Journal of Oceanology and Limnology

https://doi.org/10.1007/s00343-025-5293-8

Research Paper

Improved XGBoost with multi-source UAV data for high-accuracy fine-scale mangrove mapping

Zhaohui CHENG[1], Yongze LI[1], Xiong SUN[1], Jiajun YUAN[1], Dazhao LIU[1, 3, *], Qinyuan XIANG[2, *]

[1] *College of Electronic and Information Engineering, Guangdong Ocean University, Zhanjiang 524088, China*

[2] *College of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang 524088, China*

[3] *Guangdong Engineering Technology Research Center for Ocean Remote Sensing and Information Technology, Zhanjiang 524088, China*

Received Aug. 18, 2025; accepted in principle Oct. 9, 2025; accepted for publication Nov. 22, 2025

Abstract Unmanned aerial vehicle (UAV) datasets can derive diverse features, providing crucial support for fine-scale mangrove species classification. However, achieving high classification accuracy remains challenging due to complex feature interactions. This study utilized multi-source UAV data, including multispectral imagery, light detection and ranging (LiDAR) point clouds, and high-resolution RGB images, from the Gaoqiao Mangrove Nature Reserve, Zhanjiang, Guangdong, South China. Three hybrid feature groups were made by integrating shared multispectral features, vegetation indices, and structural features with texture features derived from principal component analysis (PCA), independent component analysis (ICA), or minimum noise fraction (MNF) dimensionality reduction. An improved Extreme Gradient Boosting (XGBoost) algorithm was developed for dominant feature selection, and random forest (RF) and XGBoost models were built for performance evaluation. The optimal results were obtained using PCA features selected by the improved XGBoost algorithm combined with the XGBoost classifier, achieving an overall accuracy of 98.48% with the user accuracy variance of only 0.000 05 among species. These findings indicate that the modified XGBoost algorithm can enhance classification accuracy and robustness, offering technical support for precise mangrove monitoring, protection, and restoration.

Keyword: mangrove fine-scale classification; improved Extreme Gradient Boosting (XGBoost); multi-source unmanned aerial vehicle (UAV) data; feature selection

1 INTRODUCTION

Mangroves are biological communities growing in coastal intertidal zones (Murdiyarso et al., 2015). They not only purify seawater, protect coastlines, and maintain ecological balance but also have strong carbon sink potential (Alongi, 2014; Santos et al., 2019; Wang et al., 2021). Currently, global mangrove ecosystems are facing varying degrees of degradation due to human activities, extreme climates, pests and diseases, and even internal competition (Lu et al., 2025). To protect and restore mangroves, ecological monitoring is a prerequisite (Abraham et al., 2022).

With the development of geographic information technology, mangrove monitoring has gradually shifted from manual field surveys to remote sensing monitoring (Xin et al., 2014). Compared with manual field surveys, remote sensing monitoring saves time and labor costs while improving monitoring efficiency (Elmahdy et al., 2020; Hu et al., 2020; Ghorbanian et al., 2021; Zhao et al., 2023). Such advancements have enabled studies on mangrove distribution and dynamics across diverse regions, including work by Liu et al. (2018) who focused on the Zhanjiang region using satellite data, alongside broader applications documented in existing literature. However, due to the limitations of free satellite remote sensing data, such as low spatial and spectral

* Corresponding authors: lldz@163.com; 545974941@qq.com

resolution, long imaging cycles, and the high cost of paid satellite remote sensing data, satellite remote sensing technology is significantly restricted in tasks requiring accurate mangrove ecological monitoring (Wang et al., 2019; Peng et al., 2020; Xia et al., 2021; Chen et al., 2022; Zeng et al., 2023).

The application and promotion of unmanned aerial vehicle (UAV) remote sensing technology have largely resolved the difficulty of accessing in-depth information on mangrove species during field surveys (Yin and Wang, 2019; Han and Li, 2021; Tian et al., 2022; Jiang et al., 2024; Zhen et al., 2024).

To further improve the accuracy of mangrove remote sensing classification, some scholars have attempted to acquire more features through multi-source remote sensing data. With the fusion of extensive feature information, the accuracy of mangrove species classification has been significantly enhanced (Cao et al., 2018; Huang et al., 2018; Li et al., 2019; Lassalle et al., 2023).

However, multi-feature fusion from UAV data inherently involves complex feature interactions—specifically manifested in three aspects: (1) feature redundancy, e.g., vegetation indices (i.e., NDVI, NDRE) derived from multispectral data share 90%+ linear correlation (calculated via Pearson coefficient in our pre-experiment), leading to redundant information that increases model computational burden without improving discriminative ability (Cai et al., 2019); (2) cross-feature interference, texture features extracted from high-resolution RGB images (e.g., Haralick Entropy) are easily affected by light detection and ranging (LiDAR)-derived elevation noise in dense canopies; for instance, the overlapping crowns of *Bruguiera gymnorhiza* and *Rhizophora stylosa* cause LiDAR elevation fluctuations, which distort the texture uniformity of RGB images, resulting in misclassification between these two evergreen mangrove species; (3) complementarity imbalance: spectral features (e.g., red-edge bands) perform well in distinguishing evergreen mangroves (e.g., *Bruguiera gymnorhiza*), but structural features (e.g., LiDAR-derived canopy height) are more critical for low-growing species (e.g., *Avicennia marina*).

These complex feature interactions not only increase the computational complexity of classification models but also contribute to classification result distortion, a common issue in fine-scale mangrove classification. Specifically, this distortion refers to the phenomenon where high overall accuracy (OA) masks low user accuracy (UA) for individual species: for instance, in our preliminary experiments using traditional XGBoost, the OA of 95.63% was driven by high UA (97.82%) for the dominant species *Bruguiera gymnorhiza*, while the UA for the less abundant *Sonneratia apetala* was only 89.15%, with a user's accuracy variance (UA_v) of 0.000 23 (far higher than the 0.000 05 of our improved method). Such distortion misleads ecological assessments, as it fails to reflect the true distribution of rare or less dominant mangrove species (e.g., *Avicennia marina*) (Collin et al., 2018; Salum et al., 2021; Xu et al., 2021). Thus, addressing both complex feature interactions and classification result distortion remains a key challenge in fine-scale mangrove classification.

UAV remote sensing data can provide diverse vegetation indices, spectral bands, and textural features. However, excessive features not only increase the computational burden of classification models but may also reduce accuracy (Cai et al., 2019; Cao et al., 2021; Wang et al., 2022). To minimize computational waste and improve accuracy, researchers employ various methods for dominant feature selection (Richter et al., 2016; He et al., 2020; Pham et al., 2020; Guo et al., 2021). Among these, Xu et al. (2021) used the XGBoost algorithm for dominant feature selection in mangroves, achieving better results in improving classification accuracy than traditional methods. However, XGBoost was not originally designed for remote sensing applications (Chen and Guestrin, 2016), leaving room for improvement in its ability to assess feature importance for mangrove remote sensing data. To date, no research has attempted to optimize the feature importance attribute of XGBoost to enhance the precision of fine-scale mangrove classification.

In this study, six sets of dominant features from multi-source UAV data of the Gaoqiao Mangrove Reserve in Zhanjiang, Guangdong Province, were extracted using both the original and improved XGBoost algorithms, and then used to train corresponding XGBoost and random forest (RF) classification models. The classification results were used. The advantages of the improved feature selection algorithm were presented, and the impact of different texture feature extraction methods in classification accuracy were analyzed, and finally a comprehensive evaluation was provided. This research aims to provide a scientific basis and technical support for advancement in mangrove ecological conservation, countermeasure against global climate change, and improvement in the accuracy and timeliness of mangrove monitoring.

2 MATERIAL AND METHOD

2.1 Study area

The study area is located in Gaoqiao Town, Zhanjiang City, Guangdong Province, China (109°46′20″E–109°46′50″E; 21°31′50″N–21°32′20″N) (Fig.1). This region features southern subtropical monsoon maritime climate in mean annual temperature of 23 °C and annual precipitation ranging from 1 700 to 1 800 mm. Mangrove coverage spans 26.87 ha. The local waters have an annual mean temperature of 23.5 °C and are subject to mixed tides dominated by irregular diurnal cycles, with a mean tidal range of 2.53 m (maximum of 6.25 m) (Li et al., 2023).

The study area experienced low anthropogenic disturbance despite of its proximity to villages, and is predominantly natural forest, supporting rich mangrove biodiversity, including *Bruguiera gymnorhiza*, *Rhizophora stylosa*, *Avicennia marina*, *Sonneratia apetala*, *Aegiceras corniculatum*, and *Kandelia obovata*, with no distinct boundaries between communities (Yu et al., 2021). Among these species, only four dominant ones form significant stands, whereas the others occur as scattered individuals without forming distinct communities. Specifically, *Bruguiera gymnorhiza* is the most abundant with the widest distribution. *Rhizophora stylosa* and *Avicennia marina* rank second and third in abundance, primarily colonizing intertidal mudflats, while *Sonneratia apetala*, which is slightly less abundant than the preceding three, forms communities concentrated along the subtidal fringe in offshore zones (Li et al., 2023).

