frontiers | Frontiers in Ecology and Evolution

TYPE Original Research
PUBLISHED 27 March 2023
DOI 10.3389/fevo.2023.1139458

OPEN ACCESS

EDITED BY
Zhouyuan Li,
Beijing Forestry University,
China

REVIEWED BY
Shihua Li,
University of Electronic Science and Technology of China,
China
Zhihui Wang,
Guangzhou Institute of Geography,
China

*CORRESPONDENCE
Yuan Zeng
✉ zengyuan@aircas.ac.cn

SPECIALTY SECTION
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Ecology and Evolution

RECEIVED 07 January 2023
ACCEPTED 10 March 2023
PUBLISHED 27 March 2023

CITATION
Li X, Zheng Z, Xu C, Zhao P, Chen J, Wu J, Zhao X, Mu X, Zhao D and Zeng Y (2023) Individual tree-based forest species diversity estimation by classification and clustering methods using UAV data.
Front. Ecol. Evol. 11:1139458.
doi: 10.3389/fevo.2023.1139458

Individual tree-based forest species diversity estimation by classification and clustering methods using UAV data

Xiuwen Li[1,2], Zhaoju Zheng[1], Cong Xu[1,2], Ping Zhao[1,2], Junhua Chen[1,2], Jinchen Wu[1,2], Xueming Zhao[1,2], Xuan Mu[1,2], Dan Zhao[1,2] and Yuan Zeng[1,2]*

[1]State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China

Monitoring forest species diversity is essential for biodiversity conservation and ecological management. Currently, unmanned aerial vehicle (UAV) remote sensing technology has been increasingly used in biodiversity monitoring due to its flexibility and low cost. In this study, we compared two methods for estimating forest species diversity indices, namely the spectral angle mapper (SAM) classification approach based on the established species-spectral library, and the self-adaptive Fuzzy C-Means (FCM) clustering algorithm by selected biochemical and structural features. We conducted this study in two complex subtropical forest areas, Mazongling (MZL) and Gonggashan (GGS) National Nature Forest Reserves using UAV-borne hyperspectral and LiDAR data. The results showed that the classification method performed better with higher values of R^2 than the clustering algorithm for predicting both species richness (0.62>0.46 for MZL and 0.55>0.46 for GGS) and Shannon-Wiener index (0.64>0.58 for MZL, 0.52>0.47 for GGS). However, the Simpson index estimated by the classification method correlated less with the field measurements than the clustering algorithm (R^2=0.44 and 0.83 for MZL and R^2=0.44 and 0.62 for GGS). Our study demonstrated that the classification method could provide more accurate monitoring of forest diversity indices but requires spectral information of all dominant tree species at individual canopy scale. By comparison, the clustering method might introduce uncertainties due to the amounts of biochemical and structural inputs derived from the hyperspectral and LiDAR data, but it could acquire forest diversity patterns rapidly without distinguishing the specific tree species. Our findings underlined the advantages of UAV remote sensing for monitoring the species diversity in complex forest ecosystems and discussed the applicability of classification and clustering methods for estimating different individual tree-based species diversity indices.

KEYWORDS

forest species diversity, classification, clustering, UAV, individual tree-based

1. Introduction

Forest biodiversity is essential in maintaining ecosystem patterns, functions and services (Balvanera et al., 2006; Thompson et al., 2009; Brockerhoff et al., 2013). Forest species diversity is a fundamental component of biodiversity, which refers to the uniformity of the number and distribution of tree species in forest ecosystems (Magurran, 1988; Vellend, 2004). However, with

the increasing pressure of human activities and climate change, it has faced severe threats, such as accelerated species extinction and increased endangered species (Iida and Nakashizuka, 1995; Haas et al., 2011). Therefore, accurate and repeated forest species diversity monitoring is important for biodiversity conservation and ecological management.

Currently, the assessment of species diversity in a certain region is mainly based on species diversity indices, among which species richness (Gaston, 2000) emphasizes the number of various species, while Shannon-Wiener index (Shannon, 1948) and Simpson index (Simpson, 1949) take into account both the amount and evenness of species. Traditional forest species diversity monitoring relies on field surveys to investigate these diversity indices (Kerr and Ostrovsky, 2003), which are labor- and material-intensive and focus on forest species distribution at the sample scale (Myers et al., 2000; Duro et al., 2007). Remote sensing has the advantages of an extensive detection range and a short data acquisition period, extending the possibilities of forest species diversity monitoring at both temporal and spatial scales (Turner et al., 2003; Skidmore et al., 2015). Near-surface remote sensing platform equipped with hyperspectral sensors and laser scanners has been a promising tool for forest species diversity monitoring in the past decade (Turner, 2014; Guo et al., 2017; Wang and Gamon, 2019; Pu, 2021). In particular, UAV (unmanned aerial vehicle) remote sensing technology, due to its advantages of flexibility and low cost, has shown great potential in species identification and biodiversity monitoring (Anderson and Gaston, 2013; Lin et al., 2019; De Almeida et al., 2021).

Hyperspectral data can obtain continuous spectral information of vegetation and has been increasingly used for monitoring forest species diversity (Féret and Asner, 2014; Ferreira et al., 2016; Laurin et al., 2016). The monitoring methods are mainly divided into two categories: supervised classification methods that directly identify forest species based on their spectroscopic characteristics (Féret and Asner, 2013; Fassnacht et al., 2016; Cao et al., 2018; Franklin and Ahmed, 2018), and spectral diversity metrics that indirectly link the variation of leaf or canopy spectra to species diversity (Palmer et al., 2002; Gholizadeh et al., 2018). Light Detection and Ranging (LiDAR) data can directly penetrate the vegetation canopy by actively emitting high-frequency pulses, so it is widely used for high-precision estimation of forest structural features, including tree height and crown diameter (Popescu, 2007; Morsdorf et al., 2009; Sankey et al., 2013; Wallace et al., 2014). Furthermore, advances in lidar remote sensing have enabled the accurate extraction of information from individual tree crowns (ITCs) (Ene et al., 2012; Zhao et al., 2014). Compared to the pixel-based approach, the ITC-based approach is more directly analogous to the field-based individual sampling method, which can better extract structural features of the canopy and minimize the signal confusion brought by non-tree pixels (Zheng et al., 2022). Based on LiDAR data, forest species diversity at the regional scale can be monitored by establishing relationships between structural features and field-measured species diversity indices (Lopatin et al., 2016; Torresani et al., 2020; Hu et al., 2021). However, the capability of forest species diversity monitoring using only hyperspectral data or LiDAR data might be limited by species complexity, image spectral mixing and canopy morphological variation (Koch, 2010; Guo et al., 2017). The fusion of hyperspectral and LiDAR data provides a synergistic ability, which can use vertical and horizontal information from each data source to acquire more robust diversity monitoring results.

Previous studies have integrated structural features extracted by LiDAR data and spectral characteristics from hyperspectral images for directly discriminating tree species by using classification techniques, including linear discrimination analysis (Alonzo et al., 2014), support vector machine (Dalponte et al., 2012), random forest (Liu et al., 2017) and spectral angle mapper (Zhao et al., 2020). Mayra et al. (2021) compared the performance of different classification methods for identifying the major tree species in a boreal forest based on airborne hyperspectral and LiDAR data. Assessing forest species diversity using remote sensing classification methods has the advantage of providing spatially explicit species distribution information for each ITC or pixel. However, it remains challenging to directly discriminate the species of all individuals accurately in complex subtropical or tropical forests due to the potential spectral or structural similarity among different species or differences existing for the same species (Price, 1994; Wang and Gamon, 2019). The confusion in classification usually increases with increasing biodiversity levels and more training data for species-rich forests is usually needed to improve the classification accuracy. Moreover, collecting sufficient training and validation data for each tree species in species-rich and topographically complex forests can be a challenging task. Although some methods are relatively capable of classifying trees with limited training samples (Christian et al., 2013; Awad, 2018), the classification results are achieved using specific images and algorithms with relatively lower transferability.

Many indirect approaches using spectral and structural information have shown great potential for monitoring forest species diversity, such as regression analysis and clustering. Regression analysis is to model the spectral and structural information directly with the measured species diversity indices, which is a mature and straightforward algorithm, but the applicability in different regions is poor (Ceballos et al., 2015). The clustering algorithm can evaluate species diversity by grouping trees with similar characteristics based on the biochemical and structural variation of different tree species (Asner et al., 2015; Padilla-Martinez et al., 2020; Pakgohar et al., 2021). Clustering can be used to identify patterns or trends in the distribution and abundance of different species within a forest ecosystem. Among them, the self-adaptive Fuzzy C-Means (FCM) clustering algorithm overcomes the disadvantage of traditional clustering methods, which require a pre-indication of the initial classes, and can determine the optimal number of clusters automatically (Bezdek et al., 1987; Li and Yu, 2009). Zhao et al. (2018) estimated the forest species richness and Shannon-Wiener index in a subtropical forest based on airborne LiDAR and hyperspectral data using seven biochemical components and tree height by the adaptive FCM clustering algorithm. The biochemical and structural parameters selected in each study area for clustering methods may be dependent, and their applicability to species diversity monitoring in other areas still needs to be determined, especially for forests with diverse species and complex compositions.

Several studies have used either classification or clustering methods to estimate forest species diversity by combining various features from remote sensing data, but it remains unclear which method is more effective for monitoring different aspects of diversity in different forest conditions. Although some studies have used spaceborne or airborne data to assess species diversity (Shen and Cao, 2017; Wan et al., 2021), they are often limited by spatial resolution or

expensive costs. UAV-borne hyperspectral and LiDAR data could provide spatially explicit information on individual trees, so it is more advantageous to explore the applicability of advanced methods in different species-rich forests by UAV data.

Therefore, the major objectives of our study are to explore the performance of individual tree-based classification and clustering methods in estimating three commonly used forest species diversity indices (species richness, Shannon-Wiener index and Simpson index) in two typical subtropical forests in China using UAV-borne hyperspectral and LiDAR data. We aim to: (1) classify tree species using the SAM classification method based on hyperspectral image and the individual tree crown segmentation results from LiDAR data, (2) estimate forest species diversity using the self-adaptive FCM clustering algorithm based on optimal biochemical vegetation indices and structural features, and (3) further compare the performance of classification and clustering methods in these two subtropical forest sites.

