JARS Journal of Applied Remote Sensing

RESEARCH PAPER

# Mangrove ecosystem species mapping from integrated Sentinel-2 imagery and field spectral data using random forest algorithm

**Nirmawana Simarmata,[a,b,c] Ketut Wikantika,[a,c] Soni Darmawan,[c,d,*] Agung Budi Harto,[a,c] Anjar Dimara Sakti,[a,c] and Aki Asmoro Santo[b]**

[a]Institut Teknologi Bandung, Geodesy and Geomatic Engineering, Faculty of Earth Sciences and Technology, Bandung, Indonesia

[b]Institut Teknologi Sumatera, Department of Geomatics Engineering, Lampung, Indonesia

[c]Institut Teknologi Bandung, Center for Remote Sensing, Bandung, Indonesia

[d]Institut Teknologi Nasional, Department of Geodesy Engineering, Faculty of Civil Engineering and Planning, Bandung, Indonesia

**ABSTRACT.** Mangroves maintain coastal balance and have the greatest potential for carbon sequestration. Most mapping studies on mangroves have focused on their extent and distribution and rarely featured mangrove species. Therefore, the objective of our study is to investigate mangrove species mapping from integrated Sentinel-2 imagery and field spectral data using a random forest (RF) algorithm. Study areas are located in East and South Lampung, Indonesia. The field samples used represented 144 points of mangrove species. The classification method used an RF algorithm and four models with varying parameters: model 1 with Sentinel-2; model 2 with both Sentinel-2 and field spectral data; model 3 with Sentinel-2, field spectral data, and spectrally transformed data; and model 4 only with spectrally transformed data. The results showed that *Rhizophora mucronata*, *Sonneratia alba*, *Avicennia lanata*, and *Avicennia marina* were the most common mangrove species in these areas, with reflectance values in the range of 0.002 to 0.493, 0.006 to 0.833, 0.014 to 0.768, and 0.002 to 0.758. Permutation importance (PI) that affects the classification model is the red band, near-infrared, and green normalized difference vegetation index, where the most PI in model 3 is 0.283. The highest level of agreement for mangrove species is found in model 3. Model 3 is the best parameter for RF classification that showed the best mapping accuracy, with the overall accuracy, user accuracy, producer accuracy, and kappa value being 81.25%, 81.68%, 81.25%, and 0.80, respectively.



**Keywords:** wetland; machine learning; field spectroradiometer; spectral reflectance; near-infrared

Paper 230288G received Jun. 10, 2023; revised Jan. 3, 2024; accepted Jan. 9, 2024; published Feb. 6, 2024.

## 1 Introduction

Mangrove ecosystems support various services,[1–3] including nutrient cycling, and serve as a vehicle for fisheries production that dominates the areas between tides along tropical[4] and subtropical coasts.[5–7] Despite mangroves making only up 0.7% of tropical forests worldwide,[8,9] they contribute ~50% of carbon stocks.[10,11] Human activities have reduced the area of mangrove

*Address all correspondence to Soni Darmawan, soni_darmawan@itenas.ac.id

forests by 30% to 50% in the last half century due to excessive exploitation/logging,[2,11,12] conversion into pond areas, and the development of the areas.[2,13]

Lampung is a coastal area covered by various mangroves[14] including *Rhizophora mucronata*, which grows on muddy soil types and, occasionally, on sandy reefs.[15] *R. mucronata* is distributed in the East Coast of Lampung, mainly in Pasir Sakti, East Lampung Regency. *Sonneratia alba* is another mangrove that grows in sandy mud. The lower leaves of *Avicennia lanata* and *A. marina* exhibit a similar morphology while being salty, exhibiting elliptical leaf tips,[16] and being distributed in sandy areas with fine mud near estuaries in Ketapang subdistrict, South Lampung.[17]

Given the several types of scattered species that reflect the development and condition of mangroves, it is necessary to identify and classify mangroves. Several methods have been used previously for identification of mangrove by only *in situ* measurements[18,19] or by utilizing satellite imagery.[20,21] *In situ* mangrove monitoring provides the most accurate information on mangrove distribution; however, data collection through field surveys remains challenging due to limited accessibility to mangroves.[22] Mangroves are located in relatively wet areas and subjected to high tides.[2] Thus remote sensing has become a practical way to map and monitor mangrove forests. Remote sensing provides land cover information using pixel-based analysis.[23]

Furthermore, integration of artificial intelligence, mathematical algorithms, and big data analysis with high-resolution sensing imaging data has become more common.[24,25] These data can be collected on a regular basis over a wide geographic area, enabling precise, and accurate monitoring of mangrove forest ecosystems.[22] The use of remote sensing technologies is expanding, along with the demand for spatial data. Remote sensing data are essential for extracting parameters and biophysical data in identifying mangrove forests.[26] These data in coastal areas can be utilized to monitor mangroves. Passive (optical) and active system synthetic aperture radar images are the two forms of remote sensing images that can be used for land monitoring.[26]

The classification and segmentation of coastal objects, including mangrove cover and tidal marsh, allows for determining the extent of each object. Using machine learning techniques, including support vector machines (SVM),[5,22,25–28] support vector regression, artificial neural network,[29] random forest (RF),[22,30,31] decision tree, symbolic regression,[32] extreme gradient boosting regression,[33,34] light gradient boosting machine (LightGBM), and extreme gradient boosting (XGBoost), allows for the retrieval of data on mangrove distribution. RF has an excellent biomass modeling ability[35] and can increase the precision of land cover mapping, wetland mapping,[36] and mangrove species classification.[5,34,37]

Mangrove identification is performed based on mangrove canopy properties. This can be analyzed using vegetation index transformation.[24,38] One technique for changing vegetation indices is the green normalized difference vegetation index (GNDVI).[17] This technique shows parameters related to vegetation,[28,39] such as green foliage biomass and area, which are essential for vegetation division. In addition, since mangroves are located in relatively wet areas, a moisture index, such as the normalized difference moisture index (NDMI)[17] and a wetness index, such as the normalized difference water index (NDWI) can be used to accurately identify them.

Earlier studies have not differentiated mangrove species and focused only on mangrove forest size and distribution.[40] Mangrove species composition and distribution data are critical to understanding mangrove ecosystem functions and ensuring sustainable mangrove conservation.[29] However, mangroves of a single species typically form tiny patches or thin strips that are invisible on satellite images. Furthermore, using machine learning models and remote sensing data for mangrove species mapping is challenging since there is no clear zoning between species due to the spectral similarity of mangroves.[28] This underlies the importance of mangrove mapping at the species level. Correctly identifying species will provide insights into the relationships between them. In addition, mangrove species classification allows the monitoring of a particular species population over time. This will help in detecting changes in the population size, distribution, and health of mangrove species. Previous studies have used machine learning models for mangrove species mapping using SVM classification with Worldview-2 images and aerial photographs.[28,41] Behera et al.[38] identified the mangrove species *Heritiera fomes*, *Excoecaria agallocha*, and *Avicennia officinalis* using reflectance/backscatter bands and vegetation indices derived from Sentinel-2, AVIRIS-NG, and Sentinel-1. Meanwhile, Paramanik et al.[42] only used AVIRIS-NG hyperspectral imagery, which successfully identified the species *H. fomes,*

*E. agallocha*, and *A. officinalis* and two of their combinations (*H. fomes–E. agallocha and A. officinalis–E. agallocha*), and Ghorbanian et al.[22] were limited to mapping mangrove ecosystems using Sentinel-1 and Sentinel-2 images using RF. For developing the classification model, the parameter used was vegetation index;[38] this is in contrast to the present study, which not only used the vegetation index but also the wetness and moisture indices. In addition, the directly measured reflectance value was also a parameter used for further modeling, which improved the accuracy of classification mapping.

Mangrove species can be categorized by changing the algorithm or applying new information.[28] Mapping and classifying mangrove species can be accomplished using a combination of RF algorithms and field spectroradiometers. Herein, the leaf area index,[43] vegetation texture and index,[44] humidity, and wetness were used as parameters. High-resolution data are useful for classifying mangrove species; nonetheless, they are not available for all areas.[45]

Previous studies have investigated mangrove reflectance using satellite data; however, observations of the spectral properties of mangrove biota using satellite images are insufficient, owing to the challenge of recognizing mangrove species solely from the canopy.[46] This is supported by previous research showing that object reflectance using a field spectroradiometer can be directly measured for mangrove species mapping.[45] Other findings have indicated that each mangrove species has a unique spectral reflectance and can be easily identified and mapped with adjacent wavelength bands in the near-infrared (NIR) region.[15]

Information on mangrove species is essential for mangrove management; however, insufficient research related to species classification has been conducted, especially in Lampung, Indonesia. Therefore, the objective of this research is to investigate mangrove species mapping using integrated Sentinel-2 imagery and field spectral data processed by an RF algorithm. The novelty of this study lies in the determination of the best parameters for RF classification, thus enhancing the accuracy of identifying mangrove species for the management, monitoring, and rehabilitation of mangroves in study areas.