2.2 Field data acquisition and initial processing

During a November 2023 field campaign, our team surveyed 78 monospecific quadrats (4 m×4 m) for four dominant mangrove species: *Bruguiera gymnorhiza*, *Rhizophora stylosa*, *Avicennia marina*, and *Sonneratia apetala*. As illustrated in Fig.2, the integrated methodology included precise UAV-based

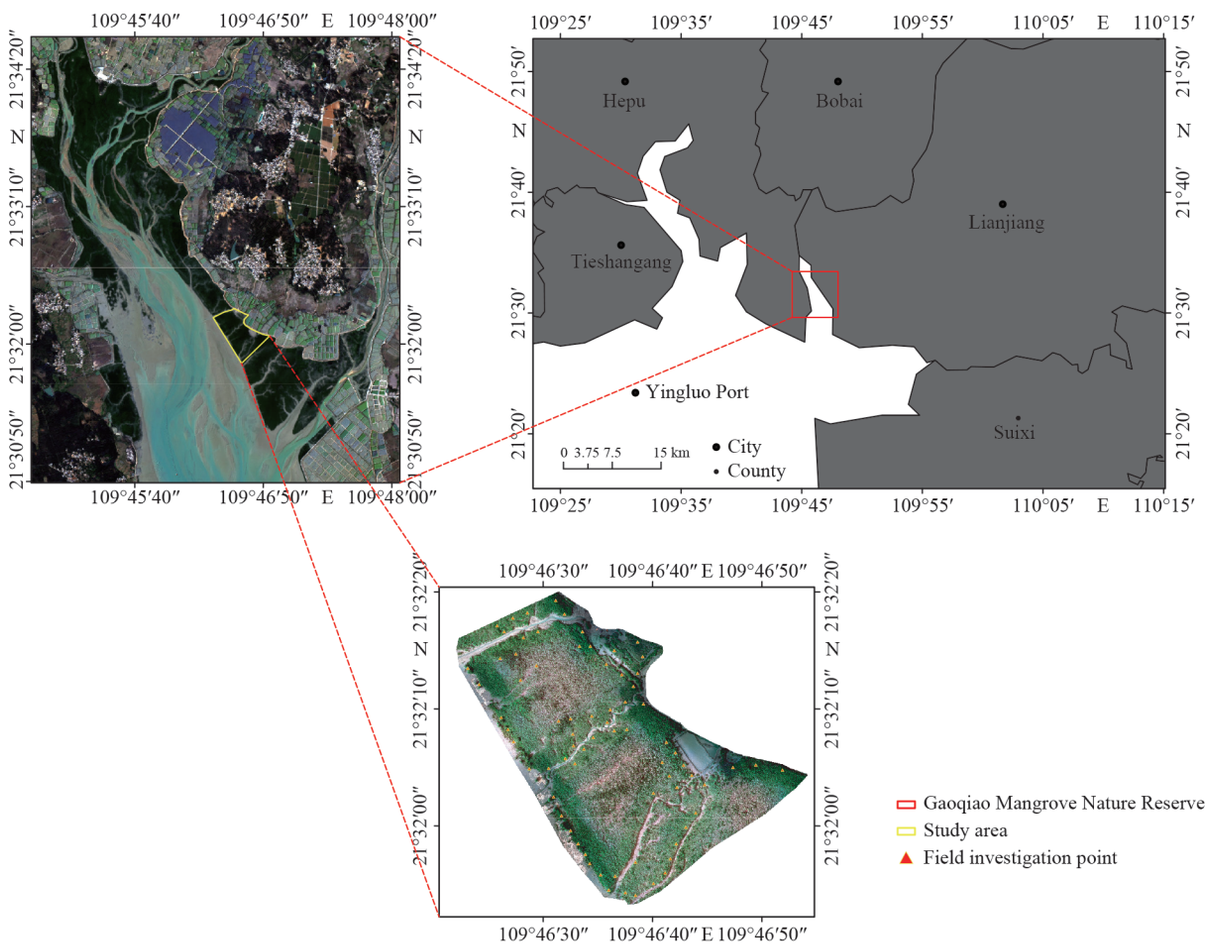


Fig.1 Location of the study area

Fig.2 Field investigation records of mangroves

a. recording geographical information; b. smartphone photography; c. low-altitude UAV flight; d. *Bruguiera gymnorhiza*; e. *Avicennia marina*; f. *Rhizophora stylosa*; g. *Sonneratia apetala*.

RTK positioning for quadrat geotagging, concurrent morphological documentation via smartphone photography, and high-resolution orthophoto acquisition through coordinated low-altitude UAV flights.

During the field campaign, UAV-based multispectral and LiDAR data were acquired simultaneously at low tide. The specifications of the UAVs are shown in Table 1. Data mosaicking was performed using DJI Terra, Pix4Dmapper, and PIE Engine, and preprocessing was conducted with ENVI and ArcGIS Pro. Additionally, the extent of mangroves was extracted in ENVI by applying the normalized difference vegetation index (NDVI) combined with visual interpretation.

2.3 Field data acquisition and initial processing

The representativeness of samples is a crucial factor influencing the performance of classification models. Based on quadrat data recorded during field surveys, pure pixel regions of each dominant species were uniformly delineated across the study area using ENVI software to establish an overall sample dataset. From this dataset, 3 500 pixels per class were allocated to the feature selection subset, while 12 500 pixels per class were allocated to the classification subset; both subsets were then partitioned into training, validation, and test sets with a fixed ratio of 7∶2∶1. Relevant information about the target

Table 1 The specifications of the UAVs

Data type	Multispectral	LiDAR	RGB
Sensor	Changguang Yuchen MS600	DJI Zenmuse L1	FC6310R
UAV platform	DJI Matrice 300 RTK	DJI Matrice 300 RTK	DJI Phantom 4 RTK
Flight altitude	100 m	110 m	60 m
Spectral wavelength/point density	450–840 nm	240 000–480 000 points/s	450–650 nm
Number of bands	6	1	3
Spatial resolution	8 cm	≈200 points/m^2	5 cm
Longitudinal overlap (%)	80	80	80
Side overlap (%)	60	60	60
Photo			

species for sample allocation is summarized in Table 2.

2.4 Feature extraction and selection

2.4.1 Feature extraction

After radiometric calibration of UAV multispectral data, original spectral features were obtained. Leveraging the dual red-edge band characteristics, multiple vegetation indices such as NDVI, red edge NDVI (NDRE), and modified red edge ratio index (mSR720) were calculated using ENVI's band math tool. Texture features were extracted using the Haralick texture extraction tool within Orfeo toolbox, based on three distinct dimensionality-reduced datasets: the first principal component analysis (PCA), the maximum-information independent component analysis (ICA), and the maximum-information minimum noise fraction data (MNF). All texture features used a 3×3 window size. Elevation and intensity features were extracted from the LiDAR data using ArcGIS Pro's LAS dataset to raster tool. A total of 115 features were ultimately derived from these multi-source UAV datasets.

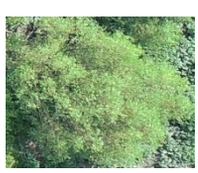
To fully exploit multi-source features, with reference to the experimental results of Fu et al. (2024b) and Xu et al. (2021), three hybrid feature groups were constructed by combining PCA-, ICA-, and MNF-based textures with a common feature pool containing spectral features, vegetation indices, and LiDAR features, forming F1 (PCA-textured), F2 (ICA-textured), and F3 (MNF-textured) groups.

2.4.2 Improved XGBoost-based feature selection method

XGBoost, a novel gradient-boosted decision tree algorithm proposed by Chen and Guestrin (2016), has shown significant potential in mangrove remote sensing feature selection, as empirically validated by Xu et al. (2021) and Pham et al. (2020). In feature selection, XGBoost uses three evaluation metrics: weight (the total number of times a feature is used for splitting across all trees), gain (the average gain produced by the feature in all trees where it is used), and cover (the average coverage of the feature within all trees). By default, weight is prioritized because it is independent of the model's parameter size and reflects the feature's role in the number of trees. Consequently, to conserve computational resources, XGBoost defaults to this metric for feature selection.

However, relying solely on weight for mangrove remote sensing feature selection has inherent theoretical limitations, rooted in two core principles of machine learning feature evaluation. The first principle is the frequency vs. contribution mismatch: weight quantifies only the split frequency of a

Table 2 Mangrove species classification system description

Classification category	Scientific name	Habitat characteristic	Photo
Dominant mangrove species 1	*Bruguiera gymnorhiza*	Widest distribution; dominates middle-high intertidal mudflats with stable sedimentation	
Dominant mangrove species 2	*Rhizophora stylosa*	Distributed in middle intertidal zones; adapted to moderate tidal erosion	
Dominant mangrove species 3	*Avicennia marina*	Colonizes low intertidal mudflats; tolerant to high salinity	
Dominant mangrove species 4	*Sonneratia apetala*	Concentrated in subtidal fringes of offshore zones; adapted to frequent tidal inundation	

feature across trees, but frequency does not equate to discriminative contribution. In high-dimensional remote sensing data, features with high split frequency may still be redundant. This redundancy arises because correlated features often share similar split patterns, leading to overcounting of their "usefulness" without actual improvement in classification performance.

The second principle is the ignorance of sample representativeness: feature importance in heterogeneous ecosystems requires not only reflecting how often a feature is used but also how well it represents the entire sample space. Weight fails to capture this representativeness, as it treats all splits equally regardless of the number of samples affected. This oversight violates the "coverage principle" in feature selection, which demands that important features should be relevant to a sufficient proportion of samples to avoid bias toward dominant data subsets.

These limitations are amplified in multi-source UAV data, where spectral, structural, and texture features exhibit distinct statistical properties. Spectral features may have moderate split frequency but high discriminative power, while structural features may have lower frequency but strong representativeness for specific vegetation strata. Relying solely on weight thus leads to suboptimal feature selection, as it cannot balance these complementary properties.

XGBoost, a novel gradient-boosted decision tree algorithm proposed by Chen and Guestrin (2016), has shown significant potential in mangrove remote sensing feature selection, as empirically validated by Pham et al. (2020) and Xu et al. (2021). In feature selection, XGBoost uses three evaluation metrics: weight (the total number of times a feature is used for splitting across all trees), gain (the average gain produced by the feature in all trees where it is used), and cover (the average coverage of the feature within all trees). By default, weight is prioritized because it is independent of the model's parameter size and reflects the feature's role in the number of trees. Consequently, to conserve computational resources, XGBoost defaults to this metric for feature selection.

To address these limitations, this study's improvement adopts a multi-criteria feature importance framework, which aligns with the theoretical requirements of high-dimensional remote sensing feature selection. This framework integrates weight, gain, and cover as complementary criteria, and this design is justified by three theoretical pillars.

The first theoretical pillar is the information complementarity principle. This principle states that gain measures the average reduction in model loss when a feature is used for splitting, directly quantifying the feature's discriminative information content. Cover, by contrast, measures the average number of samples affected by splits using the feature, quantifying the feature's representative information content. In remote sensing classification, these two types of information are complementary to weight—which reflects split frequency. Together, the three metrics capture both the frequency, discriminative power, and representativeness of features, addressing the limitations of single-criterion evaluation.

