) 1338–1346, https://doi.org/10.1109/tgrs.2003.813135.

[27] M. Soltanikazemi, S. Minaei, H. Shafizadeh-Moghadam, A. Mahdavian, Field-scale estimation of sugarcane leaf nitrogen content using vegetation indices and spectral bands of Sentinel-2: application of random forest and support vector regression, Comput. Electron. Agric. 200 (2022) 107130, https://doi.org/10.1016/j.compag.2022.107130.

[28] J. Gao, B. Meng, T. Liang, Q. Feng, J. Ge, J. Yin, C. Wu, X. Cui, M. Hou, J. Liu, H. Xie, Modeling alpine grassland forage phosphorus based on hyperspectral remote sensing and a multi-factor machine learning algorithm in the east of Tibetan Plateau, China, ISPRS J. Photogrammetry Remote Sens. 147 (2019) 104–117, https://doi.org/10.1016/j.isprsjprs.2018.11.015.

[29] R. Darvishzadeh, C. Atzberger, A. Skidmore, M. Schlerf, Mapping grassland leaf area index with airborne hyperspectral imagery: a comparison study of statistical approaches and inversion of radiative transfer models, ISPRS J. Photogrammetry Remote Sens. 66 (2011) 894–906, https://doi.org/10.1016/j.isprsjprs.2011.09.013.

[30] D. Zhao, J. Zhen, Y. Zhang, J. Miao, Z. Shen, X. Jiang, J. Wang, J. Jiang, Y. Tang, G. Wu, Mapping mangrove leaf area index (LAI) by combining remote sensing images with PROSAIL-D and XGBoost methods, Remote Sensing in Ecology and Conservation 9 (2022) 370–389, https://doi.org/10.1002/rse2.315.

[31] B. Chen, X. Mu, P. Chen, B. Wang, J. Choi, H. Park, S. Xu, Y. Wu, H. Yang, Machine learning-based inversion of water quality parameters in typical reach of the urban river by UAV multispectral data, Ecol. Indic. 133 (2021) 108434, https://doi.org/10.1016/j.ecolind.2021.108434.

[32] L. Liang, L. Di, L. Zhang, M. Deng, Z. Qin, S. Zhao, H. Lin, Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method, Rem. Sens. Environ. 165 (2015) 123–134, https://doi.org/10.1016/j.rse.2015.04.032.

[33] Jun Li, H. Xie, Jianhong Li, G. Yang, Y. Xie, J. Wang, C. Zhou, S. Zou, Influences of anthropogenic acids on carbonate weathering and CO2 sink in an agricultural karst wetland (South China), Ecol. Indic. 150 (2023) 110192, https://doi.org/10.1016/j.ecolind.2023.110192.

[34] Z. Lao, B. Fu, Y. Wei, T. Deng, W. He, Y. Yang, H. He, E. Gao, Retrieval of chlorophyll content for vegetation communities under different inundation frequencies using UAV images and field measurements, Ecol. Indic. 158 (2024) 111329, https://doi.org/10.1016/j.ecolind.2023.111329.

[35] A.R. Kotlash, B.C. Chessman, Effects of water sample preservation and storage on nitrogen and phosphorus determinations: implications for the use of automated sampling equipment, Water Res. 32 (1998) 3731–3737, https://doi.org/10.1016/s0043-1354(98)00145-6.

[36] J. Murphy, J.P. Riley, A modified single solution method for the determination of phosphate in natural waters, Anal. Chim. Acta 27 (1962) 31–36, https://doi.org/10.1016/s0003-2670(00)88444-5.

[37] H.-H. Yeoh, Y.-C. Wee, Leaf protein contents and nitrogen-to-protein conversion factors for 90 plant species, Food Chem. 49 (1994) 245–250, https://doi.org/10.1016/0308-8146(94)90167-8.

[38] G. Candiani, G. Tagliabue, C. Panigada, J. Verrelst, V. Picchi, J.P. Rivera Caicedo, M. Boschetti, Evaluation of hybrid models to estimate chlorophyll and nitrogen content of maize crops in the framework of the future CHIME mission, Remote Sens. 14 (2022) 1792, https://doi.org/10.3390/rs14081792.

[39] M.D. Raya-Sereno, M. Alonso-Ayuso, J.L. Pancorbo, J.L. Gabriel, C. Camino, P. J. Zarco-Tejada, M. Quemada, Residual effect and N fertilizer rate detection by high-resolution VNIR-SWIR hyperspectral imagery and solar-induced chlorophyll fluorescence in wheat, IEEE Trans. Geosci. Rem. Sens. 60 (2022) 1–17, https://doi.org/10.1109/tgrs.2021.3099624.