## 2 Materials and Methods

### 2.1 Study Site

The study was chosen along the coastlines of Lampung, Indonesia, specifically targeting the east[47] and south areas and located between 5°25′30″S–5°51′0″S and 105°30′0″E–105°52′0″E. Tidal swamp plains are found along the east coast with an elevation of ~0.5 to 1 m above the mean sea level, and sedimentation areas based on the rising tide characterize the research site in these two districts (East and South Lampung) (Fig. 1).

### 2.2 Datasets

#### 2.2.1 *Image data*

The image data used in this study were from the Sentinel-2A image, which has 13 channels and various resolutions.[48] Sentinel-2A imagery is suitable for classifying land cover as they contain four channels with a spatial resolution of 10 m. Sentinel-2A observes the Earth in various spectral ranges, including visible light and NIR. The Sentinel-2A data parameters are listed in Table 1.

#### 2.2.2 *Field data*

The spectral values of each mangrove species were collected from August 6 to 15, 2022. The sample locations and number of sample points were determined based on the raw pixel values of mangrove objects on Sentinel-2A, specifically the red band, green band, and NIR band.[45] Additionally, the determination considers access to the sampling location as well since some locations are difficult to reach. The pixel values obtained are plotted onto a map to determine the reflectance value of mangrove species in the field.[45] The obtained reflectance data were used for preparing a spectral library that served as the basis for the construction of classification models. The number of trees taken was 144, with an average tree height of 7 to 15 m. Based on the results of field measurements, four mangrove species were observed, namely:

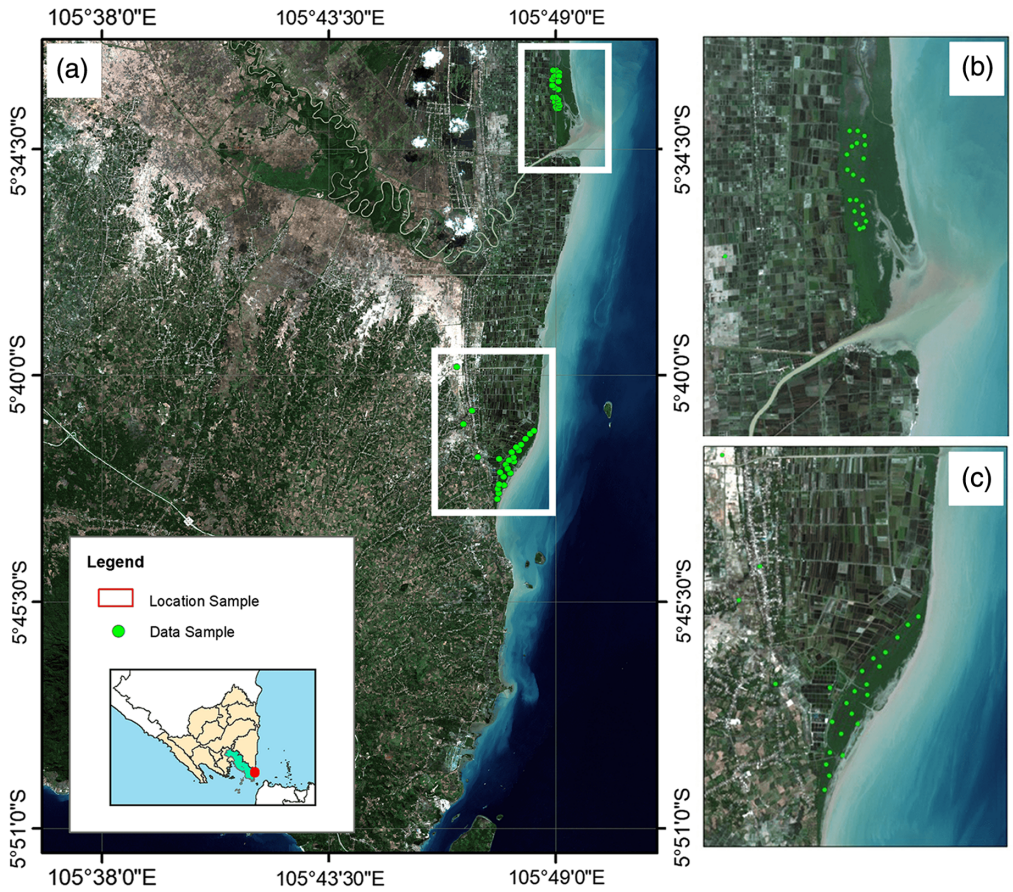


**Fig. 1** (a) Study area (RGB Sentinel-2A: August 6–15, 2022), (b) sampling points in Pasir Sakti, and (c) sampling points in Ketapang.

**Table 1** Sentinel-2A data parameters.

| Parameter | Dataset | Spatial resolution (m) |
|---|---|---|
| Red band (1) | Sentinel-2A imagery | 10 |
| Green band (2) | Sentinel-2A imagery | 10 |
| NIR band (3) | Sentinel-2A imagery | 10 |
| Red band (4) | Field spectrometer | 10 |
| Green band (5) | Field spectrometer | 10 |
| NIR band (6) | Field spectrometer | 10 |
| Density (GNDVI) (7) | Sentinel-2A imagery | 10 |
| Humidity (NDMI) (8) | Sentinel-2A imagery | 10 |
| Water index (NDWI) (9) | Sentinel-2A imagery | 10 |

*A. marina, A. lanata, S. alba*, and *R. mucronata.* Each species was observed in several different locations by measuring the diameter at breast height (DBH) of several trees in each location. The average DBH of *S. alba* in Ketapang was 55 cm, that of *R. mucronata* in Pasir Sakti was 17 cm, and that of both *A. marina* and *A. lanata* in Ketapang was 91 cm. The number of samples in each plot varies depending on the size of the tree as the larger the tree DBH, the fewer the number of trees. The range of the number of trees in a sample plot was 10 to 15 trees.

The sample points were evenly distributed in mangrove areas to represent each mangrove species. The spectral reflectance of the mangroves was measured using a field spectroradiometer,[45] a commonly used method to analyze the spectrum of light reflected or emitted.[49] Reflectance data were collected at the leaf level by selecting mangrove leaves that represent various species or conditions in the area. The spectroradiometer was placed above the leaf to measure the reflectance of light at various wavelengths, ranging from ultraviolet (315 nm) to SWIR (1100 nm).

The sample plots are determined based on the regulations of the head of the geospatial information agency number 3 of 2014, which include technical guidelines for collecting and processing mangrove geospatial data as well as the Sentinel-2A imagery data. The plot size used is $10 \times 10$ m.[50] Each plot had a different number of samples based on the diameter and size of the mangrove canopy. The number of samples in a plot was 5 to 10. The types of species in the study area were similar in some areas because they shared common traits that grouped them together in certain parts of the region. Measurements using a field spectroradiometer can identify mangrove species based on the resulting spectrum patterns.

### 2.3 Methodological Approach

SNAP 9.0.0 version was used for Sentinel-2A image preprocessing, whereas Endmapbox was used for processing of field spectroradiometer data, and quantum GIS was used for RF classification. The spectral bands used were red, green, and NIR. The methodology included radiometric and geometric correction, vegetation index transformation, moisture, and water indices, direct spectral measurements in the field, classification, accuracy testing, and mangrove species mapping. Fig. 2 outlines the research process.

Accuracy testing with a confusion matrix method is required for RF algorithm classification.[51] It is possible to determine whether the classification results are sourced from two separate classes by comparing the derivatives of each predicted class in the matrix to the derivative of the actual class in each row. For remote sensing picture classification, the most effective and practical validation tool is the confusion matrix method.[52] The vast majority of operations are consolidated into the error matrix, which use producer accuracy (PA), user accuracy (UA), and overall accuracy (OA) as indicators; this method is therefore successful.[53] A test of the classification results'