2. Materials and methods

2.1. Study area

Subtropical forest in China is a hotspot of tree species richness and a priority area for forest species diversity monitoring (Li et al., 2009; Liu et al., 2018). We conducted research in two study areas, both of which are typical subtropical forests of China, but their forest species compositions and environmental conditions are different. The first study area is located in the Mazongling National Nature Reserve (MZL, 115°41′37′–115°42′5′E, 31°15′25′–31°15′44′N) in Jinzhai county, Anhui province of China (Figure 1). The study area covers about 23.8 ha with an elevation varying from 1,000 m to 1,184 m above sea level. This region is characterized by a subtropical monsoon climate. The average annual temperature is about 13 ~ 15°C, and the average annual precipitation is 1,510 mm (Fan et al., 2022). Mazongling National Nature Reserve has abundant forest resources, and the study area contains more than 10 dominant tree species, including *Quercus glandulifera*, *Platycarya strobilacea*, *Castanea mollissima* and *Lindera glauca*.

The second study area is situated in the Minya Konka National Park (also known as Gonggashan, GGS, 102°3′50′–102°4′28′E, 29°36′2′–29°36′15′N) in Ganzi (Garzê) Tibetan Autonomous Prefecture, Sichuan province of China with an elevation varying from 1959 m to 2,247 m above sea level (Figure 1). This study area is located in the transitional zone from the subtropical belt to the temperate belt of the eastern Tibetan Plateau, covering an area of approximately 20.5 ha. The average annual temperature is about 4.2°C, and the average annual precipitation is 1947 mm (Zhou et al., 2013). The forest canopy across this study area comprises more than 15 dominant tree species, including *Fagus longipetiolata*, *Jasminum nudiflorum*, *Ailanthus altissima*, *Cercidiphyllum japonicum* and *Bothrocaryum controversum*.

2.2. Data acquisition and preprocessing

2.2.1. UAV-borne hyperspectral and LiDAR data

The UAV-borne hyperspectral data were collected on September 18 and October 15, 2020, using the Cuber UHD185 Firefly imaging spectrometer (Cubert GmbH, Ulm, Baden-Württemberg, Germany) onboard a DJI Matrice 300 aircraft (Da Jiang innovate technology Ltd., Shenzhen, China) under cloudless conditions. The sensor comprises 125 visible and near-infrared spectral channels ranging from 450 nm to 946 nm with an 8 nm spectral resolution. The sensor was equipped on the UAV platform and flew at an altitude of 80 m, resulting in a 7 cm spatial resolution. The preprocessing of the images consisted of four preliminary steps. First, the Agisoft PhotoScan software (Agisoft LCC Co. St. Petersburg, Russia) was used for image mosaic. Spectral radiation calibration was the second step to convert the spectral response into the true spectral radiance. Then, the reflectance spectrum was calculated from the reference spectra of calibration plates and the spectral radiance. Finally, the geometric correction was performed in the Image Registration Workflow tool of ENVI5.3 software (Gai, 2019).

The UAV-borne LiDAR data were obtained simultaneously with the hyperspectral dataset acquisition using the LiAir VH Pro scanner (Green Valley Inc., Beijing, China) operating at a wavelength of 905 nm. The scanner provided a 70.4° horizontal (cross-track) and 77.2° vertical (along-track) angle of view. The height accuracy of the laser scanner was 5 cm and had an 80% flight strip overlap. The average point density in MZL is more than 117 points/m^2, and the average point density in MZL is more than 168 points/m^2. The point cloud data were noise filtered and classified into ground and vegetation returns using the software TerraSolid (Terrasolid, Helsinki, Finland). We generated the digital elevation model (DEM) based on the classified ground points and constructed the digital surface model (DSM) from the first pulse reflections of the LiDAR point clouds, and subtracted a canopy height model (CHM) with a resolution of 0.1 m (Zhao et al., 2013). The UAV-LiDAR data were normalized based on the ground points to remove the influence of terrain undulations on the height values. Besides, the vegetation point clouds with a normalized height below 2 m were removed to reduce the effect of background factors such as shrubs and grasses.

2.2.2. Field measurements

Field measurements were collected simultaneously with the UAV data acquisition in September–October 2020, and a supplemental survey was conducted in July 2022. A total of 26 square sample plots (30 × 30 m) within these two study areas were acquired. Differentially-corrected GPS determined the coordinates of the four corners of each sample plot. Tree parameters were measured in each sample plot, including tree species name, diameters at breast height (DBH), crown base height, tree height, crown classes (dominant, co-dominant, intermediate and suppressed trees) and crown diameters in two directions (south–north and east–west) for all individual trees with DBH ≥ 5 cm. The plot-level forest canopy closure and leaf area index (LAI) were also obtained by hemispherical photographs taken by a fish-eye camera along two diagonals. In addition, to validate individual tree segmentation and carry out classification research, we measured the location of each tree in two of the sample plots in MZL and four of the sample plots in GGS by integrating the Real Time Kinematic (RTK) GPS/GLONASS System with the total station.

We selected 10 dominant tree species in MZL and 15 dominant tree species in GGS and collected top-of-canopy leaves for these dominant tree species to measure their biochemical components and spectral properties. We measured 10 major biochemical components, including chlorophyll a and b (Chl-a, Chl-b), total carotenoids (Car),

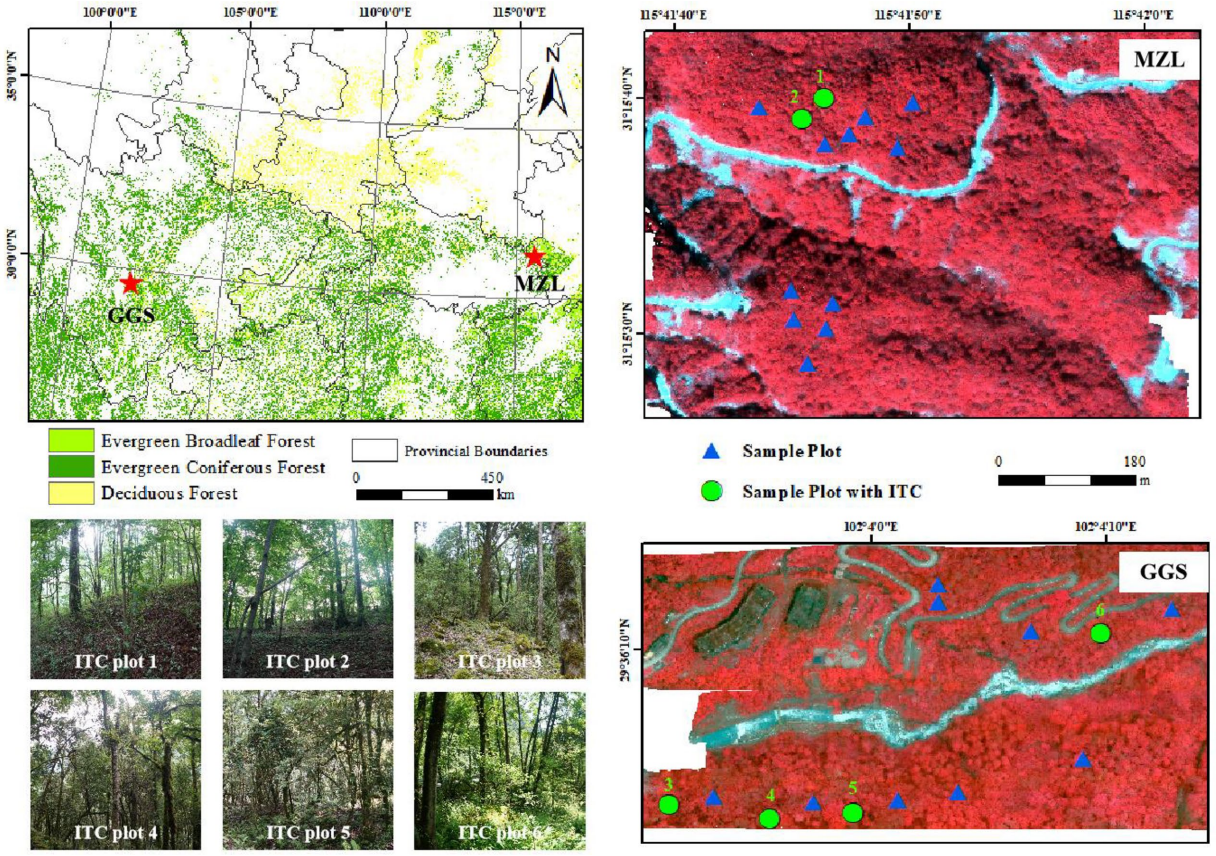

FIGURE 1
The location of two study areas (top left) with 10m spatial resolution from ChinaCover2020 (Wu et al., 2017), and MZL study area (top right) and GGS study area (bottom right) with imaging spectroscopy data acquired from Cuber UHD185 Firefly imaging spectrometer (Red: 866nm, Green: 654nm, Blue: 566nm). The blue triangles indicate the locations of field-measured sample plots. The green circles indicate the locations of individual tree crown (ITC) validation plots. Six photographs of these ITC plots are shown on the bottom left.

total carbon (C), nitrogen (N), phosphorus (P), cellulose (Cel), lignin (Lig), specific leaf area (SLA) and equivalent water thickness (EWT) similar as the previous study (Zheng et al., 2021). Leaves of each species with a mass of more than 150 g were selected and stored in plastic bags on ice and immediately transported to the laboratory for component analysis and spectroscopic measurement. Hemispherical reflectance spectra with 350–2,500 nm wavelengths were measured on 10 fresh leaves of each dominant tree species using a leaf clip coupled with the ASD FieldSpec 4 portable spectroradiometer (ASD Inc., Boulder, CO, United States). The bands with a wavelength of less than 400 nm and more than 2,400 nm were removed to eliminate the influence of instrument noise, and the spectra between 400 to 2,400 nm were smoothed by the Savitzky–Golay filter (Savitzky and Golay, 1964).

2.2.3. Species diversity indices

We used species richness, Shannon-Wiener index, and Simpson index to represent forest species diversity and calculated them within each sample plot based on the field measurements. Species richness refers to the total number of species in the sample plot. Shannon-Wiener index (Shannon, 1948) and Simpson index (Simpson, 1949) can reflect species richness and evenness of species distribution. They are comprehensive indicators reflecting the degree of species diversity. The Shannon-Wiener index is more sensitive to the number of species, and the Simpson index is more sensitive to the evenness of enriched species (Nagendra, 2002). The calculation formula of the Shannon-Wiener index (H) and Simpson index (D) were as follows:

$$H = \sum_{i=1}^{n} - p_i \ln p_i \quad (1)$$

$$D = 1 - \sum_{i=1}^{n} p_i^2 \quad (2)$$

where n is the total number of species in the sample plot, and p_i is the proportional abundance of the species i.