The second theoretical pillar is consistency with XGBoost's ensemble logic. XGBoost's performance relies on the collective contribution of multiple decision trees, where each tree's split is guided by three factors: loss reduction (gain), sample coverage (cover), and split frequency (weight). Integrating these three metrics thus reflects the algorithm's inherent ensemble logic, as each metric captures a distinct aspect of feature contribution: local loss reduction through gain, global sample impact through cover, and overall usage frequency through weight.

The third theoretical pillar is adaptation to multi-source data properties. Multi-source UAV data (multispectral, LiDAR, RGB) have distinct information densities: spectral data are rich in discriminative information, which is captured by gain; structural data (LiDAR) are rich in representative information, captured by cover; and texture features are rich in split frequency information, captured by weight. Integrating all three metrics ensures that the feature selection process adapts to these heterogeneous information sources, avoiding the "information imbalance" caused by single-criterion evaluation.

To verify this theoretical framework, a comparative experiment was conducted using the F1 feature group (PCA-texture+spectral+vegetation index+ LiDAR features, constructed in Section 2.4.1). The experiment included four scenarios corresponding to different combinations of the three criteria, designed to test their complementarity. The first scenario adopts the traditional XGBoost method, which uses weight as the sole evaluation criterion and is consistent with the algorithm's default setting. The second scenario integrates only weight and gain,

lacking representative information. The third scenario integrates only weight and cover, lacking discriminative information. The final scenario is this study's proposed method, which integrates weight, gain, and cover as complementary criteria.

To enhance the performance of the XGBoost algorithm in mangrove feature selection tasks, this study modified the algorithm's feature importance evaluation metrics. This modification allows it to consider a feature's cumulative split frequency across all trees, its average information gain in trees where it is utilized, and its mean coverage within host trees during feature screening. The newly proposed expression for the feature importance evaluation parameter is shown in Eq.1.

$$I_i = 0.5 \times G_i \times \frac{W_i}{\sum_{j=1}^{n} W_j} + 0.5 \times C_i \times \frac{W_i}{\sum_{j=1}^{n} W_j}. \quad (1)$$

In the formulation, I_i denotes the importance evaluation parameter of feature i. G_i represents the average information gain of feature i across trees where it is used for splitting. W_i indicates the cumulative split frequency of feature i. $\sum_{j=1}^{n} W_j$ signifies the total split count of all features. $W_i / \sum_{j=1}^{n} W_j$ corresponds to the proportion of splits using feature i relative to the total number of feature splits. C_i represents the average coverage of feature i within host trees.

The equal weight assigned to G_i and C_i, each accounting for 0.5 relative to the other, is not arbitrary but theoretically justified by the information entropy balance principle and remote sensing classification objectives. The first theoretical basis is the information entropy balance principle. In information theory, the "usefulness" of a feature can be quantified by its information entropy. Gain (G_i) corresponds to discriminative entropy, which represents the reduction in class uncertainty, while cover (C_i) corresponds to representative entropy, representing the reduction in sample space uncertainty. Assigning equal weight to G_i and C_i ensures that these two entropy components contribute equally to the overall feature importance (I_i), avoiding "entropy dominance". For example, excessive discriminative entropy may lead to overfitting, while excessive representative entropy may lead to underdiscrimination. For multi-source UAV data, this balance is critical because spectral features (which have high discriminative entropy) and structural features (which have high representative entropy) need to be equally valued to avoid biasing the feature selection toward one data type. The second theoretical basis is alignment with mangrove classification objectives. The primary goal of fine-scale mangrove classification stated in Section 1 of this study is to simultaneously achieve high OA and low UA_v. This goal requires balancing two key aspects: class-level discriminability (ensured by G_i, which enhances OA) and species-level balance (ensured by C_i, which reduces UA_v). A lopsided weight assignment, such as prioritizing G_i over C_i, would prioritize discriminability over balance, increasing OA but raising UA_v by neglecting rare mangrove species. Conversely, prioritizing C_i over G_i would prioritize balance over discriminability, lowering OA. The equal weight assignment of 0.5 to both G_i and C_i is the only ratio that theoretically optimizes this dual objective, as it does not favor either criterion and thus maintains the tradeoff between OA and UA_v. To validate this theoretical reasoning, a sensitivity analysis was conducted under the multi-criteria framework (consistent with the modified XGBoost method in this study). Using the F1 feature group (PCA-texture+spectral+vegetation index+LiDAR features, constructed in Section 2.4.1 of this study), the analysis tested five different ratios between G_i and C_i: 0–1, 0.25–0.75, 0.5–0.5, 0.75–0.25, and 1–0. As shown in Table 3, the ratio where G_i and C_i are equal (0.5–0.5) achieves the highest OA (98.14%) and the lowest UA_v (0.000 05), confirming that equal weight aligns with the theoretical expectation of balancing discriminability and representativeness.

2.4.3 Feature selection

In training classification models, an excessive number of features can lead to data redundancy. This not only hinders loss function convergence but also degrades model performance, thereby reducing classification accuracy. Consequently, when dealing with many features, it is necessary to algorithmically quantify the importance of each feature. Subsequently, features with lower importance are pruned based on predefined removal criteria, achieving discriminative feature selection (He et al., 2020; Guo et al., 2021).

To clarify the optimal threshold in the aforementioned "predefined removal criteria", we conducted multiple pruning intensity comparison experiments on the F1 feature group (PCA-texture+spectral+vegetation index+LiDAR); this group integrates

Table 3 Performance comparison of different feature importance evaluation scenarios

Experimental group	OA (%)	AD (%)	QD (%)	UA_v
Traditional XGBoost (weight only, algorithm default)	96.77	2.54	0.69	0.000 09
Proposed method (this study: weight integrated with gain)	96.79	2.53	0.68	0.000 08
Proposed method (this study: weight integrated with cover)	96.82	2.45	0.73	0.000 09
Proposed method (this study: weight integrated with gain and cover at a ratio of 0.25–0.75)	96.96	2.31	0.73	0.000 08
Proposed method (this study: weight integrated with gain and cover at a ratio of 0.5–0.5, core method)	98.14	1.46	0.40	0.000 05
Proposed method (this study: weight integrated with gain and cover at a ratio of 0.75–0.25)	96.82	2.51	0.67	0.000 08

OA: overall accuracy; AD: allocation disagreement; QD: quantity disagreement; UA_v: user's accuracy variance.

multi-source data and is the core object for subsequent validation. The experimental samples and classifier parameters were consistent with subsequent controlled experiments to avoid variable interference.

The evaluation metrics focused on three types of core measured data: OA for measuring the reliability of feature selection; UA_v as the key indicator of classification distortion (lower values indicate stronger stability of species discrimination), which is prioritized over allocation disagreement (AD) and quantity disagreement (QD); the number of retained features for balancing "information sufficiency" and "redundancy degree", preventing insufficient discriminative ability from too few features or increased model computational burden from too many features.

Based on the above metrics, five experimental groups with different pruning intensities were designed: two groups of over-loose pruning (60% NDVI similarity+80% inter-feature correlation, 60% NDVI similarity+90% inter-feature correlation), two groups of over-strict pruning (80% NDVI similarity+ 90% inter-feature correlation, 80% NDVI similarity+ 95% inter-feature correlation), and one balanced group (70% NDVI similarity+90% inter-feature correlation). By comparing the classification performance of each group, "70%+90%" was finally determined as the optimal threshold combination for the predefined removal criteria, providing a basis for the scientificity of subsequent feature screening.

To evaluate the performance of the proposed methodology, controlled experiments were conducted for feature screening tasks using both the original XGBoost algorithm and our approach. Computing feature importance with both methods requires training an XGBoost model. For optimal feature selection efficacy, fifteen data subsets were randomly sampled from the complete dataset; adaptive grid-search cross-validation was then employed for parameter tuning. The three hybrid feature groups were subsequently fed into the trained XGBoost model, generating six sets of feature importance data, as shown in Fig.3. Finally, features were pruned if exhibiting either >70% similarity to NDVI or pruning low-importance features if they showed >90% correlation to other features, with similarity/ correlation quantified via the Pearson correlation coefficient, with the specific formula given in Eq.2.

$$r=\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2}}. \quad (2)$$

In this equation, r denotes the Pearson correlation coefficient of features X and Y, n represents the number of training samples, X_i and Y_i denote the feature values of features X and Y for sample i, respectively, $\bar{X}$ and $\bar{Y}$ represent the mean values of features X and Y across n samples, respectively, $\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)$ indicates the covariance between features X and Y, $\sqrt{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}$ and $\sqrt{\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2}$ denote the variability of features X and Y across all training samples, respectively.

2.5 Classification method and evaluation metrics

2.5.1 Classification algorithm

XGBoost, a powerful ensemble machine learning method, performs strongly in remote sensing classification (Richter et al., 2016; Pham et al., 2020). Although RF, a representative bootstrap aggregation approach based on decision trees, underperforms XGBoost, it significantly outperforms traditional machine learning algorithms such as support vector machines (SVM), minimum distance classifiers, and Naive Bayes (Cao et al., 2021; Wang et al., 2022). Building on the feature screening results described above, this study trained both XGBoost and RF classifiers using the six pruned