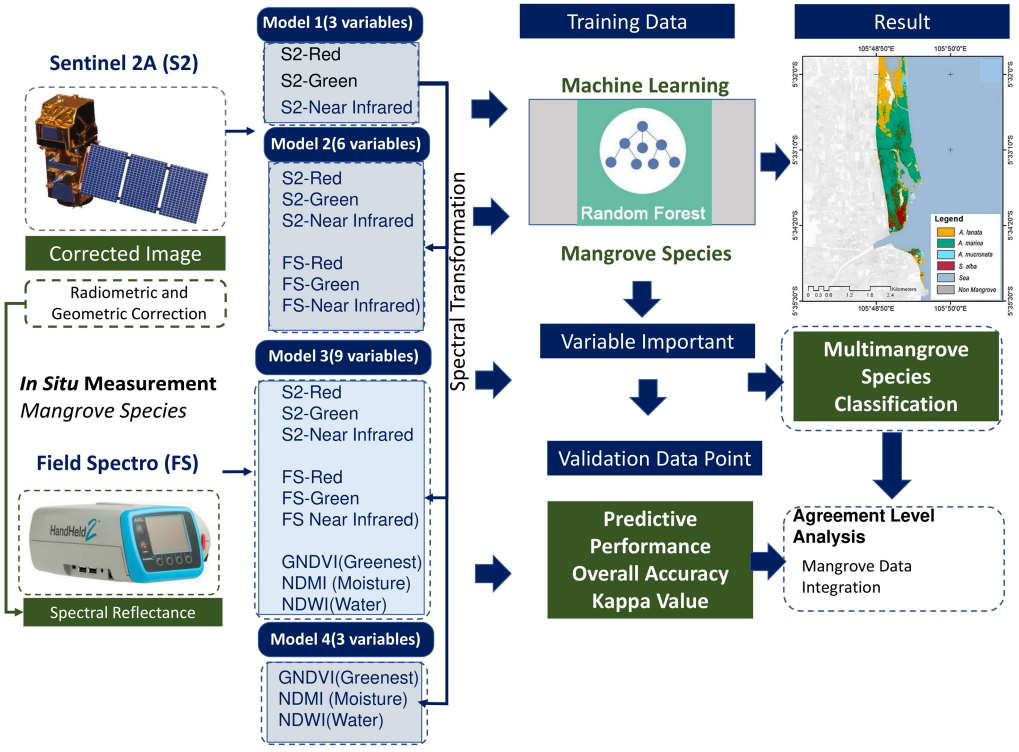


**Fig. 2** Flowchart of the methodology applied in this study.

correctness was performed to gauge the degree of precision of the usage map produced by the digital classification technique. Although the samples from the training area and accuracy test were different, the accuracy of the accuracy test sample was more commonly acknowledged because it was taken in a different area.[54]

### 2.3.1 *Image preprocessing*

The preprocessing analysis was divided into two main stages: radiometric and geometric corrections. In the radiometric correction analysis, the type of image used was Sentinel-2A MSI Multispectral Instrument, level 1C,[55] a product that has not been radiometrically and atmospherically corrected; therefore, pixel errors due to atmospheric influences must be minimized. The best image with <10% cloud cover was chosen for preprocessing. The selected recording period was between January 1 and December 31, 2022. The best data were the Sentinel-2A image recording on August 7, 2022, which had the minimum cloud cover such that the objects beneath were clearly visible. For each of the image's multispectral channels, radiometric calibration was carried out by translating digital values (DN) into radians.

The radian image was converted to top of atmosphere reflectance after radiometric calibration. The main objective was to correct for differences in reflectance values due to variations in the Earth–Sun distance on each recording date.[56] These differences can be significant owing to the differences in geographical conditions and the time of image recording. The fast line-of-sight atmospheric analysis of spectral hypercubes atmospheric correction method was used for the correction to reduce atmospheric impacts.[56] Sentinel-2A image processing was used to facilitate the analysis of the mangrove cover during the geometric correction stage. The geometric correction used in this study was image-to-map, with the reference data being an Indonesian Landform map based on ground control points (GCPs) collected directly in the field. The interpolation method used was a nearest-neighbor algorithm that only retrieves the nearest pixel value shifted to a new position. Six GCPs were used in the geometric correction, with a root-mean-square error value of 0.33 pixel.

### 2.3.2 *Mangrove reflectance*

The ASD HandHeld 2: hand-held visible NIR spectroradiometer (ASD Inc., Alpharetta, Georgia, United States) was used for mangrove reflectance measurements. The spectroradiometer was calibrated with a white reference prior to use[45] and recalibrated in cases where a significant difference in light intensity was noticed during use or when a "saturation alert" warning was issued.[15] A total of 144 reflectance values were obtained, and measurements were performed between 09.00 and 14.00 WIB to reduce the influence of weather at the research site. The angular position of the spectroradiometer sensor was set at 45 deg to the direction of sunlight to avoid shadows on the target object.[57] Measurements were made on land and partly using boats owing to difficult access to the sites.

The presence of mangroves above the water was sufficient to affect the spectral value due to the possibility of water reflection interfering with the mangrove reflectance value. The supported data output format was .txt, with data collection conducted from August 6 to 15, 2022. The Sentinel-2A image recording date was August 7, 2022. Table 2 lists a comparison of the wavelengths used in the Sentinel-2A image and the spectroradiometer. The results of the *in situ* measurements were used to develop a spectrum library[58] to compare the spectral image to the reference spectrum. The spectrum library was used as the reference to compare the Sentinel-2A image reflectance value to the field reflectance value.[59] The pixel values of mangrove species were used as the target spectrum for spectral matching-based object classification.

Table 2 lists the spectral width or resolution of the spectroradiometer used for the measurement of the reflectance of mangrove species. The spectral resolution of the spectroradiometer spans seven wavelength bands, ranging from 315 to 1100 nm. However, for mangrove measurements, the wavelengths used were green (525 to 605 nm), red (655 to 725 nm), and NIR (725 to 750 nm) and adjusted to correspond with the pre-existing wavelengths in the Sentinel-2A image.

**Table 2** Wavelength bands used by Sentinel-2A and the spectroradiometer.

| Sentinel-2A | | Spectroradiometer | |
|---|---|---|---|
| Wavelength (nm) | Band name | Wavelength (nm) | Band name |
| 433 to 453 | Coastal aerosol | 315 to 400 | Violet |
| 458 to 523 | Blue | 400 to 525 | Blue |
| 543 to 578 | Green | 525 to 605 | Green |
| 650 to 680 | Red | 605 to 655 | Yellow |
| 698 to 713 | Red edge 1 | 655 to 725 | Red |
| 733 to 748 | Red edge 2 | 725 to 750 | NIR |
| 773 to 793 | Red edge | 750 to 1100 | SWIR |
| 785 to 899 | NIR | | |
| 855 to 875 | NIR narrow | | |
| 1565 to 1655 | SWIR | | |

### 2.3.3 *Spectral transformation development*

Three index transformations were used in this study: GNDVI, NDWI, and NDMI, which are strongly associated with mangrove species characteristics and represent different groups of parameter types. GNDVI was created to assess the level of the vegetation's greenness using values generated by digital signal processing of brightness data from many satellite sensor data channels.[60] Conversely, in densely vegetated areas with good environmental conditions, the ratio of the two channels is exceptionally high (maximum) in the NIR band, whereas vegetation reflectance drops in the red band. The water spectral reflectance pattern decreases in the infrared (IR) and red light bands.[61]

The index provides a number between –1 and 1, representing vegetation cover density. In general, an index close to 1 indicates dense vegetation, whereas that below 0 indicates water and clouds.[62,63] The algorithm in remote sensing applications measures the greenness of vegetation using NIR and red wavelengths. The equation for GNDVI is as follows:[64]

$$\mathrm{GNDVI} = \frac{(\mathrm{NIR} - \mathrm{green})}{(\mathrm{NIR} + \mathrm{green})}. \tag{1}$$

Vegetation density was used to determine parameters for mapping mangrove species. The results of the transformation of the vegetation index with GNDVI are used to classify density into low, medium, and high. According to GNDVI analysis, the value range for each class was low (−0.91 to 0), moderate (0.01 to 0.45), and high (0.46 to 0.95), respectively. The visual representation of the vegetation density maps based on several indices is shown below. Vegetation index classes are determined based on the range of GNDVI values; the absence of vegetation is classified as low, moderate to dense coverage is labeled as medium, and vegetation with a high-density coverage is classified as high.[65]

For mapping water bodies, NDWI was the most suitable index. In the visible-to-IR spectrum, water has high absorption and low reflectance[66] Based on this phenomenon, this index takes advantage of the green and NIR colors in remote sensing photos. Due to its sensitivity to built-up terrain and tendency to overestimate water, NDWI can effectively improve water information.[67] The following equation was used to determine NDWI:[68]

$$\mathrm{NDWI} = \frac{(\mathrm{green} - \mathrm{NIR})}{(\mathrm{green} + \mathrm{NIR})}. \tag{2}$$

NDMI, computed as the ratio of the difference in the amounts of refracted radiation in the NIR and SWIR zones, describes the water level of the crop. The interpretation of the NDMI's

**Table 3** Variables combinations used in each model.

| Model | Number of variables |
|---|---|
| Model 1 (red, green, and NIR) | 3 |
| Model 2 [red, NIR, green (Sentinel-2A), red, NIR, green (field spectral data)] | 6 |
| Model 3 [red, NIR, green (Sentinel-2A), red, NIR, green (field spectral data), GNDVI, NDWI, and NDMI)] | 9 |
| Model 4 (GNDVI, NDWI, and NDMI) | 4 |

Abbreviations: GNDVI, green normalized difference vegetation index, NDWI, normalized difference water index, NDMI, normalized difference moisture index.

absolute value enables the quick identification of agricultural or field areas experiencing water stress issues. The NDMI is also simple to comprehend; regardless of the crop, its value ranges from −1 to 1, with each number denoting a different agronomic conditions.[67,69] NDMI was determined using the following equation:[30]

$$\mathrm{NDMI} = \frac{(\mathrm{NIR} - \mathrm{SWIR1})}{(\mathrm{NIR} + \mathrm{SWIR2})}. \tag{3}$$

### 2.3.4 *Model development*

Before using the RF algorithm to classify species, parameters related to species identification were extracted to obtain classification results with reasonable accuracy.[51] Then parameter testing was accomplished by analyzing the correlation between the parameters. The RF algorithm model was based on several parameters categorized into four classification models (Table 3).