2.3. Methods

2.3.1. Individual tree crown segmentation

Based on the 0.1 m CHM data, we used a watershed algorithm combined with morphological crown control to separate the individual tree crowns (ITCs) (Wang et al., 2004; Chen et al., 2006; Zhao et al., 2014). Firstly, a morphological crown closing operator was used to determine the crown area and obtain the binary image of the

canopy. Next, a local extremum algorithm was used to detect the positions of the potential individual treetop. The actual individual treetop positions and potential crown shapes were calibrated through two watershed transformations and image reconstruction operations. Finally, the crown shapes were determined using an adaptive optimized morphological crown opening operator.

2.3.2. Spectral angle mapper (SAM) classification

Among many supervised classification methods, the SAM classification was used for its better performance in the hyperspectral data (Park et al., 2004; Yang et al., 2008; Zhang and Li, 2014). The SAM algorithm is a physically based spectral classification that uses an n-dimensional angle to match the extracted endmember spectra (Kruse et al., 1993; Park et al., 2007; Mohajane et al., 2017). The SAM algorithm determines the spectral similarity though calculating the angle between the spectrum vectors. Smaller angles correspond to closer matches to the endmember spectrum.

We calculated the average spectrum of each canopy based on the ITC segmentation results in the sample plots. Firstly, pixels with NDVI <0.2 and canopy height < 2 m were removed from the hyperspectral images to reduce the effect of background factors such as canopy gaps. A total of 2 ITC plots with 14 tree species in MZL (covering more than 90% of local tree species) and a total of 4 ITC plots with 22 tree species in GGS (covering more than 75% of local tree species) were used to establish the endmember spectral library. Then we determined the average spectrum of each species in the two study areas and used the SAM algorithm to classify them according to the established spectral library. With the classification results, we acquired the species diversity indices of each sample plot and used for validation.

2.3.3. Self-adaptive fuzzy C-means (FCM) clustering algorithm

The optimal biochemical components selection followed two principles (Zhao et al., 2016): (1) biochemical components can be well inverted by the spectrum. (2) these biochemical components are sufficient to distinguish different tree species. The partial least squares regression (PLSR) was used to determine the relationships between the *in-situ* leaf spectral and the biochemical measurements and explore whether the biochemical components of tree species can be quantitatively estimated by their spectral signals. The PLSR method combines the advantage of principal component analysis, canonical correlation analysis, and multiple linear regression analysis. It was performed using JMP14.0 statistical software.

After the optimal biochemical components were determined, the corresponding vegetation indices (VIs) from the hyperspectral data could be selected through the existing vegetation index models to estimate the biochemical components. Due to the lack of influential lignin invention bands, we finally identified nine canopy-scale VIs to indicate Chl (Chl-a and Chl-b), Car, C, N, P, Cel, SLA and EWT based on the literature (Table 1). Many studies have confirmed that the standard deviation of VIs in an area can reflect the species diversity in this region (Cayuela et al., 2006; Stickler and Southworth, 2008; Costanza et al., 2011), so we calculated the standard deviation of VIs for all ITCs at the plot scale, and performed Spearman correlation analysis with the species diversity indices (corrplot, R-package) to select the optimal VIs. The VI for each ITC was calculated by extracting the VI of the central pixel of each ITC. These canopy-level biochemical VIs were then converted into leaf-scale biochemical VIs by dividing the canopy-level biochemical VIs by the ITC's LAI to eliminate the effects caused by the canopy structure (Zarco-Tejada et al., 2001; Zhao et al., 2018). ITC's LAI was calculated by establishing the relationship between forest gap fraction (GF) and LAI according to Beer–Lambert Law (Richardson et al., 2009), as shown in Formula (3):

TABLE 1 Vegetation indices corresponding to the biochemical components.

Biochemical component	Vegetation index	Formula	Reference
Chl	TCARI/ OSAVI	TCARI / OSAVI = $3[(R_{750.6}-R_{704.6})-0.2(R_{750.6}-R_{550.6})(R_{750.66}/R_{704.6})] / (1+0.16)(R_{750.66}-R_{704.6})/(R_{750.66}+R_{704}+0.16)$	Daughtry et al. (2000) and Wu et al. (2008)
	VOG1	VOG1 = $R_{979.95}/R_{720.88}$	Vogelmann et al. (1993)
EWT	WBI	WBI = R_{895}/R_{972}	Penuelas et al. (1993)
Car	CRI	CRI = $1/R_{510}-1/R_{550}$	Gitelson et al. (2002)
Cel	PRI	PRI = $(R_{531}-R_{570})/(R_{531}+R_{570})$	Gamon et al. (1992)
N	CCCI	CCCI = $(0.7415R_{790}-0.6965R_{720})/(0.0319R_{790}-0.281R_{720})$	El-Shikha et al. (2007)
P	NDSI	NDSI = $(R_{553}-R_{518})/(R_{553}+R_{518})$	Patil et al. (2007)
SLA	RVI	RVI = R_{750}/R_{705}	Jordan (1969)
C	PSRI	PSRI = $(R_{680}-R_{500})R_{750}$	Merzlyak et al. (1999)

$$\begin{cases} GF = \dfrac{n_{ground}}{n_{ground}+n_{vegetation}} \\ LAI = -\cos(\theta) * \dfrac{\ln(GF)}{k} \end{cases} \quad (3)$$

where n_{ground} is the number of extracted ground points, $n_{vegetation}$ is the number of vegetation points, k is the extinction coefficient and takes a value of 0.5 if the vegetation is considered to follow the spherical leaf angle distribution, θ is zenith angle (LiDAR scanning angle) and GF is gap fraction.

We extracted 58 structural features for each ITC, including canopy cover, leaf area index, and height variables (statistical parameters related to point cloud height value) based on UAV-LiDAR data and the ITC-segmented ITC boundaries using LiDAR 360 software (Supplementary Table S1). Then we calculated the standard deviation of these structural variables in each sample plot. Finally, the Spearman correlation coefficient test with species

diversity indices was performed to obtain the optimal structural features.

Self-adaptive Fuzzy C-Means (FCM) clustering algorithm was applied to calculate the species richness (the number of clusters) based on the optimal biochemical VIs derived from the hyperspectral image and optimal structural features obtained from LiDAR data for each ITC. Each cluster was considered to be a specific but unidentified species. Then the Shannon-Wiener index and the Simpson index can be derived from the cluster amount and the ITC number of each cluster in the sample plot [Formula (1) and (2)]. The field-measured values of species diversity indices of 26 sample plots in two study areas were then compared with the forest biodiversity prediction results to verify the estimation accuracy of the clustering algorithms.

The standard Fuzzy C-Means algorithm transforms the cluster into a nonlinear optimization problem and achieves the number of categories through iteration (Bezdek et al., 1987). Self-adaptive Fuzzy C-Means (FCM) clustering algorithm was developed from the standard Fuzzy C-Means algorithm (Li and Yu, 2009). Self-adaptive FCM automatically determines the optimal number of clusters by using a new validity function without relying on the number of pre-set categories and prior knowledge, solving the acute problem of the clustering algorithm to the initial value. The validity function is defined as:

$$\begin{cases} L(c) = \dfrac{\sum_{i=1}^{c}\sum_{j=1}^{n}\dfrac{u_{ij}^{m}v_i-\bar{x}^2}{c-1}}{\sum_{i=1}^{c}\sum_{j=1}^{n}\dfrac{u_{ij}^{m}x_j-v_i^{-2}}{n-c}} \\ \bar{x} = \dfrac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^{m}x_j \end{cases} \quad (4)$$

where c is the number of clusters, m is the fuzzy weighting exponent, $X=\{x_1,x_2,\ldots,x_n\}$ is a sample data set, $V=\{v_1,v_2,\ldots,v_n\}$ is the cluster center dataset, u_{ij} represents the membership if the j-th sample point belongs to the i-th class, $\bar{x}$ is the central vector of all data, and $d_{ij}=x_j-v_i$ is the Euclidean distance between the j-th sample point and j-th cluster center.

3. Results

3.1. Individual tree crown segmentation

The ITC segmentation results of all 26 sample plots show that the amounts of segmented ITCs are quantitatively close to the ground-measured tree number (MZL: $R^2=0.76$, RMSE = 5.41; GGS: $R^2=0.82$, RMSE = 7.17; Figure 2). Due to the effects of crown overlap, point cloud density, small crowns, multi-stemmed trees and other reasons, some extra trees (over-segmentation) and missed trees (under-segmentation) can be found in the ITC segmentation results. The segmented and measured position of ITCs in sample plot 2 of MZL (71 segmented vs. 74 field-measured ITCs) and sample plot 4 of GGS (55 segmented vs. 54 field-measured ITCs) are shown in Figure 2 as an example. The over-segmentation phenomenon occurs in broad-leaved trees with large crowns and non-prominent treetops. In contrast, the under-segmentation phenomenon is caused by the overlapping crowns owing to the high forest canopy density. Multiple overlapping crowns are considered as one crown and are not isolated.

3.2. Forest species diversity prediction based on classification method

The SAM classification algorithm was applied to obtain the tree species of each ITC based on hyperspectral image and ITC boundaries from LiDAR data. As illustrated in Figure 3, a total of 14 endmembers in MZL and a total of 22 endmembers in GGS were extracted directly from the hyperspectral image. These tree species' endmembers were significantly different from each other and thus could be used for classification. Figure 4 shows the tree species classification results of two typical sample plots using the SAM algorithm.

The performance of the relationships between the predicted values and the three field-measured species diversity indices (species richness, Shannon-Wiener index, and Simpson index) is shown in Figure 5 (Blue colors). In MZL, the SAM classification algorithm demonstrated positive and significant predictive validity for species richness ($R^2=0.62$, RMSE = 1.44), Shannon-Wiener index ($R^2=0.64$, RMSE = 0.16) and Simpson index ($R^2=0.44$, RMSE = 0.05). In GGS, the estimated values and the measured species diversity indices were positively and significantly correlated only for species richness ($R^2=0.55$, RMSE = 2.87) and Shannon-Wiener index ($R^2=0.52$, RMSE = 0.24). The classification-based prediction of the Simpson index was positively correlated with the field measurements, but the correlation was not significant ($R^2=0.44$, $p=0.01$).