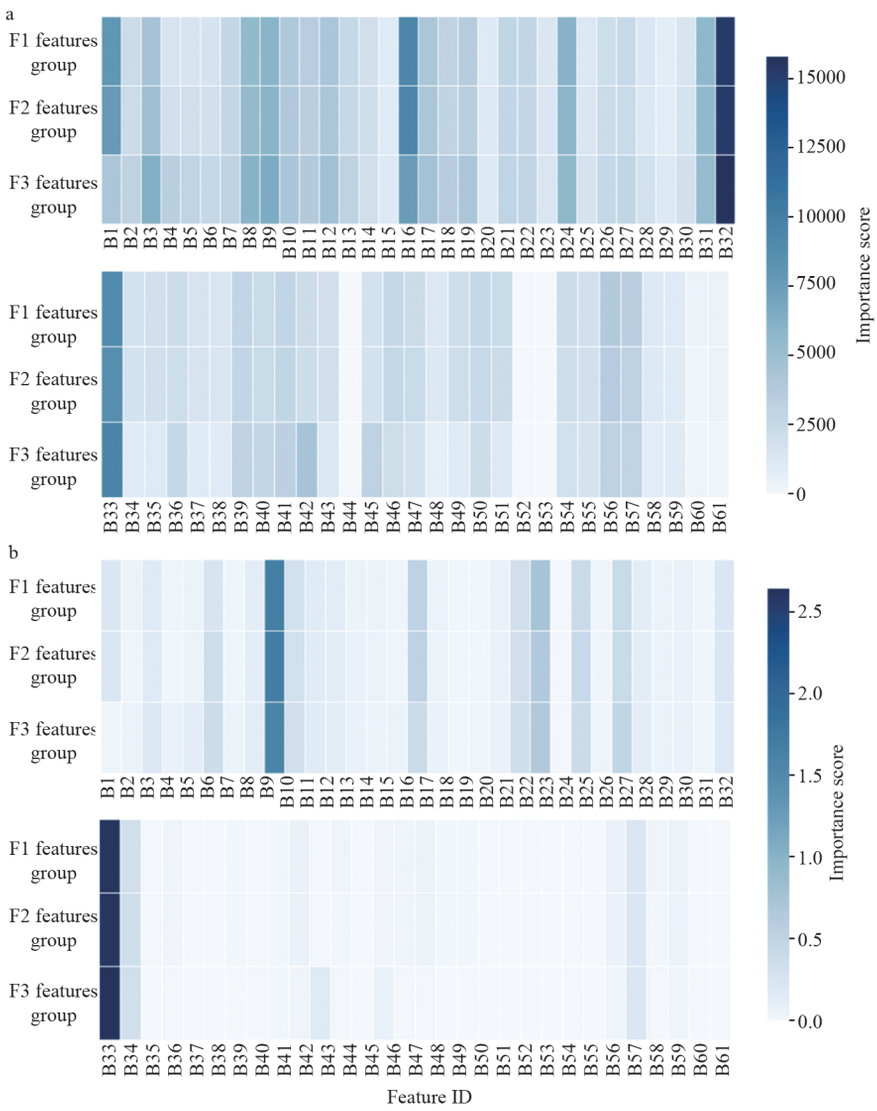


Fig.3 Heat map of feature importance

a. feature importance from Extreme Gradient Boosting; b. feature importance from the proposed method in this study.

feature groups generated during that process, validating the performance and generalizability of the proposed method.

XGBoost classifiers were implemented using Python third-party libraries, with the core XGBoost framework accessed via its official documentation and repository (https://xgboost.readthedocs.io/en/stable/). For RF model training, the Google Earth Engine (GEE) platform was utilized through its official code editor interface (https://code.earthengine.google.com/). To ensure theoretically optimal performance across all models, adaptive grid search with cross-validation was employed during the training phase for parameter tuning.

To further verify the generalizability of the improved feature selection method, this study additionally introduced two ensemble learning classifiers for comparative experiments, referring to the research ideas of Fu et al. (2024a) The same adaptive grid-search cross-validation strategy as used for XGBoost and RF was adopted for parameter tuning. The two newly added classifiers are specifically: 1) locally deployed CatBoost classifier (implemented in Python environment); 2) cloud-deployed Gradient Boosting Decision Tree (GBDT, implemented via GEE platform). Both classifiers used the F1 feature group (PCA-texture+spectral+vegetation index+LiDAR features) constructed in Section 2.4.1 and a 7∶2∶1 sample division ratio (training set∶validation set∶test set) to ensure consistency with the original experimental design and comparability of results.

2.5.2 Accuracy assessment metrics

Remote sensing classification employs standard accuracy metrics: OA, producer's accuracy (PA), UA, and Kappa coefficient. However, because the Kappa coefficient is highly correlated with OA, their combined use yields redundant information. Therefore, this study replaced the Kappa coefficient with AD and QD (Lyons et al., 2018), and additionally introduced UA_v as a quantitative indicator of classification result distortion.

Definition of classification result distortion: a phenomenon where the model's high OA is driven by extremely high accuracy of dominant species, while the accuracy of non-dominant or rare species is significantly suppressed, leading to an overestimation of the model's actual performance in distinguishing all target species. This distortion is quantified by UA_v—lower UA_v values indicate more balanced UA distribution across species and weaker classification distortion. The calculation formula of UA_v is shown in Eq.3.

$$UA_v = \frac{1}{k}\sum_{i=1}^{k}\left(UA_i - \overline{UA}\right)^2, \quad (3)$$

where k denotes number of target mangrove species (4 in this study: *Bruguiera gymnorhiza*, *Rhizophora stylosa*, *Avicennia marina*, *Sonneratia apetala*. UA_i represents the user's accuracy of species i. $\overline{UA}$ indicates the average user's accuracy of all species.

Higher AD/QD values indicate poorer classifier performance. When either AD or QD exceeds 10%, the resulting model instability necessitates critical reassessment of its deployment viability (Warrens, 2015). In this study, the UA_v was used to compare the uniformity of species-specific accuracy across different feature groups and feature selection methods—for instance, the optimal combination (F1 feature group+proposed method) achieved an extremely low UA_v, while other feature groups also showed significant reductions in UA_v compared to the original XGBoost, confirming the method's effectiveness in suppressing distortion.

With this, the entire experimental workflow—spanning data acquisition, feature extraction and selection, classification implementation, and accuracy evaluation—has been fully established, and the detailed flowchart is presented in Fig.4.

2.5.3 UAV data integration workflow in GEE

To enable the application of UAV-derived features in GEE for RF classifier training. All UAV-derived features required for classification were first integrated into a single dataset using the "Build Layer Stack" tool in ENVI. The merged dataset of all UAV-derived features was subsequently exported as a GeoTIFF file in ENVI. The exported GeoTIFF file was directly uploaded to the "Image Assets" module in GEE via the Code Editor's Asset Manager, completing the integration of all UAV-derived features into the GEE environment and laying the foundation for subsequent RF classifier training. For the complete GEE script used in this study, including calling the merged UAV features, training the RF classifier, and outputting classification results, readers can access the shared link: https://code.earthengine.google.com/efa6165590730403bfebfd96c97ebd36.

3 RESULT

3.1 Mangrove classification result and mapping

Figure 5 presents mangrove fine-classification results using features selected from three hybrid feature groups by two feature selection methods: XGBoost and the proposed method. These features were classified with XGBoost and RF classifiers, yielding OA between 92.30% and 98.48%. The optimal classification performance was achieved when features selected from the F1 group via the proposed method were input into the XGBoost classifier, with OA=98.48%, AD=1.23%, and QD=0.29%. Conversely, the poorest performance was observed when features selected from the F3 group via the original XGBoost algorithm were classified using the RF classifier, resulting in OA=92.30%, AD=6.52%, and QD=1.18%. All experimental combinations maintained overall disagreement (OD) below 10%, demonstrating empirically reliable classification outcomes.

The F1 group's ability to achieve the optimal OA of 98.48% is directly related to the "70% NDVI similarity+90% inter-feature correlation" threshold adopted during feature screening. To verify the rationality of this threshold, this study designed five comparison experiments with different pruning intensities for the F1 group (PCA-texture+spectral+vegetation index+LiDAR), with experimental samples and classifier parameters consistent with the controlled experiments in Section 2.4.3 to avoid variable interference. The experimental results are shown in Table 4, which details the OA, UA_v, number of retained features, and supplementary AD/QD indicators of the F1 group under different threshold combinations.

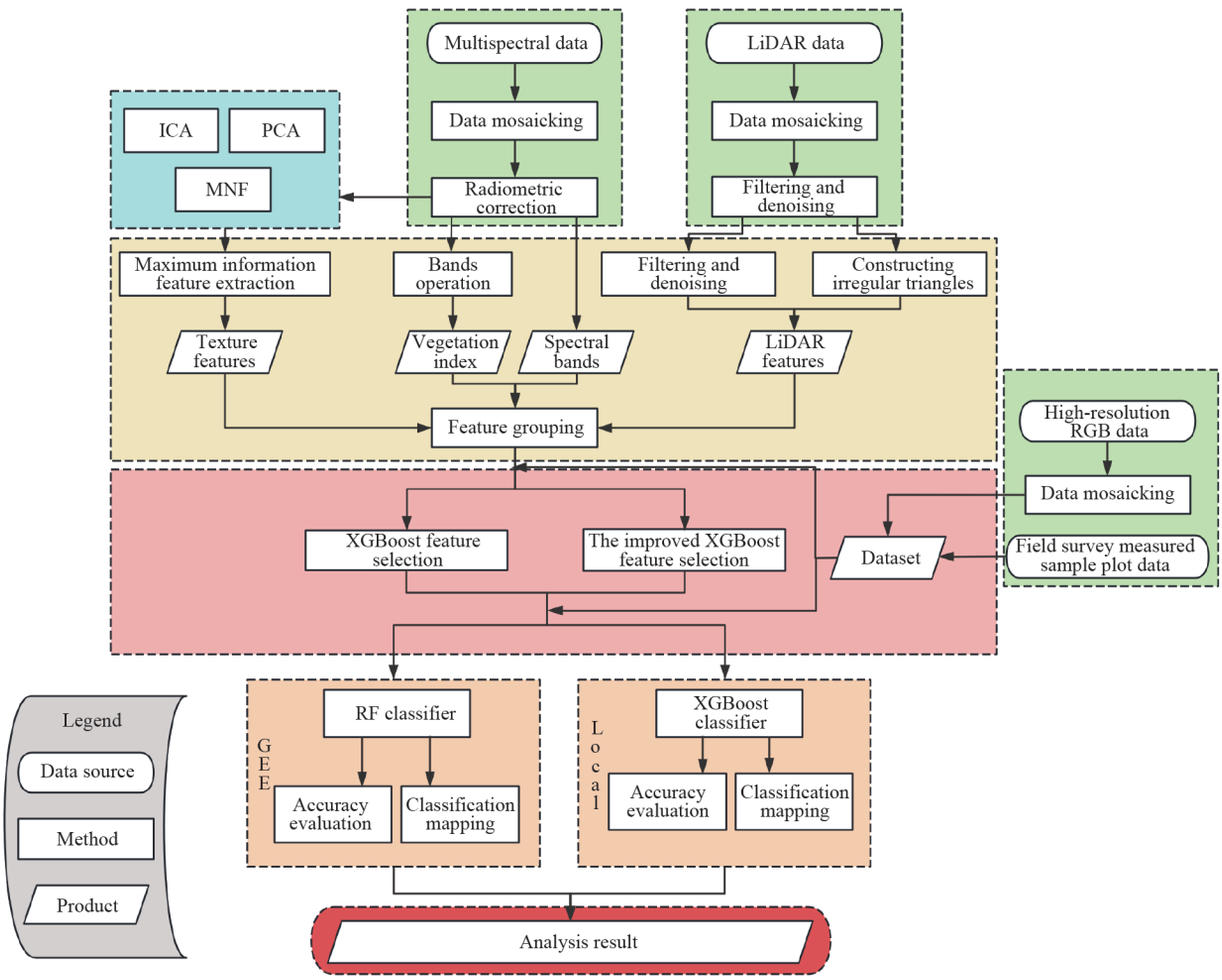


Fig.4 Flowchart of the experiment in this study

It can be seen from Table 4 that different pruning intensities significantly affect the performance of the F1 group: over-loose pruning groups (G1, G2) had UA_v at 0.000 10 due to residual highly redundant features, resulting in insufficient classification stability; Group G2 retained only 20 features, leading to insufficient information and an OA 0.81% lower than Group G5. Over-strict pruning groups (G3, G4) had OA below 97% due to excessive removal of complementary features.