Model 1 used green, red, and NIR band parameters; model 2 used green, red, and NIR band parameters as well as additional field spectroradiometer measurements. During model preparation, reflectance value parameters in the field were determined by plotting the reflectance values of mangrove species in the field onto a map. The reflectance data obtained were used to prepare a spectrum library that was used to construct a classification model.[45] Spectral reflectance analysis of mangrove species can be effectively performed using the ASD HandHeld 2 device in the wavelength range of 350 to 940 nm. This wavelength range was chosen based on the characteristics of the spectrometer specifications used.

Spectra representing mangrove species collected in the spectral reference source were resampled to match the center wavelength of the Sentinel-2A image band. The developed model was named model 3 and was based on red, green, and NIR band parameters, field spectroradiometer measurements, GNDVI, NDWI, and NDMI, meanwhile, model 4 uses the parameters GNDVI, NDWI, and NDMI. Each node was divided by RF using a randomized selection of input features or predictive variables.

## 2.4 Mangrove Classification Using RF

Each node in an RF model is divided into a random subset of input characteristics or predictive variables. In addition, to increase tree diversity, for building trees from various training data, RF employs bagging or bootstrap aggregation.[37,70,71] RF requires the selection of attributes (samples) and pruning methods.[72,73] The RF algorithm creates multiple bootstrap samples by randomly sampling the training data with replacement. The equation for bootstrap sampling is as follows:

$$D' = D * \left(1 - \frac{1}{e}\right)^n, \tag{4}$$

where $D'$ is the bootstrap sample, $D$ is the original dataset, $n$ is the number of examples in the bootstrap sample, and $e$ is expressed as the base of the natural logarithm (~2.71828).

Typically, an RF algorithm assesses the quality of a split in a decision tree using impurity metrics like the Gini index or entropy.[74] The formula for the Gini index is as follows:

$$\text{Gini index} = 1 - \Sigma(p_i)^2, \tag{5}$$

where $p_i$ is the proportion of instances of class $i$ at a particular node.

The entropy equation is as follows:[37]

$$\text{entropy} = -\Sigma(p_i * \log_2 p_i), \tag{6}$$

where $p_i$ is the proportion of instances of class $i$ at a particular node.

The ultimate prediction in the RF method was created by voting together the predictions of many decision trees.[4] For classification tasks, the majority vote was considered the final prediction. In classification, to improve processing effectiveness, it is crucial to understand how each variable affects the outcomes.[75] Variable importance, or the permutation importance (PI) value, or the mean decrease accuracy value was used to determine the contribution of a variable to the classification outcome. The more important the variable is, the greater its permutation value will be. The importance of the variables increases with the accuracy drop:[38]

$$\text{variable importance} = \text{OOB permutation} - \text{OOBbas}. \tag{7}$$

Out-of-bag (OOB) permutation is a measurement of variable importance that is determined by permuting variable data values not used for tree building. Out-of-bag basic (OOBbas) represents a basic measurement of variable importance without permutation. Counting the instances of the variable in the decision tree group is a simplistic method of determining the variable's relevance. The importance of the variable increases with its influence. When determining the relative relevance of a variable, the regression coefficient's absolute value was employed; the higher the coefficient value is, the larger the contribution of the variable in question to the biomass estimate for each unit change in the variable.[38,76]

PI is an algorithm used to obtain feature importance information by permuting (reordering the data set) the features used in training the prediction model. The process involves training a prediction model, permutation of features in the data, and re-evaluating the model. If a feature does not contribute much to the performance of a model, then reordering the data will have no significant effect; conversely, a feature with a significant contribution will greatly affect the performance of the model if the data were reordered.

### 2.5 Training Sampling and Predictive Performance

The sampling method used in this research is purposive random sampling. The purposive random sampling method was tailored to certain criteria to make the selected sample more representative.[71,75,77] Purposive random sampling was used to collect samples based on the consideration of built-up development and obtain samples that are representative of the study area. The obtained sample points were used for model training and accuracy tests.[77] The sample selection was performed by considering local knowledge and field checks. The number of image pixels in the mangrove area ranged from 396,115 pixels at location 1 to 216,738 at location 2. The classification process requires two types of data: training and testing. Training data are a sample of the entire population in the field used to build a model. In contrast, testing data are representative points of objects in the field used to validate the correctness of the results. The following equation was used to determine how many training samples are needed:

$$n = \frac{N}{1 + N \times (e)^2}, \tag{8}$$

where $n$ is the number of samples, $N$ is the population, and $e$ is the margin of error (percentage allowance for the accuracy of sampling errors that are acceptable). Based on the aforementioned technique and the number of pixels in the study, there were 400 pixels in each of the training samples. The following equation was used to determine the number of test points in this study:[51]

$$N = \frac{Z^2(p)(q)}{E^2}, \tag{9}$$

where $p$ is defined as the expected percentage accuracy, $q = 100 - p$, $E$ is the acceptable error, and $Z = 2$, calculated from the normal deviation of 1.96 at the 95% confidence level, $N$ is the

number of training samples obtained using the above method, and the number of pixels in each sample in this study is 144.

Predictive performance measures the model's ability to produce predictive results in accordance with the actual data.[78] A good model has high predictive performance, i.e., the predictions made match the reality or original data. This test involves 70% training data and 30% testing data. Predictive performance of mangrove classification refers to the extent to which a classification model can make accurate predictions related to the category or type of mangrove based on the parameters used as input for modeling.[54,78,79]

## 3 Results

### 3.1 Spectral Library of Mangrove Species

The mangrove species found in this study area were *R. mucronata, S. alba, A. lanata*, and *A. marina*. Data were collected from 144 sample points spread across the East Coast of East and South Lampung. A spectroradiometer was used to sample several leaves of different mangrove plants (Fig. 3). The results of reflectance measurements showed that *A. lanata*, *R. mucronata*, and *A. marina* exhibited reflectance values ranging from 0.014 to 0.768, 0.002 to 0.493, and 0.002 to 0.758, respectively; the reflectance values for *S. alba* ranged from 0.006 to 0.833. *Rhizophora* species exhibited a lower reflectance value than *A. marina* at wavelengths of 325 to 675 nm; nevertheless, spectral values at wavelengths of 680 to 700 nm were similar. Furthermore, the spectral values increased at wavelengths of 725 to 1075 nm. As evidenced by the large variance of the curve compared to that of the visible spectrum, the measurement findings demonstrate that the NIR wavelengths are particularly sensitive to the measuring distance.[57]

The reflectance value can be influenced by several factors, one of which is the time of data collection due to the angle of the altitude of the Sun at different sample points. In addition, the sensor direction relative to the nadir and the characteristics and conditions of the object can also influence the reflectance value. *Rhizophora* species showed higher reflectance values at wavelengths of 325 to 475 nm, and the values increased in the wavelengths of 525 to 575 nm. Further, they decreased at wavelengths of 625 to 675 nm and increased at 725 to 1075 nm. The reflectance

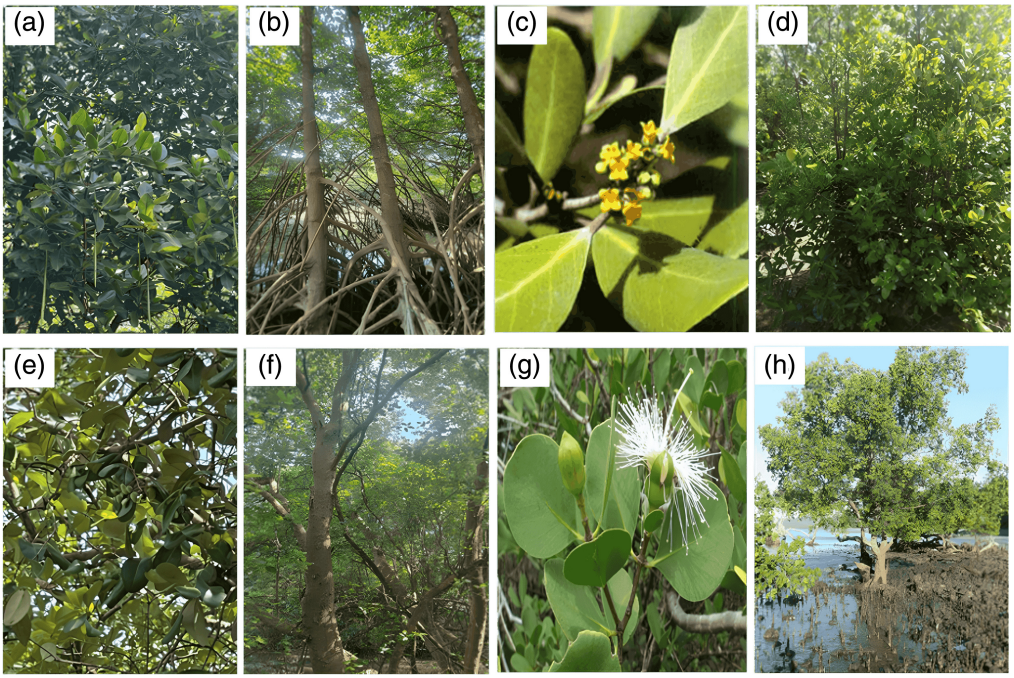


**Fig. 3** Shapes of leaves, roots, and flowers: (a) leaves and fruits of *R. mucronata* , (b) roots and stems of *R. mucronata*, (c) leaves and fruits of *A. lanata*, (d) roots and stems of *A. lanata*, (e) leaves and fruits of *A. marina*, (f) roots and stems of *A. marina*, (g) leaves and fruits of *S. alba*, and (h) roots and stems of *S. alba*.