3.3. Forest species diversity validation of clustering algorithm

The estimation accuracies of biochemical components based on leaf spectra of tree species are shown in Table 2. It demonstrates that: (1) In MZL, Chl-a, Chl-b, EWT, Car, SLA and C could be strongly predicted by leaf spectra based on PLSR models ($R^2=0.78$–0.82). Cel, N, Lig and P are also relatively quantified by spectral reflectance ($R^2=0.44$–0.74). (2) In GGS, Chl-a, Chl-b, EWT and SLA could be well estimated by spectral signatures ($R^2=0.62$–0.73). Car, N and C also perform a relatively positive relationship with spectral properties ($R^2=0.30$–0.46). Cel, Lig and P have no obvious correlation with spectral reflectance ($R^2<0.30$).

Based on the optimal biochemical components that are spectrally obtainable (Table 2), we determined 9 biochemical VIs (TCARI/OSAVI, VOG1, CRI, WBI, CCCI, RVI, PRI, NDSI and PSRI) in MZL to indicate Chl (Chl-a and Chl-b), Car, EWT, N, SLA, Cer, P and C, respectively. As for GGS, we selected 7 biochemical VIs (TCARI/OSAVI, VOG1, WBI, RVI, CRI, CCCI and PSRI) to express Chl (Chl-a and Chl-b), EWT, SLA, Car, N and C, respectively. Supplementary Figures S1, S2 show the relationships between the standard deviation of ITC's biochemical VIs of sample plots and the species diversity indices in the two study areas. In MZL, there was a positive correlation between the standard deviation of 7 ITC-based VIs (WBI, TCARI/OSAVI, PRI, RVI, CCCI, VOG1 and PSRI) and species richness at the sample plot scale. These 7 VIs were selected as the optimal biochemical VIs. In

10.3389/fevo.2023.1139458

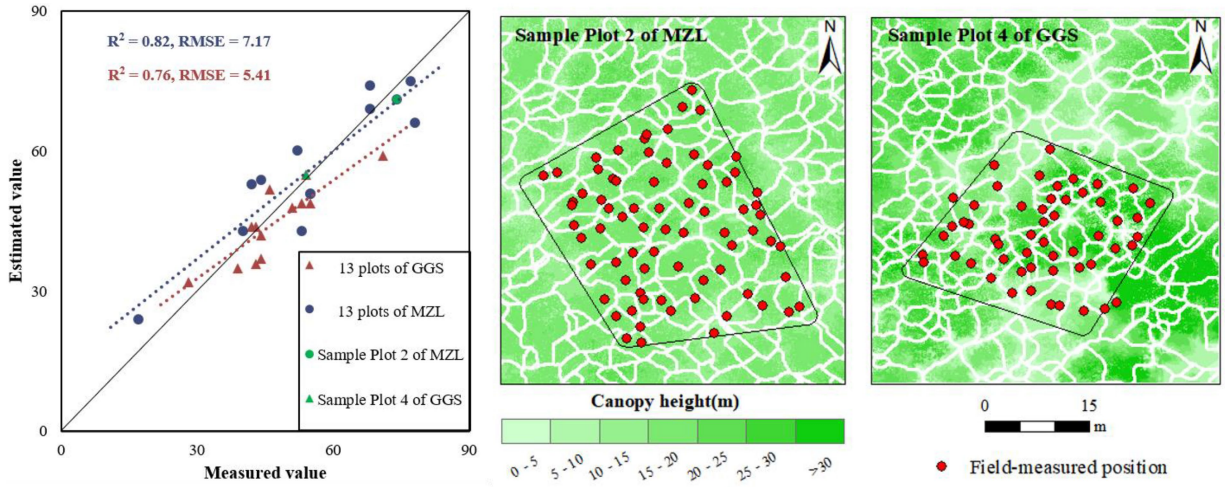

FIGURE 2
Scatter diagram for verification of ITCs (left) and results of individual tree separation in sample plot 2 of MZL (middle) and sample plot 4 of GGS (right) (black boxes represent the boundaries of plots, white polygons refer to the segmented tree crowns, and red points represent the field-measured positions at the base of tree stems).

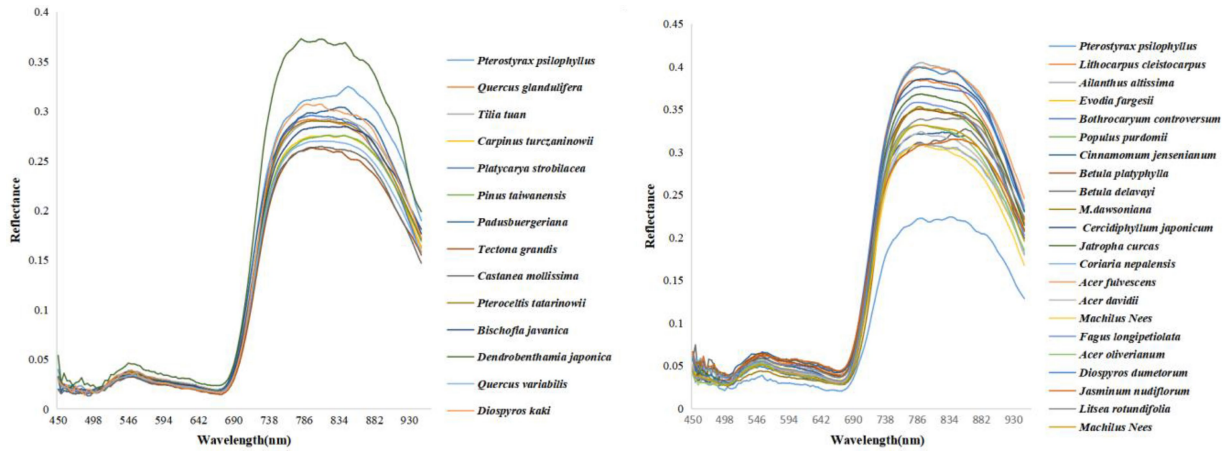

FIGURE 3
Endmember spectral library of dominant tree species in two study areas (left: MZL, right: GGS).

GGS, only 3 ITC's VIs (TCARI/OSAVI, RVI and PSRI) partially correlated with the species diversity indices. We regarded them as the optimal biochemical VIs.

The standard deviation of ITC-based structural features was weakly correlated with the species diversity indices in MZL. We finally conducted the two most relevant characteristics, namely canopy cover (CC) and density metric 30% (DM 30%) as the optimal structural features. In GGS, many structural features were positively and significantly correlated with species diversity indices. We finally determined five optimal structural features, including the interquartile range of accumulated elevation (Elev AIQ), coefficient of variance of elevation (Elev CV), the variance of elevation (Elev Var), density metric 20% (DM 20%) and density metric 30% (DM 30%), which showed high correlation with species diversity indices (Spearman correlation was above 0.5). Supplementary Figures S3, S4 show the relationships between the standard deviation of ITC's structural features of sample plots and the species diversity indices for MZL and GGS, respectively.

We applied the Self-adaptive FCM algorithm to estimate the three species diversity indices at 26 sample plots in two study areas based on the optimal biochemical VIs and optimal structural features for each ITC. The results are shown in Figure 5 (Red colors). In MZL, the clustering algorithm demonstrated positive and significant predictive validity for Shannon-Wiener index ($R^2 = 0.58$, RMSE = 0.22) and Simpson index ($R^2 = 0.83$, RMSE = 0.06). The estimated species richness was lower than the field-measured value (RMSE = 2.47) and performed relatively unsatisfactory inversion results ($R^2 = 0.46$, $p = 0.01$). In GGS, the estimated value and the measured species diversity indices were positively and significantly correlated only for Simpson index ($R^2 = 0.62$, RMSE = 0.07). The prediction results for

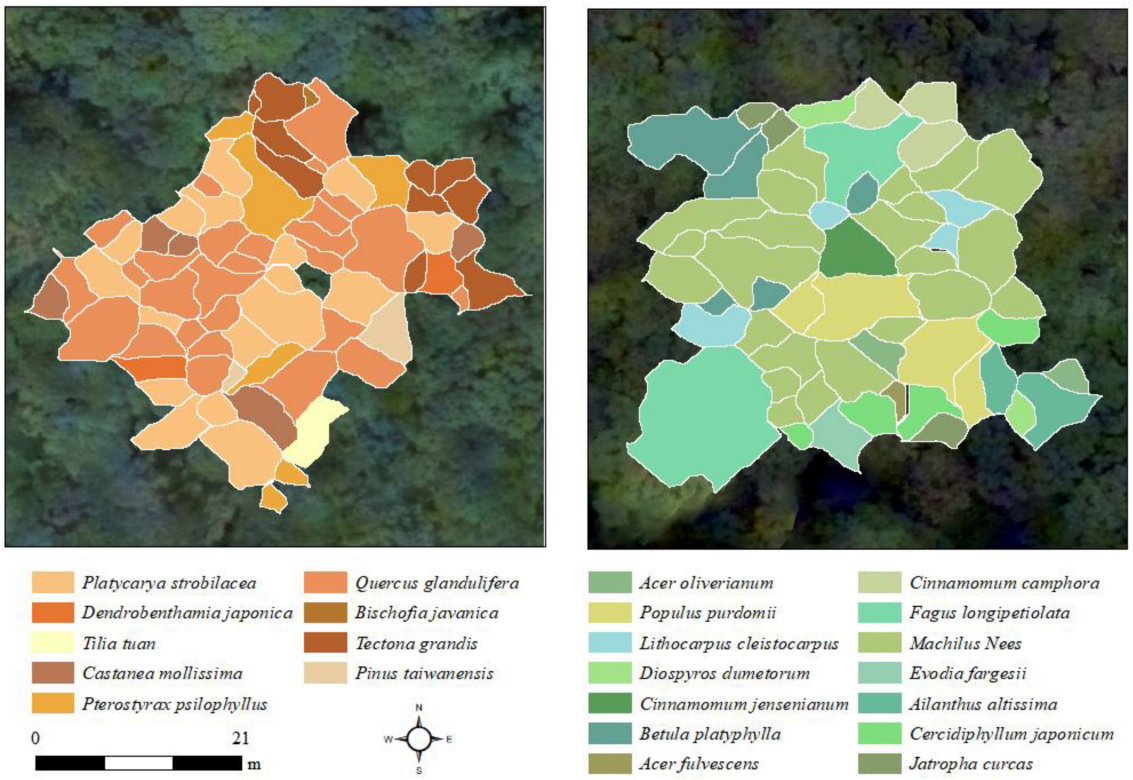

FIGURE 4
Tree species classification results of one sample plot in MZL (left) and one sample plot in GGS (right). Different geometric polygons represent identified ITCs, with white borders indicating the boundaries of the ITCs.

species richness ($R^2 = 0.46$, RMSE = 3.94) and Shannon-Wiener index ($R^2 = 0.47$, RMSE = 0.28) were positive but not significant enough ($p = 0.01$).