Only the balanced Group G5 (70%+90% threshold) achieved a balance of "high OA-low distortion-moderate number of features": its OA was consistent with the core results of the main text (98.48%), UA_v was controlled at 0.000 05, and 28 features were retained. It not only effectively eliminated redundancy but also fully retained multi-source complementary information, confirming that this threshold is a key parameter for the F1 group to achieve optimal performance.

Figure 6a–f shows the results of RF classification using three hybrid feature groups processed by two feature selection methods. The highest PA values for each mangrove species were achieved with the following configurations: *Rhizophora stylosa* (PA= 95.07%) and *Sonneratia apetala* (PA=95.56%) using F2 features processed by the proposed method; *Bruguiera gymnorrhiza* (PA=91.12%) with F3 features processed by the proposed method; and *Avicennia marina* (PA=94.99%) with F1 features processed by the proposed method. In terms of UA, the peak performances were as follows: *Rhizophora stylosa* (UA=94.62%) and *Sonneratia apetala* (UA= 94.34%) using F3 features processed by the proposed method; *Bruguiera gymnorhiza* (UA= 90.63%) with F2 features processed by the proposed method; and *Avicennia marina* (UA=96.24%) with F1 features processed by XGBoost. Notably, F3 features group consistently exhibited minimal UA_v across species, with values of 0.000 77 under

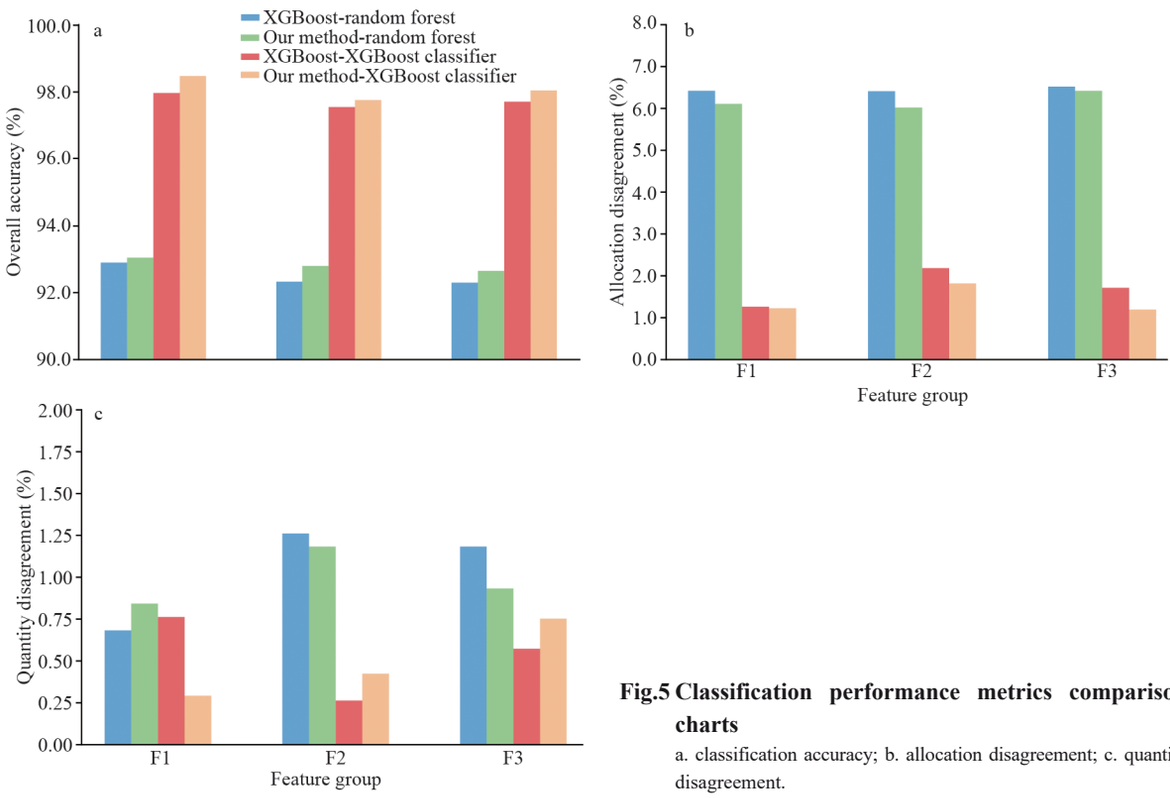


Fig.5 Classification performance metrics comparison charts

a. classification accuracy; b. allocation disagreement; c. quantity disagreement.

Table 4 Measured performance of F1 feature group under different pruning intensities

Group ID	Pruning intensity	Threshold combination (NDVI similarity+inter-feature correlation)	OA (%)	UA_v	Number of filtered features	AD (%)	QD (%)
G1	Over-loose 1	60%+80%	97.64	0.000 10	25	1.60	0.76
G2	Over-loose 2	60%+90%	97.67	0.000 10	20	1.71	0.62
G3	Over-strict 1	80%+90%	96.52	0.000 07	31	2.86	0.62
G4	Over-strict 2	80%+95%	96.88	0.000 03	33	2.57	0.55
G5	This study	70%+90%	98.48	0.000 05	28	1.23	0.29

XGBoost processing and 0.000 81 under the proposed method, representing the lowest variability observed in the study.

Figure 6g–l shows the results of XGBoost classification using three hybrid feature groups processed by two feature selection methods. The highest PA values for each mangrove species were achieved with the following configurations: *Rhizophora stylosa* (PA=98.61%) and *Bruguiera gymnorrhiza* (PA=98.75%) using F3 features processed by the proposed method; *Avicennia marina* (PA=99.03%) with F1 features processed by the proposed method; *Sonneratia apetala* (PA= 98.99%) with F2 features processed by the proposed method. In terms of UA, the peak performances were achieved as follows: *Rhizophora stylosa* (UA= 97.31%) and *Bruguiera gymnorrhiza* (UA=98.47%) using F1 features processed by the proposed method; *Avicennia marina* (UA=99.90%) with F3 features processed by XGBoost; and *Sonneratia apetala* (UA=98.98%) with F1 features processed by XGBoost. Notably, F1 feature groups consistently exhibited minimal UA_v across species, with values of 0.000 25 under XGBoost processing and 0.000 05 under the proposed method, representing the lowest variability observed in the study.

We can see in Table 5 the OA, AD, QD, and UA_v of the supplementary experiments. It is evident that although the same feature group and dataset were used, the GBDT classifier deployed on GEE was weaker than the RF classifier. The catboost classifier deployed locally is also weaker than the XGBoost classifier. Similar to its previous experiments, using the research method in this article for screening is