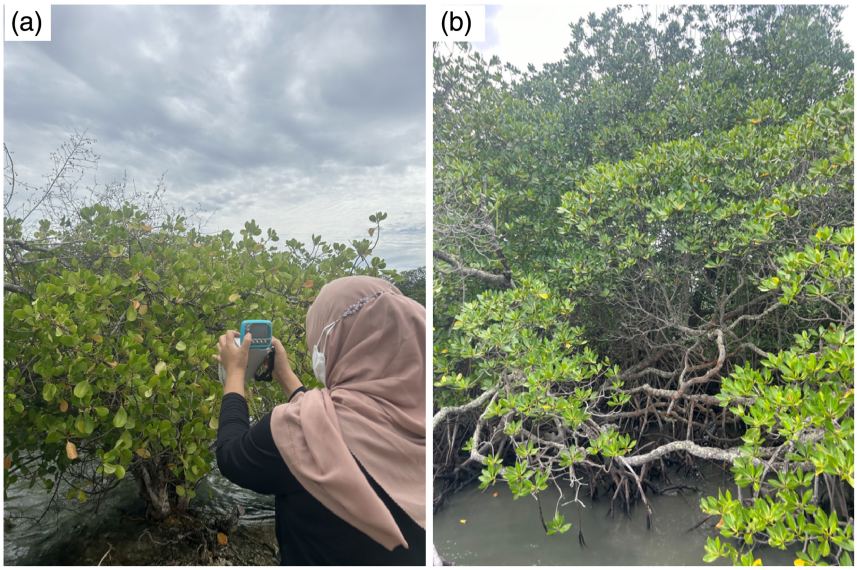


**Fig. 4** Field sampling of mangrove species: (a) spectral measurements using a field spectroradiometer and (b) types of Rhizopora mangrove species measured at the study site.

patterns of *Rhizophora* and *A. marina* were relatively similar, except that lower reflectance was observed in the wavelength range of 325 to 475 nm.

The pattern of these two species could be distinguished since the values for *R. mucronata* were higher than those for *A. marina*. The spectroradiometer used in the field had a sensitivity of only 1075 nm. In contrast, the Sentinel-2A image has a sensitivity of 2190 nm with the SWIR band, allowing it to be compared in the visible IR (green, red, and blue) and NIR bands with wavelengths measured with a field spectroradiometer. Figure 4 shows the spectral data collection process at one of the locations in the mangrove ecosystem.

Visible near-infrared spectroradiometer and Sentinel-2A field data processing and extraction results showed the same curve pattern, called a single-peak curve, albeit with different wavelengths. There were differences in the reflectance values of Sentinel-2A image processing from the time of measurement in the field. This is likely because the recording time is the main factor influencing the difference in species at the peak of the curve. A factor influencing the intensity of sunlight that the sensor can receive is the method of measuring samples in the field, such as the location of the measured leaf parts, field conditions, and density of each mangrove tree.[80] Figure 5 illustrates the box plots of the reflectance values of mangrove species.

The wavelength of light used differed for each mangrove species. At 443 to 665 nm, *A. marina* and *A. lanata* species had the highest median values; at 740 to 940 nm, *S. alba* species had the highest median wavelength and was followed by *A. marina* and *A. lanata* species. In general, *R. mucronata* species have the lowest median value among other species. The results from the box plot analysis showed differences between the species. The differences in the median, IQR, and presence of outliers indicate the variations in reflectance intensity and elucidate the potential unique characteristics of each species in the field spectroradiometer measurements.

Our findings illustrated the unique characteristic reflection patterns in the light spectrum of each mangrove species, thus allowing easy recognition and mapping of these species. This recognition and mapping can be accomplished using wavelength ranges adjacent to the NIR region, which facilitates the identification process based on the different light reflectance patterns produced by each mangrove species.[15]

According to Behera et al.,[38] the NIR and SWIR bands show considerable differences in reflectance intensity, with *A. officinalis* species showing higher reflectance than *H. fomes* and *E. agallocha*, despite these two species having identical spectral reflectance. In line with our findings, *A. marina* is known to exhibit high reflectance values in the NIR region at lower visible wavelengths. Meanwhile, *S. alba* tended to show higher values at all wavelengths.

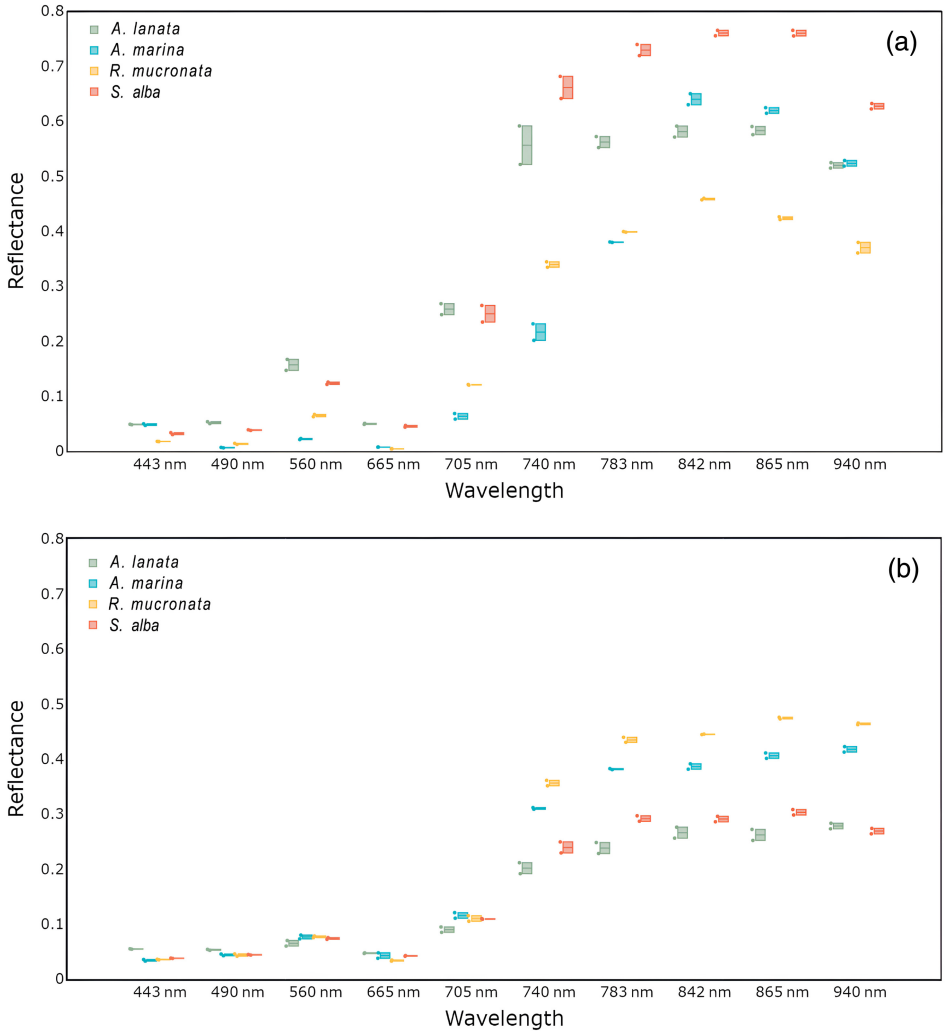


**Fig. 5** Box plots of reflectance values: (a) measurement results of reflectance values using a field spectroradiometer and (b) results of reflectance value analysis using Sentinel-2A imagery. The median in each box plot indicates average reflectance value. The bottom horizontal line of the box presents the first quartile (Q1), and the top horizontal line presents the third quartile (Q3).