4. Discussion

Our results showed that the classification method performed better with higher values of R^2 than the clustering algorithm for predicting species richness (0.62 > 0.46 for MZL and 0.55 > 0.46 for GGS) and the Shannon-Wiener index (0.64 > 0.58 for MZL, 0.52 > 0.47 for GGS) in two study areas (Figure 5). However, the Simpson index estimated by the classification method correlated less with the field measurements than the clustering algorithm ($R^2 = 0.44$ and 0.83 for MZL and $R^2 = 0.44$ and 0.62 for GGS). This is probably due to the Simpson index weights rare species less and dominant species more than Shannon-Wiener index (Magurran, 1988; Daly et al., 2018), so the clustering algorithm taking dominant species/traits more into account are expected to predict the Simpson index accurately. Some previous studies have suggested that the Shannon-Wiener index is more closely related to species richness, while the Simpson index is more distantly correlated with richness (Nagendra, 2002; Costanza et al., 2011; Leinster and Cobbold, 2012). Our outcomes further indicated that the classification method is more advantageous in identifying rare species and estimating species richness, while the clustering method performs better in indicating the evenness of species. Constrained by the limited number of sample plots, it could be considered to use more independent validation plots to verify the advantages of classification and clustering methods in predicting species diversity indices in the future study.

We demonstrated that the individual tree-based SAM classification could be used to monitor the species diversity of complex forests and have the ability to distinguish the non-dominant species (Figure 4). This is mainly because SAM classification could distinguish similar spectra of tree species for classifying species based on hyperspectral data (e.g., *Platycarya strobilacea* and *Tilia tuan* in this experiment, Figure 3) when the endmember spectral library of dominant tree species is available (Awad, 2018; Zhao et al., 2020). However, when the spectra of non-dominant trees and dominant trees are very similar (such as *Carpinus turczaninowii* and *Castanea mollissima* in this study, Figure 3), SAM classification may also incorrectly classify them, which brings some challenges to the estimation of Shannon-Weiner and Simpson index. To better estimate species diversity using the SAM classification approach, it is necessary to extract their distinguishable bands to accurately classify these tree species. Moreover, forest structure has been identified as an essential indicator of forest species diversity (Ishii et al., 2004; Zeng et al., 2008; Guo et al., 2017; Torresani et al., 2020). We used the SAM classification to monitor species diversity based on the ITCs' spectral signal from UAV-hyperspectral data without considering the input of structural characteristics. The fusion of

10.3389/fevo.2023.1139458

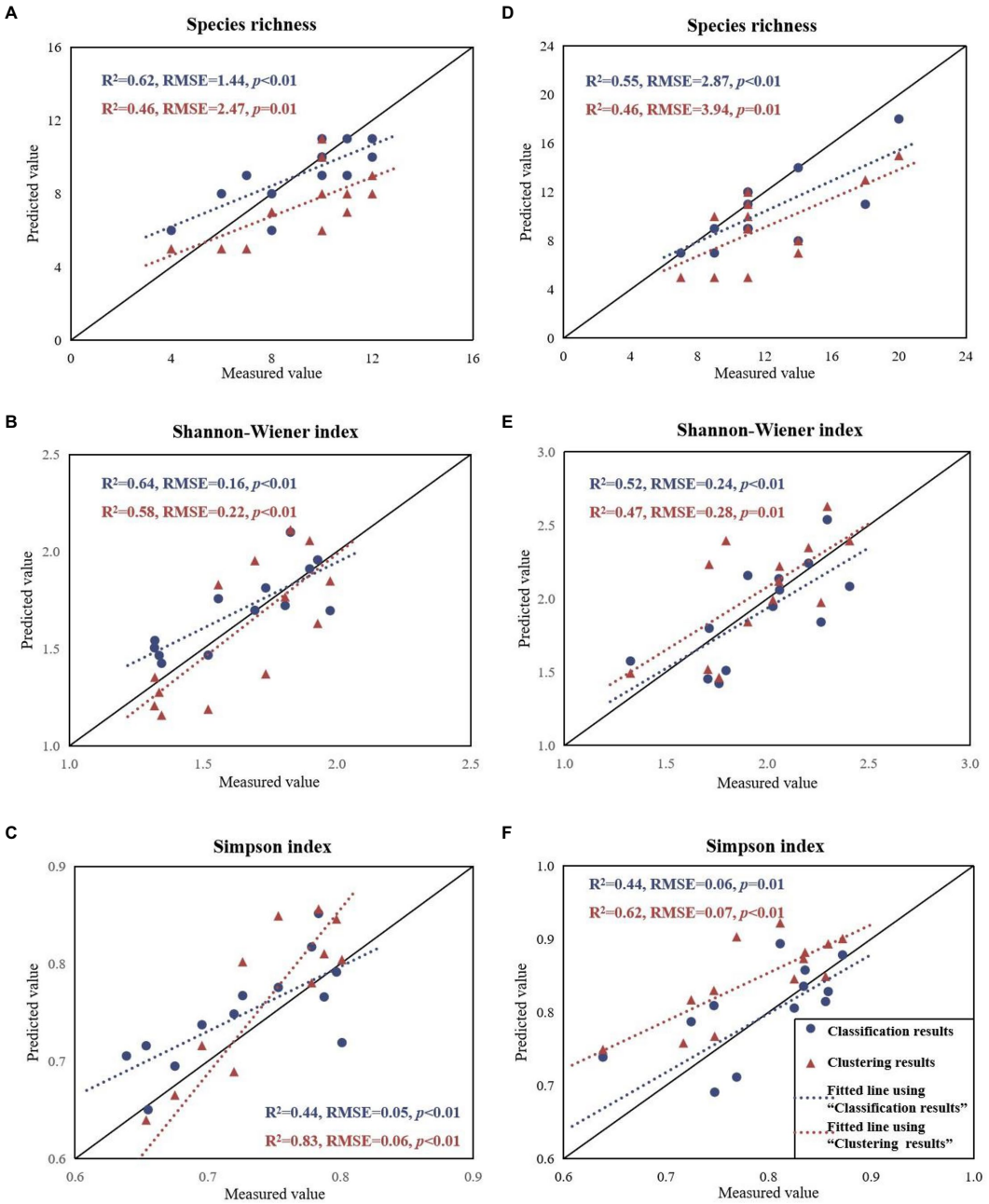

FIGURE 5
Field-measured species diversity indices compared with the predicted values based on classification and clustering approaches for MZL (left) and GGS (right).

TABLE 2 Estimation results of leaf biochemical components.

Biochemical component		Chl-a	Chl-b	EWT	Car	Cel	*N*	Lig	*P*	SLA	*C*
MZL	R^2	0.80	0.81	0.78	0.82	0.44	0.74	0.67	0.61	0.80	0.81
	RMSE	1.87	0.96	2.76	0.30	5.99	0.35	3.43	0.02	31.35	1.25
GGS	R^2	0.68	0.69	0.62	0.46	0.24	0.34	0.25	0.24	0.73	0.30
	RMSE	2.60	1.24	4.37	0.40	3.71	0.40	3.54	0.03	34.88	1.72

spectral and structural features could increase the dissimilarity among tree species and improve classification accuracy (Torabzadeh et al., 2019). Therefore, whether integrating LiDAR-derived tree structural parameters into the supervised classification of hyperspectral data can improve species diversity monitoring is worth to be further investigated.

Our results for both study areas demonstrated that forest diversity patterns could be rapidly acquired by the Self-adaptive FCM clustering algorithm based on individual tree-based variations in biochemical and structural features without distinguishing the tree species, which is similar to the previous clustering research (Féret and Asner, 2014; Schafer et al., 2016; Zhao et al., 2018). A maximum number of 11 and 15 tree species can be identified in the sample polt of MZL and GGS based on different optimal feature compositions using the clustering algorithm (7 optimal biochemical VIs and 2 optimal structural features for MZL, 3 optimal biochemical VIs and 6 optimal structural features for GGS, Figure 6). The outcomes of our feature selection further illustrated the spectral and structural heterogeneity of different regions and also emphasized the applicability of our clustering method in subtropical forests. Compared to the previous studies using Random Forest (RF) algorithm to select the optimal features, we underlined the strength of correlation analysis between the variation of biochemical VIs or structural features and species diversity indices at sample plot scale (Xie et al., 2019; Adhikari et al., 2020; de Almeida et al., 2021). The RF algorithm filters the optimal features according to the importance of the variables, while our feature selection method considers the basic biochemical and structural principles of forest (Hall, 2000; Strobl et al., 2008). However, the biochemical composition and structural characteristics of the same tree species vary considerably depending on individual development and landscape topography, introducing much uncertainty in selecting parameters for different forests.

Our results demonstrated better performance for forest species richness estimation in complex forests based on UAV-borne data (RMSE: 1.44 to 2.47 for MZL, 2.87 to 3.94 for GGS) than previous studies using airborne data (RMSE: 4.0 and 6.74) (Hernandez-Stefanoni et al., 2014; Zhao et al., 2018). Coarser image spatial resolution (typically between 1 to 10 m) and relatively lower point density (usually between 4 to 10 points/m^2) of airborne data can make it difficult to identify or segment trees with smaller canopies, and image spectral mixing may also be an issue (Medina et al., 2013; Sankey et al., 2017). This affects the accuracy of forest species diversity monitoring, as the spectral and structural differences between species may not be accurately captured (Ustin et al., 2004; Lesak et al., 2011; Naidoo et al., 2012). In contrast, UAV-borne LiDAR data with higher point cloud density (more than 100 points/m^2) could discriminate and detect individual trees with satisfactory accuracies (Figure 2). The spectral mixture problem would be solved with the ultra-high resolution UAV-borne imagery (Somers et al., 2011; Ronay et al., 2022), but how to better represent the spectral features of each ITC and avoid potential noise caused by intra-crown shade still need to be further studied (Rocchini et al., 2010). Given the lower flight altitude than conventional airborne platforms, the UAV-borne hyperspectral images are less affected by the atmosphere, leading to improved image quality and easier processing. UAV remote sensing has improved the timeliness of data acquisition, but it has limitations such as limited payload, short flight life, and more fabulous mosaic and geocode efforts (Nex and Remondino, 2014; Matese et al., 2015; Pu, 2021). In addition, due to the "top-down" operation method of UAV-borne platforms, the data for the understory in dense forest areas are often missing. Therefore, combining the advantages of different monitoring tools, such as ground-based LiDAR to complement and verify each other can provide more information for related forest diversity research.