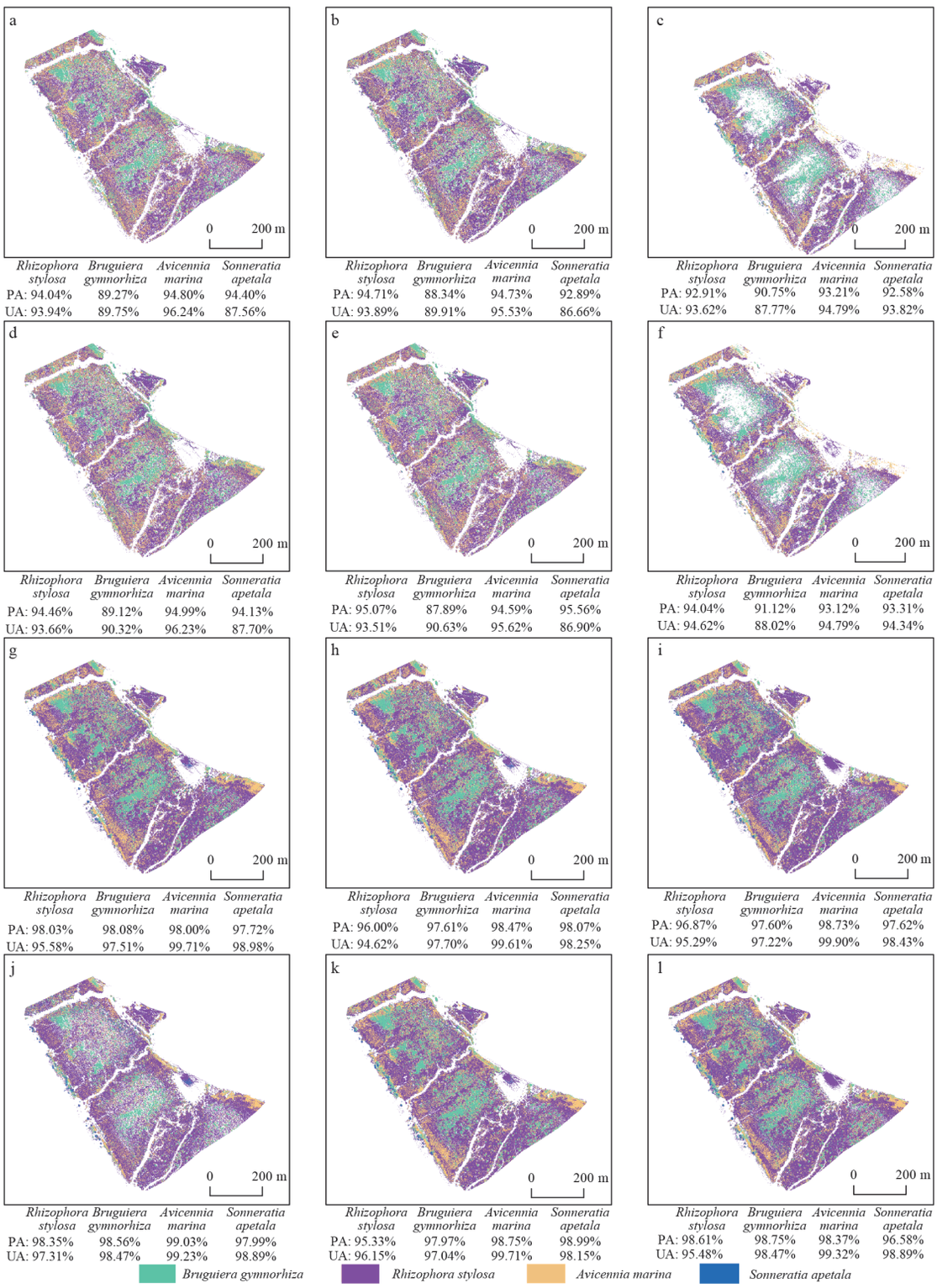


Fig.6 Mangrove species classification maps generated by 2 classifiers based on 3 feature combinations processed by 2 feature selection methods

a. RF classifier+F1 feature group+XGBoost method; b. RF classifier+F2 feature group+XGBoost method; c. RF classifier+F3 feature group+XGBoost method; d. RF classifier+F1 feature group+proposed method in this study; e. RF classifier+F2 feature group+proposed method in this study; f. RF classifier+F3 feature group+proposed method in this study; g. XGBoost classifier+F1 feature group+XGBoost method; h. XGBoost classifier+F2 feature group+XGBoost method; i. XGBoost classifier+F3 feature group+XGBoost method; j. XGBoost classifier+F1 feature group+proposed method in this study; k. XGBoost classifier+F2 feature group+proposed method in this study; l. XGBoost classifier+F3 feature group+proposed method in this study.

Table 5 Performance comparison of different classifiers supplementary experiments

Classifier-feature selection method	OA (%)	AD (%)	QD (%)	UA_v
CatBoost-original XGBoost	96.86	2.45	0.69	0.000 08
CatBoost-proposed method	96.91	2.45	0.64	0.000 07
GBDT(GEE)-original XGBoost	87.89	11.26	0.85	0.001 00
GBDT(GEE)-proposed method	92.75	6.54	0.71	0.000 80

superior to using the original XGBoost algorithm. This shows that the improved method has good robustness and can be adapted to different deployment platforms and classification frameworks.

3.2 Feature importance analysis

Given the consistent superiority of XGBoost classifiers over RF classifiers under identical feature selection methods and feature groups, this study focuses exclusively on XGBoost-based feature importance analysis for mangrove classification. Figure 6 reveals distinct patterns among the top 10 important features in XGBoost-selected feature groups. The F1 and F2 feature groups share eight common features, with group-specific distinctions rooted in dimensionality reduction techniques. Specifically, the F1 feature group exclusively incorporates PCA-derived low grey-level run emphasis and run percentage, while the F2 feature group exclusively includes ICA-derived low grey-level run emphasis and run percentage. The F3 feature group shares seven features with F1 and F2 but lacks the blue band, instead containing unique high-importance features: the Rededge720, the modified chlorophyll absorption ratio index 2 (MCARI2), and the MMF-derived mean values.

Similarly, for classification results based on feature groups selected by this methodology, the top 10 important features reveal that the F1 and F2 feature groups share eight common features. The differences between them arise from the dimensionality reduction techniques: specifically, the F1 feature group exclusively includes PCA-derived Run Percentage and Haralick Correlation, while the F2 feature group contains ICA-derived counterparts. The F3 feature group shares seven features with both F1 and F2 feature groups but lacks the blue band; instead, it contains unique high-importance features such as Rededge750, Rededge720, and NDVI.

Figure 7 also illustrates distinct feature selection patterns under comparable feature group conditions for both methodologies. When classifying with the F1 feature group, the top 10 important features selected by XGBoost and our methodology share eight common features. The unshared features are NDVIRE2 and MCARI1 for XGBoost, and RDVI and PCA-derived Haralick Correlation for our methodology. Similarly, for the F2 feature group, the top 10 features exhibit eight common features. XGBoost consistently selects NDVIRE2 and MCARI1, while our methodology selects RDVI and ICA-derived Haralick Correlation. For the F3 feature group, the top 10 features share only six common features. XGBoost selects NDVIRE2, MCARI1, MCARI2, and Mean (MMF), while our methodology selects RDVI, Rededge750, NDVI, and Low Grey-Level Run Emphasis (MNF).

4 DISCUSSION

4.1 Interpretation of classification performance

Comparative analysis of all 12 experimental groups reveals that XGBoost classifiers achieve superior OA, AD, and species-specific PA and UA metrics compared to RF classifiers when employing identical feature selection methods and feature groups. The only deviation was observed in the QD metric, where XGBoost classifiers underperformed RF classifiers when applied to the F1 feature group selected by the XGBoost algorithm. These findings align with the conclusions of Cai et al. (2019) and Pham et al. (2020), confirming the superior accuracy and robustness of XGBoost for mangrove remote sensing applications.

Under identical feature selection methods and classifiers, the F1 feature group consistently achieved superior OA compared to other feature groups. Over 50% of experimental results demonstrated better performance in QD and species-specific PA and UA with the F1 group, though AD slightly underperformed versus the F2 group. This indicates the F1 group's comprehensive superiority in providing effective informational support for classifiers. However, RF classification of XGBoost-processed F1 features exhibited higher variance in species-level UA compared to F3 groups, revealing classification distortion caused by disproportionately high accuracy in specific species.

The comprehensive advantage of the F1 group not only stems from the complementarity of multi-source data (spectral+LiDAR+texture) but also the rationality of its feature screening threshold as a key premise. Comparison of different pruning intensities

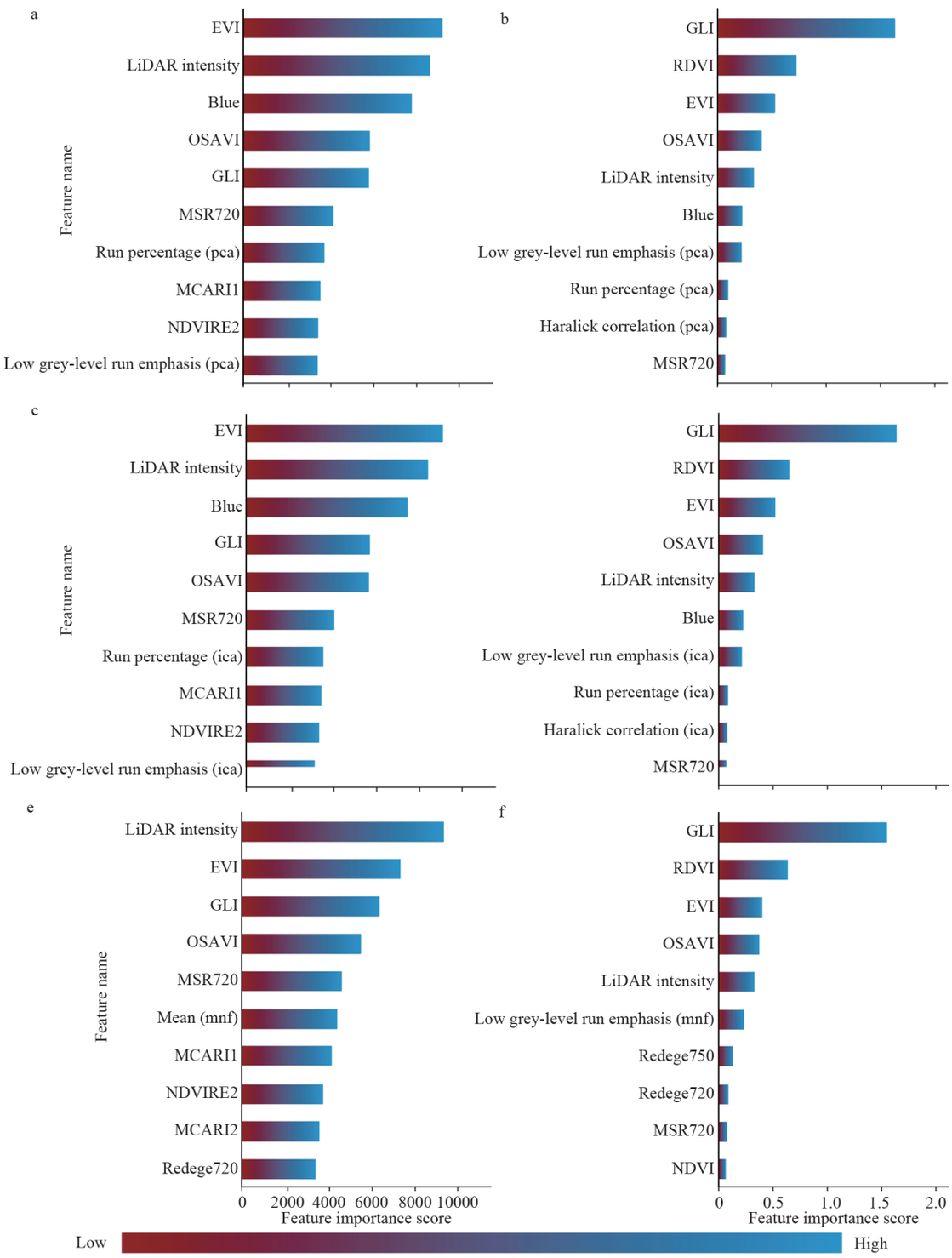


Fig.7 Top ten feature importance ordination plots

a. feature importance of F1 feature group selected by XGBoost algorithm; b. feature importance of F1 feature group selected by the proposed method in this study; c. feature importance of F2 feature group selected by XGBoost algorithm; d. feature importance of F2 feature group selected by the proposed method in this study; e. feature importance of F3 feature group selected by XGBoost algorithm; f. feature importance of F3 feature group selected by the proposed method in this study. EVI: enhanced vegetation index; OSAVI: optimized soil-adjusted vegetation index; GLI: green leaf index; MSR720: modified simple ratio at 720 nm; MCARI1: modified chlorophyll absorption in reflectance index 1; NDVIRE2: red edge 2 normalized difference vegetation index; RDVI: ratio difference vegetation index; MCARI2: modified chlorophyll absorption in reflectance index 2; NDVI: normalized difference vegetation index.

showed that: if the threshold was too loose, redundant features would interfere with the model's discrimination of similar species, leading to UA_v rising above 0.000 10 and weakening the synergistic effect of multi-source information; if the threshold was too strict, key features would be lost, resulting in an OA loss of over 1.6% and failing to reflect the complementary advantage of LiDAR structural information and spectral information.