### 3.2 Mangrove Species Classification

Species classification was tested using vegetation index parameters GNDVI, NDWI, and NDMI. Correlation with GNDVI, NDWI, and NDMI index values was measured using a field spectroradiometer. Comparing vegetation indices calculated from images with vegetation index data collected directly in the field can be used to evaluate the accuracy and relevance of images in representing actual vegetation conditions. The GNDVI parameter had a correlation value ($R^2$) of 0.71, indicating that the density parameter had an influence of 71%, whereas other factors contributed to the remaining 29%. The NDWI correlation value was 0.6, indicating that the density parameter had an influence of 60%, and other factors influenced the remaining 40%. The NDMI correlation coefficient was 0.66, indicating that the density parameter had an influence of 66%, and other factors influenced the remaining 34%.

The classification of mangrove species analyzed using the RF algorithm, resulted in four types of mangrove species, namely: *A. lanata, R. mucronata, A. marina*, and *S. alba.* The classification model consisted of four models with different numbers of parameters as mentioned in Table 3. Model 1 shows that in the Pasir Sakti area, *A. marina* and *R. mucronata* dominated; meanwhile, *S. alba* dominated in the Ketapang area. Results from the model 2 classification show that Pasir Sakti are mostly populated by *A. marina* and *R. mucronata* species, whereas *A. lanata* and *S. alba* species are widely distributed in Ketapang. Additionally, model 3 shows similar

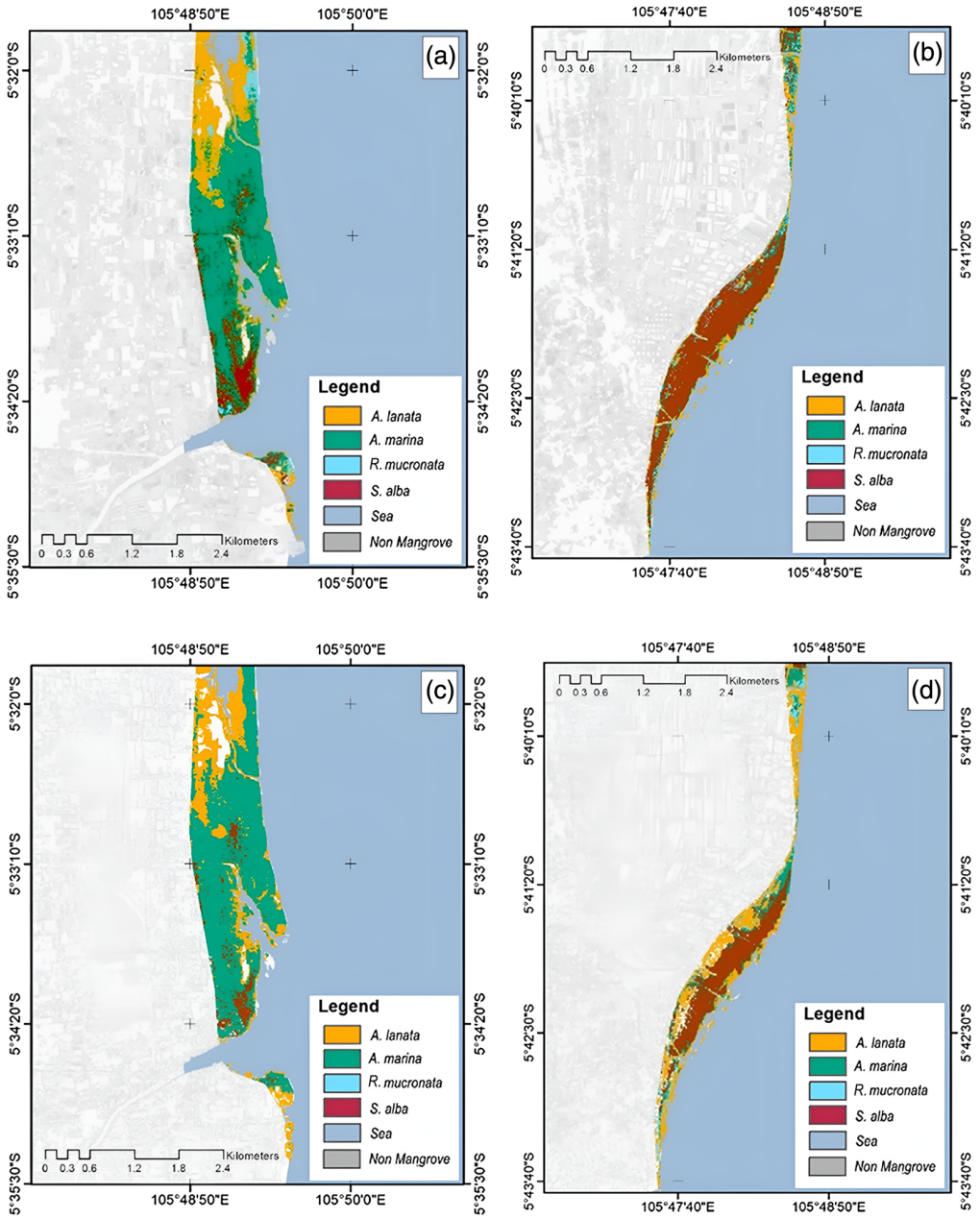


**Fig. 6** Classification of mangrove species using RF with model 1: (a) mangrove species map in Pasir Sakti, (b) mangrove species map Ketapang, (c) classification of mangrove species using RF with model 2 in Pasir Sakti, and (d) classification of mangrove species using RF with Model 2 in Ketapang.

results, where *A. marina* and *R. mucronata* dominated the Pasir Sakti area, whereas *A. lanata* and *S. alba* comprised most of the mangrove population in Ketapang. In contrast, for model 4, *A. marina* and *A. lanata* are dominant in Pasar Sakti, whereas *S. alba* and *A. lanata* dominate in Ketapang. The distribution of mangroves based on the classification model is shown in Figs. 6 and 7.

In the coastal area of South Lampung, most of the mangroves were dominated by *S. alba* and *A. lanata*, whereas in East Lampung, the dominant species were *A. marina* and *R. mucronata*. The types of mangroves found in the research location had characteristics that were clustered because there were planted mangroves. The RF algorithm can be used to distinguish between several mangrove species, such as *R. mucronata*, *S. alba*, *A. lanata*, *A. marina*, and non-mangroves. This classification contrasts species classes analyzed using spectral value inputs and

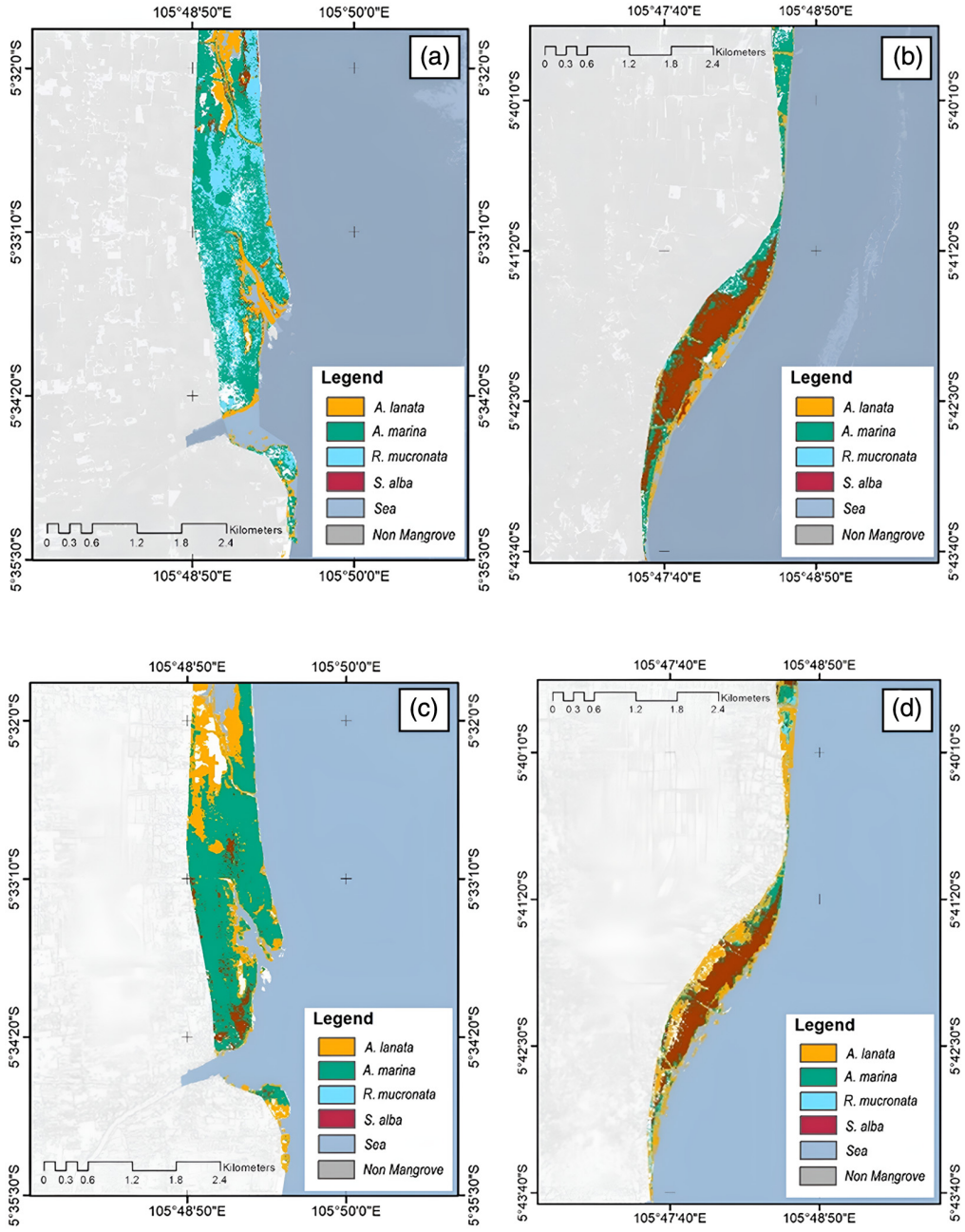


**Fig. 7** Classification of mangrove species using RF with model 3: (a) mangrove species map in Pasir Sakti, (b) mangrove species map Ketapang, (c) classification of mangrove species using RF with model 4 in Pasir Sakti, and (d) classification of mangrove species using RF with model 4 in Ketapang.