[40] A. Simic, J.M. Chen, T.L. Noland, 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, Int. J. Rem. Sens. 32 (2011) 5621–5644, https://doi.org/10.1080/01431161.2010.507257.

[41] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Comput. Phys. Commun. 181 (2010) 259–270, https://doi.org/10.1016/j.cpc.2009.09.018.

[42] N.M. Knox, A.K. Skidmore, H.H.T. Prins, G.P. Asner, H.M.A. van der Werff, W.F. de Boer, C. van der Waal, H.J. de Knegt, E.M. Kohi, R. Slotow, R.C. Grant, Dry season mapping of savanna forage quality, using the hyperspectral Carnegie Airborne Observatory sensor, Rem. Sens. Environ. 115 (2011) 1478–1488, https://doi.org/10.1016/j.rse.2011.02.007.

[43] A. Zhang, S. Yin, J. Wang, N. He, S. Chai, H. Pang, Grassland chlorophyll content estimation from drone hyperspectral images combined with fractional-order derivative, Remote Sens. 15 (2023) 5623, https://doi.org/10.3390/rs15235623.

[44] J.G. Ferwerda, A.K. Skidmore, Can nutrient status of four woody plant species be predicted using field spectrometry? ISPRS J. Photogrammetry Remote Sens. 62 (2007) 406–414, https://doi.org/10.1016/j.isprsjprs.2007.07.004.

[45] E.B. Knipling, Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation, Rem. Sens. Environ. 1 (1970) 155–159, https://doi.org/10.1016/s0034-4257(70)80021-9.

[46] P.J. Curran, Remote sensing of foliar chemistry, Rem. Sens. Environ. 30 (1989) 271–278, https://doi.org/10.1016/0034-4257(89)90069-2.

[47] T. Lillesand, R.W. Kiefer, J. Chipman, Remote Sensing and Image Interpretation, Wiley, 2003.

[48] H. Marchner, Mineral Nutrition on Higher Plants, second ed., Academic Press, New York, 1995.

[49] N.M. Hughes, W.K. Smith, Attenuation of incident light in Galax urceolata (Diapensiaceae): concerted influence of adaxial and abaxial anthocyanic layers on photoprotection, Am. J. Bot. 94 (2007) 784–790, https://doi.org/10.3732/ajb.94.5.784.

[50] B. Qian, H. Ye, W. Huang, Q. Xie, Y. Pan, N. Xing, Y. Ren, A. Guo, Q. Jiao, Y. Lan, A sentinel-2-based triangular vegetation index for chlorophyll content estimation, Agric. For. Meteorol. 322 (2022) 109000, https://doi.org/10.1016/j.agrformet.2022.109000.

[51] J. Estévez, M. Salinero-Delgado, K. Berger, L. Pipia, J.P. Rivera-Caicedo, M. Wocher, P. Reyes-Muñoz, G. Tagliabue, M. Boschetti, J. Verrelst, Gaussian processes retrieval of crop traits in Google Earth Engine based on Sentinel-2 top-of-

atmosphere data, Rem. Sens. Environ. 273 (2022) 112958, https://doi.org/10.1016/j.rse.2022.112958.

[52] Y. Wang, H. Fang, Y. Zhang, S. Li, Y. Pang, T. Ma, Y. Li, Retrieval and validation of vertical LAI profile derived from airborne and spaceborne LiDAR data at a deciduous needleleaf forest site, GIScience Remote Sens. 60 (2023), https://doi.org/10.1080/15481603.2023.2214987.

[53] H. Fang, F. Baret, S. Plummer, G. Schaepman-Strub, An overview of global leaf area index (LAI): methods, products, validation, and applications, Rev. Geophys. (1985) 57 (2019) 739–799, https://doi.org/10.1029/2018rg000608.

[54] B. Combal, F. Baret, M. Weiss, A. Trubuil, D. Macé, A. Pragnère, R. Myneni, Y. Knyazikhin, L. Wang, Retrieval of canopy biophysical variables from bidirectional reflectance: using prior information to solve the ill-posed inverse problem, Rem. Sens. Environ. 84 (2003) 1–15, https://doi.org/10.1016/s0034-4257(02)00035-4.

[55] J. Zhu, J. Lu, W. Li, Y. Wang, J. Jiang, T. Cheng, Y. Zhu, W. Cao, X. Yao, Estimation of canopy water content for wheat through combining radiative transfer model and machine learning, Field Crops Res. 302 (2023) 109077, https://doi.org/10.1016/j.fcr.2023.109077.

[56] L. Wan, Y. Liu, Y. He, H. Cen, Prior knowledge and active learning enable hybrid method for estimating leaf chlorophyll content from multi-scale canopy reflectance, Comput. Electron. Agric. 214 (2023) 108308, https://doi.org/10.1016/j.compag.2023.108308.