This pattern of "over-loose leads to distortion, over-strict leads to inaccuracy" confirms the balanced value of 70% NDVI similarity+90% inter-feature correlation. This threshold not only eliminates highly redundant features but also fully retains key complementary features such as texture and LiDAR canopy height, providing parameter guarantee for the F1 group to exert the advantage of multi-source data. Ultimately, it supports the F1 group to achieve the highest OA of 98.48% and low UA_v of 0.000 05.

Under identical feature groups and classifiers, the proposed method outperformed XGBoost-based feature selection in key metrics, with more balanced accuracy across species. This improvement stems from addressing the inherent limitations of traditional feature selection, rather than simple parameter tuning.

Against complex feature interactions, traditional XGBoost relies primarily on feature weight to evaluate importance, which often leads to overemphasis on a single feature type. To fix this, our method enriches the evaluation logic by incorporating feature gain and coverage—two metrics that reflect a feature's actual discriminative contribution and sample representation, respectively. This refined evaluation mechanism (Eq.1) can effectively retain complementary features and accurately eliminate redundant ones, reducing cross-feature interference without losing critical information.

For classification result distortion, traditional feature selection methods often prioritize OA alone, which tends to sacrifice the classification performance of minority species, resulting in the common issue of "high OA masking low minority species accuracy". Our method addresses this by introducing user accuracy variance to guide feature selection: it ensures balanced accuracy across species by avoiding overemphasis on dominant species traits, thereby enabling reliable classification for all target mangrove species.

These optimizations align with the core demand of fine-scale mangrove classification (Fu et al., 2023), i.e., balancing feature utility and classification stability. Unlike rule-based screening or single-metric feature selection methods, our approach adapts to the complexity of multi-source UAV data, which explains its superior performance in practical applications.

4.2 Significance of feature importance patterns

Accuracy analysis indicates that for the comprehensively optimal F1 group, classification using the proposed methodology exhibits marginally lower UA for *Rhizophora stylosa* and *Sonneratia apetala* compared to XGBoost. However, it demonstrates marked superiority in all other metrics, particularly higher OA and significantly reduced species-level UA_v. For the F2 and F3 feature groups, evaluation metrics confirm that the proposed methodology outperforms XGBoost in over half of the experiments. Collectively, under comparable hybrid feature group conditions, the improved method shows superior performance for mangrove fine classification, with consistently higher OA and reduced species-level UA_v further attesting to its enhanced capability to mitigate classification distortion.

4.3 Limitation of the classification methods

Regarding data processing, a fixed 3×3 window size was used for texture feature extraction across all dimensionality reduction methods (PCA, ICA, MNF) to ensure standardization. Variations in window size may differentially impact classification outcomes across mangrove species, suggesting that parameter fine-tuning could improve accuracy. For classifier optimization, adaptive grid-search cross-validation was conducted on a subset of the dataset to achieve optimal parameterization. Although this subset was excluded from subsequent training, testing, and prediction phases, its use slightly compromises classifier generalizability.

Future research will address these limitations by expanding the sampling scope to include multiple mangrove ecosystems across different latitudes and climatic zones, enhancing the spatial representativeness and robustness of the classification model. Additionally, incorporating other machine learning algorithms such as LightGBM and CatBoost will enable adaptive optimization of feature extraction parameters based on species-specific traits. Integrating multi-temporal UAV data will help capture seasonal dynamics in mangrove vegetation, while developing a standardized

spectral and structural feature library specific to coastal wetlands will improve the transferability of the method to other similar ecosystems.

4.4 Practical implication

The improved XGBoost method integrated with multi-source UAV data effectively addresses key challenges in fine-scale mangrove classification, including complex feature interactions and classification result distortion. It enables reliable identification of both dominant and non-dominant mangrove species, providing essential support for targeted ecological protection and restoration. Compared with traditional monitoring approaches, this method is more efficient for fine-scale mangrove surveys. Additionally, its adaptability to diverse scenarios, such as local computing environments, cloud platforms like GEE, and different classifiers, has been verified, allowing extension to larger-scale mangrove monitoring and serving as a practical, high-precision tool for mangrove ecological management.

5 CONCLUSION

This study conducted fine-scale mangrove classification in the Gaoqiao Mangrove Nature Reserve, Zhanjiang, Guangdong Province, using multi-source UAV data. We constructed hybrid feature sets, performed feature selection via an improved XGBoost algorithm, and evaluated classification performance using XGBoost and RF classifiers. The key conclusions are as follows:

(1) By optimizing XGBoost's feature importance metrics, the improved method enhances multi-source feature screening. When selecting the F1 group (PCA-texture+spectral+vegetation index+LiDAR features) and paired with the XGBoost classifier, it achieves optimal performance: 98.48% OA, 1.23% AD, and 0.29% QD, outperforming the original XGBoost and RF models. This validates its effectiveness in extracting critical features for fine-scale mangrove classification.

(2) Traditional multi-feature fusion may cause distortion, but the improved method reduces inter-species UA_v via optimized selection logic. This advantage holds across different feature groups and classifiers: under identical feature groups and classifiers, the improved method consistently outperforms the original XGBoost in over 50% of experiments, with higher OA, lower AD, and superior species-specific PA/UA. This indicates its robustness and universality, not limited to specific feature or classifier combinations.

(3) The improved XGBoost-based feature selection method proposed in this study, combined with multi-source UAV data, not only achieves high-accuracy fine-scale mangrove classification but also exhibits strong scenario adaptability. Through cloud-based platform validation, the method still maintained stable performance—for example, the GBDT classifier on GEE, when using features selected by the improved method, achieved higher accuracy than that with the original XGBoost. This adaptability allows the method to be extended from small-scale surveys to regional-scale mangrove monitoring, further reinforcing its practical value for mangrove ecological conservation, restoration, and long-term dynamic assessment.

In summary, this study addresses key challenges in fine-scale mangrove classification, such as "accuracy bottlenecks due to complex feature interactions" and "classification result distortion", through improved feature selection algorithms and multi-source data fusion strategies. Its adaptability across features and classifiers, combined with the validated feasibility of UAV data processing on GEE, further enhances practical value, providing high-precision technical support for mangrove species identification, ecological conservation, and restoration.

We recognize, however, that the current study is not without limitations: although the dataset and sampling design are sufficient for preliminary validation of the proposed method, the study is confined to a single mangrove reserve, which may limit the model's generalizability to other geographical contexts with varying environmental conditions, and the fixed window size used for texture feature extraction, while standardized, may not fully capture the morphological characteristics of different mangrove species with varying canopy structures. For future research directions, we aim to expand the sampling scope to include multiple mangrove ecosystems across different latitudes and climatic zones to enhance the spatial representativeness and robustness of the classification model; we also plan to incorporate other machine learning algorithms such as LightGBM and CatBoost to enable adaptive optimization of feature extraction parameters based on species-specific traits. Furthermore, integrating multi-temporal UAV data will help capture seasonal dynamics in mangrove vegetation, while developing a standardized spectral and structural feature library specific to coastal wetlands will improve the

transferability of the method to other similar ecosystems.

6 DATA AVAILABILITY STATEMENT

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

7 ACKNOWLEDGMENT

Thank our supervisor and classmates for their assistance in this study.

References

Abraham J, Cheng L J, Mann M E et al. 2022. The ocean response to climate change guides both adaptation and mitigation efforts. *Atmospheric and Oceanic Science Letters*, **15**(4): 100221, https://doi.org/10.1016/j.aosl.2022.100221.

Alongi D M. 2014. Carbon cycling and storage in mangrove forests. *Annual Review of Marine Science*, **6**(1): 195-219, https://doi.org/10.1146/annurev-marine-010213-135020.

Cai L F, Wu D S, Fang L M et al. 2019. Tree species identification using XGBoost based on GF-2 images. *Forest Resources Management*, (5): 44-51, https://doi.org/10.13466/j.cnki.lyzygl.2019.05.009. (in Chinese with English abstract)

Cao J J, Leng W C, Liu K et al. 2018. Object-based mangrove species classification using unmanned aerial vehicle hyperspectral images and digital surface models. *Remote Sensing*, **10**(1): 89, https://doi.org/10.3390/rs10010089.

Cao J J, Liu K, Zhuo L et al. 2021. Combining UAV-based hyperspectral and LiDAR data for mangrove species classification using the rotation forest algorithm. *International Journal of Applied Earth Observation and Geoinformation*, **102**: 102414, https://doi.org/10.1016/j.jag.2021.102414.

Chen G W, Jin R J, Ye Z J et al. 2022. Spatiotemporal mapping of salt marshes in the intertidal zone of China during 1985-2019. *Journal of Remote Sensing*, **2022**: 9793626, https://doi.org/10.34133/2022/9793626.

Chen T Q, Guestrin C. 2016. XGBoost: a scalable tree boosting system. *In*: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery, San Francisco, California, USA. p. 785-794, https://doi.org/10.1145/2939672.2939785.

Collin A, Lambert N, Etienne S. 2018. Satellite-based salt marsh elevation, vegetation height, and species composition mapping using the superspectral WorldView-3 imagery. *International Journal of Remote Sensing*, **39**(17): 5619-5637, https://doi.org/10.1080/01431161.2018.1466084.