vegetation index/density values with those analyzed using spectral value inputs from field measurements and density. The classification results for each species revealed that each species had a clustering pattern within one area. This suggests the influence of the environment on the characteristics of the mangrove species. Environmental factors that affect mangrove growth in a location include coastal physiography (topography), tides (length, duration, and range), waves and currents, climate (light, rainfall, temperature, and wind), salinity, dissolved oxygen, soil, and nutrients.[81]

Based on the results of the classification of mangrove species, Table 4 lists the respective mangrove areas as per models 1 to 4. Models 1 to 3 covered 254.61, 331.75, 307.06 ha, and 361.42, respectively. *A. marina* had the highest abundance. *R. mucronata* covered the smallest

**Table 4** Estimation of mangrove areas (ha) and % by each model.

| | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
|---|---|---|---|---|---|---|---|---|
| Species | Area | Percentage | Area | Percentage | Area | Percentage | Area | Percentage |
| *R. Mucronata* | 35.82 | 5.40 | 129.31 | 18.44 | 109.72 | 16.69 | 27.45 | 5.03 |
| *S. Alba* | 211.00 | 31.82 | 104.02 | 14.84 | 17.83 | 2.71 | 147.8 | 27.09 |
| *A. Lanata* | 161.61 | 24.37 | 331.75 | 47.32 | 222.98 | 33.91 | 8.90 | 1.63 |
| *A. Marina* | 254.61 | 38.40 | 136.00 | 19.40 | 307.06 | 46.69 | 361.42 | 66.24 |

area in models 1 and 4, whereas *A. lanata* covered the smallest area in Model 4. For *R. mucronata*, the area covered in model 2 differed significantly from those in the other two models.

### 3.3 Predictive Performance of Classification

Prediction results indicate a significant variation in accuracy among the different models. Indeed, model 1 showed an accuracy rate of 70.69%; model 2 increased the accuracy to 76.81%; model 3 achieved an accuracy of 81.25%; model 4 had an accuracy rate of 79.17%. A notable change occurred from models 1 to 2, where the accuracy increased by ~6.12%. However, the difference of 2.08% in accuracy between models 3 and 4 was not significant. This change in accuracy suggests that the use of certain parameters in the model has a major impact on the prediction performance. Model 2 might have involved some adjustments or additional parameters that improved its accuracy compared with model 1. Similarly, model 3, while having a modest difference in accuracy from model 4, might have also involved more careful parameter configuration to achieve significant performance improvements.[78] The kappa value of models 1 to model 4 was 0.696, 0.778, 0.80, and 0.78, respectively (Fig. 8).

Sahani and Ghosh[82] indicated that the classification of the kappa value is based on the level of agreement: <0.05 (zero), 0.05 to 0.20 (very bad), 0.20 to 0.40 (bad), 0.40 to 0.55 (fair), 0.55 to 0.70 (good), 0.70 to 0.85 (very good), 0.85 to 0.99 (excellent), and 0.99 to 1 (perfect).[82] Thus the kappa value in this study can be classified in the "good to excellent" category, where the kappa value of all models ranged between 0.696 and 0.80.

The classification results show that the higher the parameters used for modeling are, the higher the accuracy and kappa values are. The correctly classified graph in Fig. 8 explains that *A. lanata* is referred to as ALA, *A. marina* is referred to as AMA, *R. mucronata* as RMU, and *S. alba* refers to SAL. Details regarding the graph comparing the accuracy and kappa values are shown in Fig. 8.

Model 1 produced an OA of 70.69% for 720 sample points. *R. mucronata* obtained UA 76.39% and PA 70.06%, *S. alba* obtained UA 68.06% and PA 66.22%, *A. lanata* obtained UA 69.44% and PA 76.92%, *A. marina* obtained UA 62.50% and PA 68.70%, and non-mangrove obtained UA 77.08 and PA 72.08%. Based on previous research, the maximum accuracy of 74% was obtained by combining AVIRIS-NG data, whereas another study explained that the OA was 69.44% for pixel-based classification and 82% for object-based classification using RF.[5,38] An accuracy test was conducted using the Sentinel-2A confusion matrix/error test (Table 5).

Model 2 exhibited an OA of 76.81% for 720 sample points. *R. mucronata* obtained UA values of 83.33% and PA of 96.77%, and *S. alba* obtained UA of 76.39% and PA of 76.92%, *A. lanata* obtained UA of 70.14% and PA of 64.33%, *A. marina* obtained UA of 74.31% and PA of 72.79%, and non-mangrove obtained UA of 79.86% and PA of 77.17%. Details on the accuracy test of model 2 are presented in Table 6.

Model 3 exhibited an OA of 81.25% for 720 sample points. *R. mucronata* obtained UA values of 84.03% and PA of 90.98%, *S. alba* obtained UA of 79.86% and PA of 83.94%, *A. lanata* obtained UA of 86.11% and PA 75.61%, *A. marina* obtained UA of 76.39% and PA of 73.33%, and non-mangrove obtained UA of 79.86% and PA of 84.56%. Details on the accuracy test of model 3 are presented in Table 7.