5. Conclusion

In this study, we compared the performance of individual tree-based classification and clustering methods with UAV-borne data for estimating the forest species diversity indices in the Mazongling and Gonggashan National Nature Forest Reserves of China. We proved

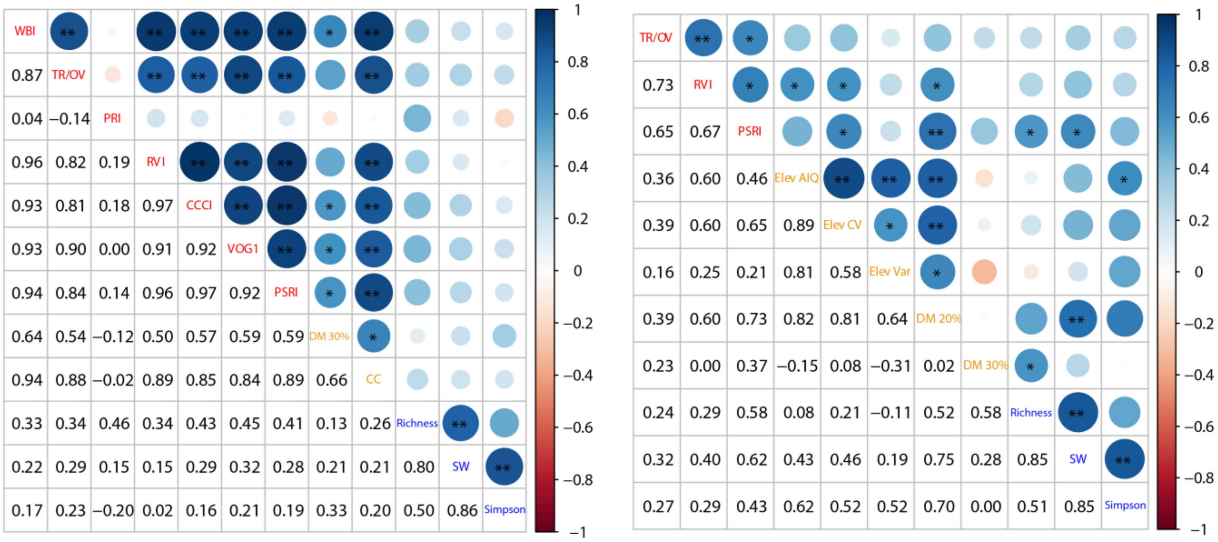

FIGURE 6
Selected features of two study areas (left: MZL, right: GGS; red labels: optimal biochemical VIs, orange labels: optimal structural features, blue labels: species diversity indices; TR/OV: TCARI/OSAVI, SW: Shannon-Wiener index).

that the SAM classification could provide more accurate predictions of species richness indices but requires spectral information of all dominant tree species. The Self-adaptive FCM clustering algorithm could achieve high-precision predictions for evenness indices (especially Simpson index), although information on specific tree species is unavailable.

The combination of UAV imaging spectroscopy and LiDAR make it possible to predict regional forest species diversity more accurately at individual canopy scale for complex forests. Future studies could improve the forest species-spectral library and explore forest species identification from multiple perspectives. Additionally, considering the variation in forest species characteristics over time, it would be valuable to further examine the accuracy of classification and clustering methods by incorporating phenological or multi-temporal features. Moreover, it would be beneficial to investigate the applicability of species diversity estimation models for forests in different ecological contexts and how high-resolution UAV data can be leveraged to bridge the scale gap between traditional field plot samplings and large-scale satellite observations.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

XL collected field data, calculated and analyzed the data, and wrote the manuscript. YZ and ZZ assisted in guiding the design of the experiment, discussing the results, and revising the manuscript. DZ and CX helped design the experiment and contributed to the manuscript. PZ, JC, JW, XZ, and XM supported the data investigation. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Key Research and Development Program of China (2020YFE0200800 and 2022YFF1302100) and the National Natural Science Foundation of China (No. 42071344).

Acknowledgments

We thank Ruiying Chang and Qiangxin Ou from the Gonggashan and Mazonglin National Nature Forest Reserves for their support with the fieldwork. We also thank Mingxing Zhang and Shaobo Yang for their assistance in the field sample collection.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2023.1139458/full#supplementary-material

References

Adhikari, H., Valbuena, R., Pellikka, P. K. E., and Heiskanen, J. (2020). Mapping forest structural heterogeneity of tropical montane forest remnants from airborne laser scanning and Landsat time series. *Ecol. Indic.* 108:105739. doi: 10.1016/j.ecolind.2019.105739

Alonzo, M., Bookhagen, B., and Roberts, D. A. (2014). Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sens. Environ.* 148, 70–83. doi: 10.1016/j.rse.2014.03.018

Anderson, K., and Gaston, K. J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Front. Ecol. Environ.* 11, 138–146. doi: 10.1890/120150

Asner, G. P., Martin, R. E., Anderson, C. B., and Knapp, D. E. (2015). Quantifying forest canopy traits: imaging spectroscopy versus field survey. *Remote Sens. Environ.* 158, 15–27. doi: 10.1016/j.rse.2014.11.011

Awad, M. M. (2018). Forest mapping: a comparison between hyperspectral and multispectral images and technologies. *J. For. Res.* 29, 1395–1405. doi: 10.1007/s11676-017-0528-y

Balvanera, P., Pfisterer, A. B., Buchmann, N., He, J. S., Nakashizuka, T., Raffaelli, D., et al. (2006). Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol. Lett.* 9, 1146–1156. doi: 10.1111/j.1461-0248.2006.00963.x

Bezdek, J. C., Hathaway, R. J., Sabin, M. J., and Tucker, W. T. (1987). Convergence theory for fuzzy c-means: counterexamples and repairs. *IEEE Trans. Syst. Man Cybern.* 17, 873–877. doi: 10.1109/tsmc.1987.6499296

Brockerhoff, E. G., Jactel, H., Parrotta, J. A., and Ferraz, S. F. B. (2013). Role of eucalypt and other planted forests in biodiversity conservation and the provision of biodiversity-related ecosystem services. *For. Ecol. Manag.* 301, 43–50. doi: 10.1016/j.foreco.2012.09.018

Cao, J. J., Leng, W. C., Liu, K., Liu, L., He, Z., and Zhu, Y. H. (2018). Object-based mangrove species classification using unmanned aerial vehicle hyperspectral images and digital surface models. *Remote Sens.* 10:89. doi: 10.3390/rs10010089

Cayuela, L., Benayas, J. M., Justel, A., and Salas-Rey, J. (2006). Modelling tree diversity in a highly fragmented tropical montane landscape. *Glob. Ecol. Biogeogr.* 15, 602–613. doi: 10.1111/j.1466-8238.2006.00255.x

Ceballos, A., Hernandez, J., Corvalan, P., and Galleguillos, M. (2015). Comparison of airborne lidar and satellite hyperspectral remote sensing to estimate vascular plant richness in deciduous mediterranean forests of Central Chile. *Remote Sens.* 7, 2692–2714. doi: 10.3390/rs70302692

Chen, Q., Baldocchi, D., Gong, P., and Kelly, M. (2006). Isolating individual trees in a savanna woodland using small footprint LIDAR data. *Photogramm. Eng. Remote Sens.* 72, 923–932. doi: 10.14358/PERS.72.8.923

Christian, B., Saini, M., Joshi, N., and Krishnayya, N. (2013). *Endmember Extraction and Classification of Tropical Trees (India) Using SFF and SAM Algorithm*. In: 2013 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS).

Costanza, J. K., Moody, A., and Peet, R. K. (2011). Multi-scale environmental heterogeneity as a predictor of plant species richness. *Landsc. Ecol.* 26, 851–864. doi: 10.1007/s10980-011-9613-3

Dalponte, M., Bruzzone, L., and Gianelle, D. (2012). Tree species classification in the southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sens. Environ.* 123, 258–270. doi: 10.1016/j.rse.2012.03.013

Daly, A. J., Baetens, J. M., and De Baets, B. (2018). Ecological diversity: measuring the unmeasurable. *Mathematics* 6:6070119. doi: 10.3390/math6070119

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., De Colstoun, E. B., and Mcmurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* 74, 229–239. doi: 10.1016/s0034-4257(00)00113-9

De Almeida, D. R. A., Broadbent, E. N., Ferreira, M. P., Meli, P., Zambrano, A. M. A., Gorgens, E. B., et al. (2021). Monitoring restored tropical forest diversity and structure through UAV-borne hyperspectral and lidar fusion. *Remote Sens. Environ.* 264:112582. doi: 10.1016/j.rse.2021.112582

Duro, D., Coops, N. C., Wulder, M. A., and Han, T. (2007). Development of a large area biodiversity monitoring system driven by remote sensing. *Prog. Phys. Geogr. Earth Environ.* 31, 235–260. doi: 10.1177/0309133307079054

El-Shikha, D. M., Waller, P., Hunsaker, D., Clarke, T., and Barnes, E. (2007). Ground-based remote sensing for assessing water and nitrogen status of broccoli. *Agric. Water Manag.* 92, 183–193. doi: 10.1016/j.agwat.2007.05.020

Ene, L., Naesset, E., and Gobakken, T. (2012). Single tree detection in heterogeneous boreal forests using airborne laser scanning and area-based stem number estimates. *Int. J. Remote Sens.* 33, 5171–5193. doi: 10.1080/01431161.2012.657363

Fan, W., Wang, J. J., Wang, H. L., Deng, P. F., Li, A. Q., Zhang, S. S., et al. (2022). Fine-root chemical traits rather than morphological traits of Chinese fir (*Cunninghamia lanceolata* (lamb.) hook.) plantations vary along an altitudinal gradient in eastern China. *Forest Syst.* 31:e010. doi: 10.5424/fs/2022312-18793

Fassnacht, F. E., Latifi, H., Sterenczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., et al. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* 186, 64–87. doi: 10.1016/j.rse.2016.08.013

Féret, J. B., and Asner, G. P. (2013). Tree species discrimination in tropical forests using airborne imaging spectroscopy. *IEEE Trans. Geosci. Remote Sens.* 51, 73–84. doi: 10.1109/tgrs.2012.2199323

Féret, J. B., and Asner, G. P. (2014). Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecol. Appl.* 24, 1289–1296. doi: 10.1890/13-1824.1