Elmahdy S I, Ali T A, Mohamed M M et al. 2020. Spatiotemporal mapping and monitoring of mangrove forests changes from 1990 to 2019 in the Northern Emirates, UAE using random forest, Kernel logistic regression and Naive Bayes Tree models. *Frontiers in Environmental Science*, **8**: 102, https://doi.org/10.3389/fenvs.2020.00102.

Fu B L, Kuang H Y, Wu Y et al. 2024a. Mangrove species classification using novel adaptive ensemble learning with multi-spatial-resolution multispectral and full-polarization SAR images. *International Journal of Digital Earth*, **17**(1): 2346277, https://doi.org/10.1080/17538947.2024.2346277.

Fu B L, Liang Y Y, Lao Z N et al. 2023. Quantifying scattering characteristics of mangrove species from Optuna-based optimal machine learning classification using multi-scale feature selection and SAR image time series. *International Journal of Applied Earth Observation and Geoinformation*, **122**: 103446, https://doi.org/10.1016/j.jag.2023.103446.

Fu B L, Zhang S R, Li H J et al. 2024b. Exploring the effects of different combination ratios of multi-source remote sensing images on mangrove communities classification. *International Journal of Applied Earth Observation and Geoinformation*, **134**: 104197, https://doi.org/10.1016/j.jag.2024.104197.

Ghorbanian A, Zaghian S, Asiyabi R M et al. 2021. Mangrove ecosystem mapping using Sentinel-1 and Sentinel-2 satellite images and random forest algorithm in Google Earth Engine. *Remote Sensing*, **13**(13): 2565, https://doi.org/10.3390/rs13132565.

Guo X X, Wang M, Jia M M et al. 2021. Estimating mangrove leaf area index based on red-edge vegetation indices: a comparison among UAV, WorldView-2 and Sentinel-2 imagery. *International Journal of Applied Earth Observation and Geoinformation*, **103**: 102493, https://doi.org/10.1016/j.jag.2021.102493.

Han D, Lee C. 2021. Mapping of coastal zones using unmanned aerial vehicle (UAV) video sequences. *Journal of Coastal Research*, **114**(Suppl 1): 400-404, https://doi.org/10.2112/JCR-SI114-081.1.

He Z, Shi Q, Liu K et al. 2020. Object-oriented mangrove species classification using hyperspectral data and 3-D Siamese residual network. *IEEE Geoscience and Remote Sensing Letters*, **17**(12): 2150-2154, https://doi.org/10.1109/LGRS.2019.2962723.

Hu L J, Xu N, Liang J et al. 2020. Advancing the mapping of mangrove forests at national-scale using Sentinel-1 and Sentinel-2 time-series data with Google Earth Engine: a case study in China. *Remote Sensing*, **12**(19): 3120, https://doi.org/10.3390/rs12193120.

Huang Z C, Yeh C Y, Tseng K H et al. 2018. A UAV-RTK Lidar system for wave and tide measurements in coastal zones. *Journal of Atmospheric and Oceanic Technology*, **35**(8): 1557-1570, https://doi.org/10.1175/JTECH-D-17-0199.1.

Jiang X P, Gao Z Q, Wang Z C. 2024. Estimation and verification of green tide biomass based on UAV remote sensing. *Journal of Oceanology and Limnology*, **42**(4): 1216-1226, https://doi.org/10.1007/s00343-023-3113-6.

Lassalle G, Ferreira M P, La Rosa L E C et al. 2023. Advances in multi-and hyperspectral remote sensing of mangrove species: a synthesis and study case on airborne and multisource spaceborne imagery. *ISPRS*

Journal of Photogrammetry and Remote Sensing, **195**: 298-312, https://doi.org/10.1016/j.isprsjprs.2022.12.003.

Li L F, Liu W, Ai J W et al. 2023. Predicting mangrove distributions in the Beibu Gulf, Guangxi, China, using the MaxEnt model: determining tree species selection. *Forests*, **14**(1): 149, https://doi.org/10.3390/f14010149.

Li Q S, Wong F K K, Fung T. 2019. Classification of mangrove species using combined WordView-3 and LiDAR data in Mai Po Nature Reserve, Hong Kong. *Remote Sensing*, **11**(18): 2114, https://doi.org/10.3390/rs11182114.

Liu D Z, Li S S, Fu D Y et al. 2018. Remote sensing analysis of mangrove distribution and dynamics in Zhanjiang from 1991 to 2011. *Journal of Oceanology and Limnology*, **36**(5): 1597-1603, https://doi.org/10.1007/s00343-018-7004-1.

Lu F, Chen Y, Liu W Q et al. 2025. Regional geochemical baseline establishment, heavy metal pollution assessment and investigation of its variation in response to human activities in mangrove intertidal sediments in Hainan, China. *Journal of Oceanology and Limnology*, published Online First, June 2025, https://doi.org/10.1007/s00343-025-4228-8.

Lyons M B, Keith D A, Phinn S R et al. 2018. A comparison of resampling methods for remote sensing classification and accuracy assessment. *Remote Sensing of Environment*, **208**: 145-153, https://doi.org/10.1016/j.rse.2018.02.026.

Murdiyarso D, Purbopuspito J, Kauffman J B et al. 2015. The potential of Indonesian mangrove forests for global climate change mitigation. *Nature Climate Change*, **5**(12): 1089-1092, https://doi.org/10.1038/nclimate2734.

Peng L H, Liu K, Cao J J et al. 2020. Combining GF-2 and RapidEye satellite data for mapping mangrove species using ensemble machine-learning methods. *International Journal of Remote Sensing*, **41**(3): 813-838, https://doi.org/10.1080/01431161.2019.1648907.

Pham T D, Yokoya N, Xia J S et al. 2020. Comparison of machine learning methods for estimating mangrove above-ground biomass using multiple source remote sensing data in the Red River Delta biosphere reserve, Vietnam. *Remote Sensing*, **12**(8): 1334, https://doi.org/10.3390/rs12081334.

Richter R, Reu B, Wirth C et al. 2016. The use of airborne hyperspectral data for tree species classification in a species-rich Central European forest area. *International Journal of Applied Earth Observation and Geoinformation*, **52**: 464-474, https://doi.org/10.1016/j.jag.2016.07.018.

Salum R B, Robinson S A, Rogers K. 2021. A validated and accurate method for quantifying and extrapolating mangrove above-ground biomass using LiDAR data. *Remote Sensing*, **13**(14): 2763, https://doi.org/10.3390/rs13142763.

Santos I R, Maher D T, Larkin R et al. 2019. Carbon outwelling and outgassing vs. burial in an estuarine tidal creek surrounded by mangrove and saltmarsh wetlands. *Limnology and Oceanography*, **64**(3): 996-1013, https://doi.org/10.1002/lno.11090.

Tian Y C, Zhang Q, Huang H et al. 2022. Aboveground biomass of typical invasive mangroves and its distribution patterns using UAV-LiDAR data in a subtropical estuary: Maoling River estuary, Guangxi, China. *Ecological Indicators*, **136**: 108694, https://doi.org/10.1016/j.ecolind.2022.108694.

Wang D Z, Wan B, Qiu P H et al. 2022. Mapping mangrove species using combined UAV-LiDAR and Sentinel-2 data: feature selection and point density effects. *Advances in Space Research*, **69**(3): 1494-1512, https://doi.org/10.1016/j.asr.2021.11.020.

Wang F M, Sanders C J, Santos I R et al. 2021. Global blue carbon accumulation in tidal wetlands increases with climate change. *National Science Review*, **8**(9): nwaa296, https://doi.org/10.1093/nsr/nwaa296.

Wang L, Jia M M, Yin D M et al. 2019. A review of remote sensing for mangrove forests: 1956-2018. *Remote Sensing of Environment*, **231**: 111223, https://doi.org/10.1016/j.rse.2019.111223.

Warrens M J. 2015. Properties of the quantity disagreement and the allocation disagreement. *International Journal of Remote Sensing*, **36**(5): 1439-1446, https://doi.org/10.1080/01431161.2015.1011794.

Xia Q, Jia M M, He T T et al. 2021. Effect of tide level on submerged mangrove recognition index using multi-temporal remotely-sensed data. *Ecological Indicators*, **131**: 108169, https://doi.org/10.1016/j.ecolind.2021.108169.

Xin K, Huang X, Hu J L et al. 2014. Land use change impacts on heavy metal sedimentation in mangrove wetlands—a case study in Dongzhai Harbor of Hainan, China. *Wetlands*, **34**(1): 1-8, https://doi.org/10.1007/s13157-013-0472-3.

Xu Y, Zhen J N, Jiang X P et al. 2021. Mangrove species classification with UAV-based remote sensing data and XGBoost. *National Remote Sensing Bulletin*, **25**(3): 737-752, https://doi.org/10.11834/jrs.20210281. (in Chinese with English abstract)

Yin D M, Wang L. 2019. Individual mangrove tree measurement using UAV-based LiDAR data: possibilities and challenges. *Remote Sensing of Environment*, **223**: 34-49, https://doi.org/10.1016/j.rse.2018.12.034.

Yu C X, Guan D S, Gang W et al. 2021. Development of ecosystem carbon stock with the progression of a natural mangrove forest in Yingluo Bay, China. *Plant and Soil*, **460**(1): 391-401, https://doi.org/10.1007/s11104-020-04819-3.

Zeng J L, Ai B, Jian Z K et al. 2023. Analysis of mangrove dynamics and its protection effect in the Guangdong-Hong Kong-Macao Coastal Area based on the Google Earth Engine platform. *Frontiers in Marine Science*, **10**: 1170587, https://doi.org/10.3389/fmars.2023.1170587.

Zhao C P, Jia M M, Wang Z M et al. 2023. Identifying mangroves through knowledge extracted from trained random forest models: an interpretable mangrove mapping approach (IMMA). *ISPRS Journal of Photogrammetry and Remote Sensing*, **201**: 209-225, https://doi.org/10.1016/j.isprsjprs.2023.05.025.

Zhen J N, Mao D H, Shen Z et al. 2024. Performance of XGBoost ensemble learning algorithm for mangrove species classification with multisource spaceborne remote sensing data. *Journal of Remote Sensing*, **4**: 0146, https://doi.org/10.34133/remotesensing.0146.