Ferreira, M. P., Zortea, M., Zanotta, D. C., Shimabukuro, Y. E., and De Souza, C. R. (2016). Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. *Remote Sens. Environ.* 179, 66–78. doi: 10.1016/j.rse.2016.03.021

Franklin, S. E., and Ahmed, O. S. (2018). Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *Int. J. Remote Sens.* 39, 5236–5245. doi: 10.1080/01431161.2017.1363442

Gai, J. (2019). Establishment of plot-yield prediction models in soybean breeding programs using UAV-based hyperspectral remote sensing. *Remote Sens.* 11:2752. doi: 10.3390/rs11232752

Gamon, J. A., Penuelas, J., and Field, C. B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* 41, 35–44. doi: 10.1016/0034-4257(92)90059-s

Gaston, K. J. (2000). Global patterns in biodiversity. *Nature* 405, 220–227. doi: 10.1038/35012228

Gholizadeh, H., Gamon, J. A., Zygielbaum, A. I., Wang, R., Schweiger, A. K., and Cavender-Bares, J. (2018). Remote sensing of biodiversity: soil correction and data dimension reduction methods improve assessment of alpha-diversity (species richness) in prairie ecosystems. *Remote Sens. Environ.* 206, 240–253. doi: 10.1016/j.rse.2017.12.014

Gitelson, A. A., Zur, Y., Chivkunova, O. B., and Merzlyak, M. N. (2002). Assessing carotenoid content in plant leaves with reflectance spectroscopy. *Photochem. Photobiol.* 75, 272–281. doi: 10.1562/0031-8655(2002)075<0272:Accipl>2.0.Co;2

Guo, X., Coops, N. C., Tompalski, P., Nielsen, S. E., Bater, C. W., and Stadt, J. J. (2017). Regional mapping of vegetation structure for biodiversity monitoring using airborne lidar data. *Eco. Inform.* 38, 50–61. doi: 10.1016/j.ecoinf.2017.01.005

Haas, S. E., Hooten, M. B., Rizzo, D. M., and Meentemeyer, R. K. (2011). Forest species diversity reduces disease risk in a generalist plant pathogen invasion. *Ecol. Lett.* 14, 1108–1116. doi: 10.1111/j.1461-0248.2011.01679.x

Hall, M. A.. (2000). *Feature Selection for Discrete and Numeric Class Machine Learning*. Stanford, CA, USA: Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000), Stanford University.

Hernandez-Stefanoni, J. L., Dupuy, J. M., Johnson, K. D., Birdsey, R., Tun-Dzul, F., Peduzzi, A., et al. (2014). Improving species diversity and biomass estimates of tropical dry forests using airborne LiDAR. *Remote Sens.* 6, 4741–4763. doi: 10.3390/rs6064741

Hu, T. Y., Sun, X. L., Su, Y. J., Guan, H. C., Sun, Q. H., Kelly, M., et al. (2021). Development and performance evaluation of a very low-cost UAV-Lidar system for forestry applications. *Remote Sens.* 13:77. doi: 10.3390/rs13010077

Iida, S., and Nakashizuka, T. (1995). Forest fragmentation and its effect on species diversity in sub-urban coppice forests in Japan. *For. Ecol. Manag.* 73, 197–210. doi: 10.1016/0378-1127(94)03484-e

Ishii, H. T., Tanabe, S., and Hiura, T. (2004). Exploring the relationships among canopy structure, stand productivity, and biodiversity of temperature forest ecosystems. *For. Sci.* 50, 342–355. doi: 10.1093/forestscience/50.3.342

Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the Forest floor. *Ecology* 50, 663–666. doi: 10.2307/1936256

Kerr, J. T., and Ostrovsky, M. (2003). From space to species: ecological applications for remote sensing. *Trends Ecol. Evol.* 18, 299–305. doi: 10.1016/s0169-5347(03)00071-5

Koch, B. (2010). Status and future of laser scanning, synthetic aperture radar and hyperspectral remote sensing data for forest biomass assessment. *ISPRS J. Photogramm. Remote Sens.* 65, 581–590. doi: 10.1016/j.isprsjprs.2010.09.001

Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., et al. (1993). The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.* 44, 145–163. doi: 10.1016/0034-4257(93)90013-n

Laurin, G. V., Puletti, N., Hawthorne, W., Liesenberg, V., Corona, P., Papale, D., et al. (2016). Discrimination of tropical forest types, dominant species, and mapping of functional guilds by hyperspectral and simulated multispectral Sentinel-2 data. *Remote Sens. Environ.* 176, 163–176. doi: 10.1016/j.rse.2016.01.017

Leinster, T., and Cobbold, C. A. (2012). Measuring diversity: the importance of species similarity. *Ecology* 93, 477–489. doi: 10.1890/10-2402.1

Lesak, A. A., Radeloff, V. C., Hawbaker, T. J., Pidgeon, A. M., Gobakken, T., and Contrucci, K. (2011). Modeling forest songbird species richness using LiDAR-derived measures of forest structure. *Remote Sens. Environ.* 115, 2823–2835. doi: 10.1016/j.rse.2011.01.025

Li, L., Huang, Z. L., Ye, W. H., Cao, H. L., Wei, S. G., Wang, Z. G., et al. (2009). Spatial distributions of tree species in a subtropical forest of China. *Oikos* 118, 495–502. doi: 10.1111/j.1600-0706.2009.16753.x

Li, Y., and Yu, F. S.. (2009). *A New Validity Function for Fuzzy Clustering*. In: International Conference on Computational Intelligence and Natural Computing, pp. 462–465.

Lin, Q. N., Huang, H. G., Wang, J. X., Huang, K., and Liu, Y. Y. (2019). Detection of pine shoot beetle (PSB) stress on pine forests at individual tree level using UAV-based hyperspectral imagery and Lidar. *Remote Sens.* 11:2540. doi: 10.3390/rs11212540

Liu, L. X., Coops, N. C., Aven, N. W., and Pang, Y. (2017). Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. *Remote Sens. Environ.* 200, 170–182. doi: 10.1016/j.rse.2017.08.010

Liu, X. J., Trogisch, S., He, J. S., Niklaus, P. A., Bruelheide, H., Tang, Z. Y., et al. (2018). Tree species richness increases ecosystem carbon storage in subtropical forests. *Proc. R. Soc. B Biol. Sci.* 285:1888. doi: 10.1098/rspb.2018.1240

Lopatin, J., Dolos, K., Hernandez, H. J., Galleguillos, M., and Fassnacht, F. E. (2016). Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in Central Chile. *Remote Sens. Environ.* 173, 200–210. doi: 10.1016/j.rse.2015.11.029

Magurran, A. E.. (1988). *Ecological Diversity and its Measurement*. Springer Netherlands.

Matese, A., Toscano, P., Di Gennaro, S. F., Genesio, L., Vaccari, F. P., Primicerio, J., et al. (2015). Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. *Remote Sens.* 7, 2971–2990. doi: 10.3390/rs70302971

Mayra, J., Keski-Saari, S., Kivinen, S., Tanhuanpaa, T., Hurskainen, P., Kullberg, P., et al. (2021). Tree species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks. *Remote Sens. Environ.* 256:112322. doi: 10.1016/j.rse.2021.112322

Medina, O., Manian, V., and Chinea, J. D. (2013). Biodiversity assessment using hierarchical agglomerative clustering and spectral Unmixing over hyperspectral images. *Sensors* 13, 13949–13959. doi: 10.3390/s131013949

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., and Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.* 106, 135–141. doi: 10.1034/j.1399-3054.1999.106119.x

Mohajane, M., Essahlaoui, A., Oudija, F., El Hafyani, M., and Teodoro, C. (2017). Mapping Forest species in the central middle atlas of Morocco (Azrou Forest) through remote sensing techniques. *ISPRS Int. J. Geo Inf.* 6:275. doi: 10.3390/ijgi6090275

Morsdorf, F., Nichol, C., Malthus, T., and Woodhouse, I. H. (2009). Assessing forest structural and physiological information content of multi-spectral LiDAR waveforms by radiative transfer modelling. *Remote Sens. Environ.* 113, 2152–2163. doi: 10.1016/j.rse.2009.05.019

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. B., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858. doi: 10.1038/35002501

Nagendra, H. (2002). Opposite trends in response for the Shannon and Simpson indices of landscape diversity. *Appl. Geogr.* 22, 175–186. doi: 10.1016/s0143-6228(02)00002-4

Naidoo, L., Cho, M. A., Mathieu, R., and Asner, G. (2012). Classification of savanna tree species, in the greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a random Forest data mining environment. *ISPRS J. Photogramm. Remote Sens.* 69, 167–179. doi: 10.1016/j.isprsjprs.2012.03.005

Nex, F., and Remondino, F. (2014). UAV for 3D mapping applications: a review. *Appl. Geomat.* 6, 1–15. doi: 10.1007/s12518-013-0120-x

Padilla-Martinez, J. R., Corral-Rivas, J. J., Briseno-Reyes, J., Paul, C., Lopez-Serrano, P. M., and Von Gadow, K. (2020). Patterns of density and production in the community forests of the Sierra Madre occidental, Mexico. *Forests* 11:307. doi: 10.3390/f11030307

Pakgohar, N., Rad, J. E., Gholami, G., Alijanpour, A., and Roberts, D. W. (2021). A comparative study of hard clustering algorithms for vegetation data. *J. Veg. Sci.* 32:e13042. doi: 10.1111/jvs.13042

Palmer, M. W., Earls, P. G., Hoagland, B. W., White, P. S., and Wohlgemuth, T. (2002). Quantitative tools for perfecting species lists. *Environmetrics* 13, 121–137. doi: 10.1002/env.516

Park, B., Windham, W. R., Lawrence, K. C., and Smith, D. P.. (2004). *Classification of Hyperspectral Imagery for Identifying Fecal and Ingesta Contaminants*. In: Monitoring Food Safety, Agriculture, and Plant Health).

Park, B., Windham, W. R., Lawrence, K. C., and Smith, D. P. (2007). Contaminant classification of poultry hyperspectral imagery using a spectral angle mapper algorithm. *Biosyst. Eng.* 96, 323–333. doi: 10.1016/j.biosystemseng.2006.11.012

Patil, V. D., Adsul, P. B., and Deshmukh, L. S. (2007). Studies on spectral reflectance under normal and nitrogen, phosphorus and pest and disease stress condition in soybean (*Glycine max* L.). *J. Indian Soc. Remote Sens.* 35, 351–359. doi: 10.1007/bf02990790

Penuelas, J., Filella, I., Biel, C., Serrano, L., and Save, R. (1993). The reflectance at the 950–970 nm region as an indicator of plant water status. *Int. J. Remote Sens.* 14, 1887–1905. doi: 10.1080/01431169308954010

Popescu, S. C. (2007). Estimating biomass of individual pine trees using airborne lidar. *Biomass Bioenergy* 31, 646–655. doi: 10.1016/j.biombioe.2007.06.022

Price, J. C. (1994). How unique are spectral signatures? *Remote Sens. Environ.* 49, 181–186. doi: 10.1016/0034-4257(94)90013-2

Pu, R. (2021). Mapping tree species using advanced remote sensing technologies: a state-of-the-art review and perspective. *J. Remote Sens.* 2021:9812624. doi: 10.34133/2021/9812624

Richardson, J. J., Moskal, L. M., and Kim, S. H. (2009). Modeling approaches to estimate effective leaf area index from aerial discrete-return LIDAR. *Agric. For. Meteorol.* 149, 1152–1160. doi: 10.1016/j.agrformet.2009.02.007

Rocchini, D., Balkenhol, N., Carter, G. A., Foody, G. M., Gillespie, T. W., He, K. S., et al. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: recent advances and open challenges. *Eco. Inform.* 5, 318–329. doi: 10.1016/j.ecoinf.2010.06.001

Ronay, I., Kizel, F., and Lati, R.. (2022). *The Effect of Spectral Mixtures on WEED Species Classification*. 24th ISPRS Congress on Imaging Today, Foreseeing Tomorrow, pp. 477–484.

Sankey, T., Donager, J., Mcvay, J., and Sankey, J. B. (2017). UAV lidar and hyperspectral fusion for forest monitoring in the southwestern USA. *Remote Sens. Environ.* 195, 30–43. doi: 10.1016/j.rse.2017.04.007

Sankey, T., Shrestha, R., Sankey, J. B., Hardegree, S., and Strand, E. (2013). Lidar-derived estimate and uncertainty of carbon sink in successional phases of woody encroachment. *J. Geophys. Res. Biogeosci.* 118, 1144–1155. doi: 10.1002/jgrg.20088

Savitzky, A., and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36, 1627–1639. doi: 10.1021/ac60214a047

Schafer, E., Heiskanen, J., Heikinheimo, V., and Pellikka, P. (2016). Mapping tree species diversity of a tropical montane forest by unsupervised clustering of airborne imaging spectroscopy data. *Ecol. Indic.* 64, 49–58. doi: 10.1016/j.ecolind.2015.12.026

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Shen, X., and Cao, L. (2017). Tree-species classification in subtropical forests using airborne hyperspectral and LiDAR data. *Remote Sens.* 9:1180. doi: 10.3390/rs9111180

Simpson, E. H. (1949). Measurement of diversity. *Nature* 163:688. doi: 10.1038/163688a0

Skidmore, A. K., Pettorelli, N., Coops, N. C., Geller, G. N., Hansen, M., Lucas, R., et al. (2015). Agree on biodiversity metrics to track from space. *Nature* 523, 403–405. doi: 10.1038/523403a

Somers, B., Asner, G. P., Tits, L., and Coppin, P. (2011). Endmember variability in spectral mixture analysis: a review. *Remote Sens. Environ.* 115, 1603–1616. doi: 10.1016/j.rse.2011.03.003

Stickler, C. M., and Southworth, J. (2008). Application of multi-scale spatial and spectral analysis for predicting primate occurrence and habitat associations in Kibale National Park, Uganda. *Remote Sens. Environ.* 112, 2170–2186. doi: 10.1016/j.rse.2007.10.013

Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T., and Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics* 9:307. doi: 10.1186/1471-2105-9-307

Thompson, I. D., Mackey, B. G., Mcnulty, S., and Mosseler, A.. (2009). *Forest Resilience, Biodiversity, and Climate Change: A Synthesis of the Biodiversity/Resilience/Stability Relationship in Forest Ecosystems*. Secretariat of the Convention on Biological Diversity, Montreal.

Torabzadeh, H., Leiterer, R., Hueni, A., Schaepman, M. E., and Morsdorf, F. (2019). Tree species classification in a temperate mixed forest using a combination of imaging spectroscopy and airborne laser scanning. *Agric. For. Meteorol.* 279:107744. doi: 10.1016/j.agrformet.2019.107744

Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Hauffe, H. C., Heym, M., et al. (2020). Height variation hypothesis: a new approach for estimating forest species diversity with CHM LiDAR data. *Ecol. Indic.* 117:106520. doi: 10.1016/j.ecolind.2020.106520

Turner, W. (2014). Sensing biodiversity. *Science* 346, 301–302. doi: 10.1126/science.1256014

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., and Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends Ecol. Evol.* 18, 306–314. doi: 10.1016/s0169-5347(03)00070-3

Ustin, S. L., Roberts, D. A., Gamon, J. A., Asner, G. P., and Green, R. O. (2004). Using imaging spectroscopy to study ecosystem processes and properties. *Bioscience* 54, 523–534. doi: 10.1641/0006-3568(2004)054[0523:Uistse]2.0.Co;2

Vellend, M. (2004). Parallel effects of land-use history on species diversity and genetic diversity of forest herbs. *Ecology* 85, 3043–3055. doi: 10.1890/04-0435

Vogelmann, J. E., Rock, B. N., and Moss, D. M. (1993). Red edge spectral measurements from sugar maple leaves. *Int. J. Remote Sens.* 14, 1563–1575. doi: 10.1080/01431169308953986

Wallace, L., Musk, R., and Lucieer, A. (2014). An assessment of the repeatability of automatic Forest inventory metrics derived from UAV-borne laser scanning data. *IEEE Trans. Geosci. Remote Sens.* 52, 7160–7169. doi: 10.1109/tgrs.2014.2308208

Wan, H. M., Tang, Y. W., Jing, L. H., Li, H., Qiu, F., and Wu, W. J. (2021). Tree species classification of Forest stands using multisource remote sensing data. *Remote Sens.* 13:144. doi: 10.3390/rs13010144

Wang, R., and Gamon, J. A. (2019). Remote sensing of terrestrial plant biodiversity. *Remote Sens. Environ.* 231:111218. doi: 10.1016/j.rse.2019.111218

Wang, L., Gong, P., and Biging, G. S. (2004). Individual tree-crown delineation and treetop detection in high-spatial-resolution aerial imagery. *Photogramm. Eng. Remote Sens.* 70, 351–357. doi: 10.14358/PERS.70.3.351

Wu, C. Y., Niu, Z., Tang, Q., and Huang, W. J. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: modeling and validation. *Agric. For. Meteorol.* 148, 1230–1241. doi: 10.1016/j.agrformet.2008.03.005

Wu, B., Qian, J., and Zeng, Y.. (2017). *Land Cover Atlas of the People's Republic of China (1:1,000,000)*. Beijing: Sinomaps Press.

Xie, Z. L., Chen, Y. L., Lu, D. S., Li, G. Y., and Chen, E. X. (2019). Classification of land cover, Forest, and tree species classes with ZiYuan-3 multispectral and stereo data. *Remote Sens.* 11:164. doi: 10.3390/rs11020164

Yang, C., Everitt, J. H., and Bradford, J. M. (2008). Yield estimation from hyperspectral imagery using spectral angle mapper (SAM). *Trans. ASABE* 51, 729–737. doi: 10.13031/2013.20649

Zarco-Tejada, P. J., Miller, J. R., Noland, T. L., Mohammed, G. H., and Sampson, P. H. (2001). Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* 39, 1491–1507. doi: 10.1109/36.934080

Zeng, Y., Schaepman, M. E., Wu, B., Clevers, J., and Bregt, A. K. (2008). Scaling-based forest structural change detection using an inverted geometric-optical model in the three gorges region of China. *Remote Sens. Environ.* 112, 4261–4271. doi: 10.1016/j.rse.2008.07.007

Zhang, X. Y., and Li, P. J. (2014). Litho logical mapping from hyperspectral data by improved use of spectral angle mapper. *Int. J. Appl. Earth Obs. Geoinf.* 31, 95–109. doi: 10.1016/j.jag.2014.03.007

Zhao, D., Pang, Y., Li, Z. Y., and Liu, L. J. (2014). Isolating individual trees in a closed coniferous forest using small footprint lidar data. *Int. J. Remote Sens.* 35, 7199–7218. doi: 10.1080/01431161.2014.967886

Zhao, D., Pang, Y., Li, Z. Y., and Sun, G. Q. (2013). Filling invalid values in a lidar-derived canopy height model with morphological crown control. *Int. J. Remote Sens.* 34, 4636–4654. doi: 10.1080/01431161.2013.779398

Zhao, D., Pang, Y., Liu, L. J., and Li, Z. Y. (2020). Individual tree classification using airborne LiDAR and hyperspectral data in a natural mixed Forest of Northeast China. *Forests* 11:303. doi: 10.3390/f11030303

Zhao, Y. J., Zeng, Y., Zhao, D., Wu, B. F., and Zhao, Q. J. (2016). The optimal leaf biochemical selection for mapping species diversity based on imaging spectroscopy. *Remote Sens.* 8:216. doi: 10.3390/rs8030216

Zhao, Y. J., Zeng, Y., Zheng, Z. J., Dong, W. X., Zhao, D., Wu, B. F., et al. (2018). Forest species diversity mapping using airborne LiDAR and hyperspectral data in a subtropical forest in China. *Remote Sens. Environ.* 213, 104–114. doi: 10.1016/j.rse.2018.05.014

Zheng, Z. J., Zeng, Y., Schneider, F. D., Zhao, Y. J., Zhao, D., Schmid, B., et al. (2021). Mapping functional diversity using individual tree-based morphological and physiological traits in a subtropical forest. *Remote Sens. Environ.* 252:112170. doi: 10.1016/j.rse.2020.112170

Zheng, Z. J., Zeng, Y., Schuman, M. C., Jiang, H. L., Schmid, B., Schaepman, M. E., et al. (2022). Individual tree-based vs pixel-based approaches to mapping forest functional traits and diversity by remote sensing. *Int. J. Appl. Earth Obs. Geoinform.* 114:103074. doi: 10.1016/j.jag.2022.103074

Zhou, J., Wu, Y. H., Jorg, P., Bing, H. J., Yu, D., Sun, S. Q., et al. (2013). Changes of soil phosphorus speciation along a 120-year soil chronosequence in the Hailuogou glacier retreat area (Gongga Mountain, SW China). *Geoderma* 195, 251–259. doi: 10.1016/j.geoderma2012.12.010