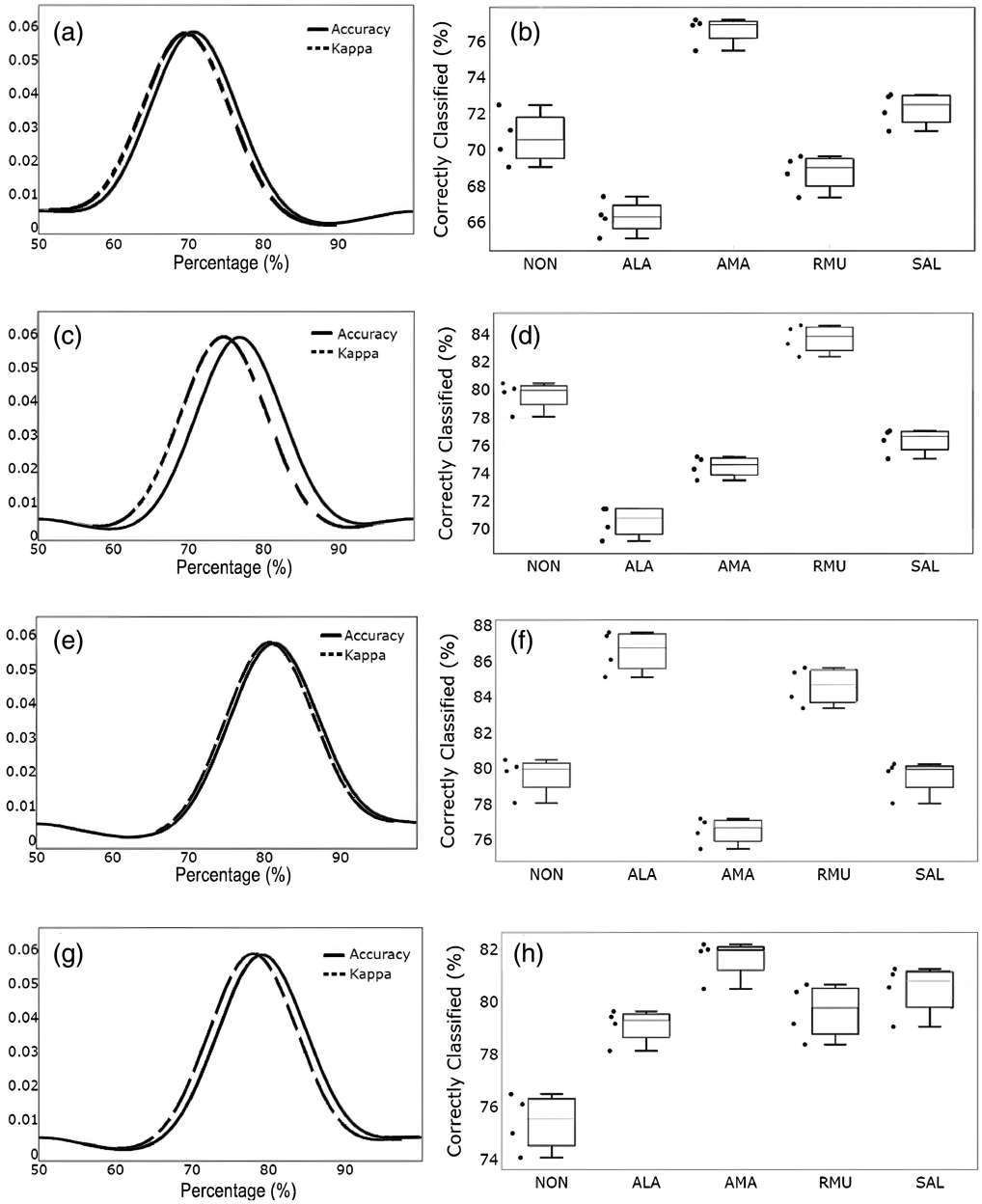


**Fig. 8** Predictive performance for the training: (a) comparison of kappa value and accuracy of model 1, (b) datasets using model 1, (c) kappa value and accuracy of model 2, (d) datasets using model 2, (e) kappa value and accuracy of model 3, (f) datasets using model 3, (g) kappa value and accuracy of model 4, and (h) datasets using model 4.

Model 4 showed an OA of 79.17% for 720 sample points. *R. mucronata* showed a UA value of 79.17% and PA of 89.06%, *S. alba* showed a UA of 80.56% and PA of 76.32%, *A. lanata* showed a UA of 79.17% and PA of 73.55%, *A. marina* showed a UA of 81.94% and PA of 72.84%, and the non-mangrove group showed a UA of 75.00% and PA of 87.80% (Table 8).

Model 3 had the highest accuracy of the four models with the most parameters. Previous studies have attempted to improve mapping accuracy, including mapping based on machine learning classification.[83] It is challenging to map mangrove species using remote sensing data from spectral reflectance patterns measured directly in the field.[57] In mangrove habitats, several water and soil quality parameters of mangrove species can be described using remote sensing data. The study also discovered that variations in mangrove reflectance at the canopy level are determined by the amount of chlorophyll in the species, environmental conditions at the time of

**Table 5** Classification accuracy of model 1.

| | Field data | | | | | | |
|---|---|---|---|---|---|---|---|
| Image data | *R. mucronata* | *S. alba* | *A. lanata* | *A. marina* | Non-mangrove | Total | UA (%) |
| *R. mucronata* | 110 | 25 | 5 | 4 | 0 | 144 | 76.39 |
| *S. alba* | 20 | 98 | 0 | 15 | 11 | 144 | 68.06 |
| *A. lanata* | 0 | 25 | 100 | 5 | 14 | 144 | 69.44 |
| *A. marina* | 21 | 0 | 15 | 90 | 18 | 144 | 62.50 |
| Non-mangrove | 6 | 0 | 10 | 17 | 111 | 144 | 77.08 |
| Total | 157 | 148 | 130 | 131 | 154 | 720 | |
| PA (%) | 70.06 | 66.22 | 76.92 | 68.70 | 72.08 | 353.98 | |
| OA (%) | 70.69 | | | | | | |

**Table 6** Classification accuracy of model 2.

| | Field data | | | | | | |
|---|---|---|---|---|---|---|---|
| Image data | *R. mucronata* | *S. alba* | *A. lanata* | *A. marina* | Non-mangrove | Total | UA (%) |
| *R. mucronata* | 120 | 6 | 6 | 12 | 0 | 144 | 83.33 |
| *S. alba* | 0 | 110 | 15 | 10 | 9 | 144 | 76.39 |
| *A. lanata* | 0 | 20 | 101 | 8 | 15 | 144 | 70.14 |
| *A. marina* | 4 | 3 | 20 | 107 | 10 | 144 | 74.31 |
| Non-mangrove | 0 | 4 | 15 | 10 | 115 | 144 | 79.86 |
| Total | 124 | 143 | 157 | 147 | 149 | 720 | |
| PA (%) | 96.77 | 76.92 | 64.33 | 72.79 | 77.18 | 388.00 | |
| OA (%) | 76.81 | | | | | | |

**Table 7** Classification accuracy of model 3.

| | Field data | | | | | | |
|---|---|---|---|---|---|---|---|
| Image data | *R. mucronata* | *S. alba* | *A. lanata* | *A. marina* | Non-mangrove | Total | UA (%) |
| *R. mucronata* | 121 | 0 | 14 | 9 | 0 | 144 | 84.03 |
| *S. alba* | 2 | 115 | 15 | 12 | 0 | 144 | 79.86 |
| *A. lanata* | 3 | 7 | 124 | 9 | 1 | 144 | 8611 |
| *A. marina* | 5 | 0 | 9 | 110 | 20 | 144 | 76.39 |
| Non-mangrove | 2 | 15 | 2 | 10 | 115 | 144 | 79.86 |
| Total | 133 | 137 | 164 | 150 | 136 | 720 | |
| PA (%) | 90.98 | 83.94 | 75.61 | 73.33 | 84.56 | 408.42 | |
| OA (%) | 81.25 | | | | | | |

**Table 8** Classification accuracy of model 4.

| | Field data | | | | | | |
|---|---|---|---|---|---|---|---|
| Image data | *R. mucronata* | *S. alba* | *A. lanata* | *A. marina* | Non-mangrove | Total | UA (%) |
| *R. mucronata* | 114 | 1 | 13 | 15 | 1 | 144 | 79.17 |
| *S. alba* | 4 | 116 | 12 | 10 | 2 | 144 | 80.56 |
| *A. lanata* | 5 | 12 | 114 | 11 | 2 | 144 | 79.17 |
| *A. marina* | 2 | 4 | 10 | 118 | 10 | 144 | 81.94 |
| Non-mangrove | 3 | 19 | 6 | 8 | 108 | 144 | 75.00 |
| Total | 128 | 152 | 155 | 162 | 123 | 720 | |
| PA (%) | 89.06 | 76.32 | 73.55 | 72.84 | 87.80 | 399.57 | |
| OA (%) | 79.17 | | | | | | |

measurement, and the background reflectance of soil and water.[46] When measuring mangrove spectral reflectance in the field, variations can be caused by lighting circumstances, canopy structure, leaf orientation, measurement distance, and background objects.[57]

Prior research related to mangrove classification obtained mangrove species from *Avicennia*, *Rhizophora*, and *Sonneratia* using *K*-means and decision tree methods.[47] The results for the mangrove class showed considerable PA and UA averages of roughly 94.4% and 94.5%, respectively, demonstrating the usefulness of the adopted strategy for precise mangrove delineation. According to Rahmawati et al.,[64] in the classification using RF algorithm, various vegetation indices were used, including enhanced vegetation index, normalized difference vegetation index (NDVI), soil-adjusted vegetation index, GNDVI, modified normalized difference water index, NDWI, index-based built-up index, and land surface water index. The mangrove area obtained was 424.48 ha with an OA of 58.45% and a kappa value of 39.59.

### 3.4 Agreement Level Analysis

Agreement level analysis was conducted by comparing and observing the similarity of pixels by overlaying datasets based on cross-walking between classes.[84] In line with the results of previous research, in this study, the classification results were calculated to determine the percentage of species similarities found in models 1 to 4. The RF algorithms were classified into four models based on the parameters used. Model 1 used red, green, and NIR band parameters; model 2 used red, green, and NIR band parameters and additional field spectroradiometer measurements; model 3 used red, green, and NIR band parameters, field spectroradiometer measurements, GNDVI, NDWI, and NDMI; and model 4 only used GNDVI, NDWI, and NDMI depicts the findings of the agreement level analysis. The mangrove area was calculated based on the level of species similarity generated by the four models using the agreement level analysis. The classification results were analyzed based on the similarity of species generated from these four models, including *A. marina*, *A. lanata*, *R. mucronata*, and *S. alba* (Figs. 9 and 10). The aim of this analysis was to compare the area of each species based on the similarity of the classification results (Table 9).

The results of the agreement level analysis show that each species had the highest percentage in model 3. Meanwhile, model 2 showed the lowest representation for *R. mucronata* (3%) and *A. lanata* (9%), whereas model 3 had the highest percentages of these species (72% and 41%, respectively). *S. alba* and *A. marina* have the lowest percentages in model 4 (22% and 20%, respectively), whereas their highest percentage was observed in model 3 (29%). However, unlike *S. alba*, *A. marina* had the highest percentage in model 1 (at 32%).

Table 10 lists the analysis results, where the preparation of this agreement level was based on the similarity of species generated by each model. Agreement level 1 indicates that *A. lanata, A. marina, R. mucronata*, and *S. alba* were found in models 1 to 4. Agreement level 2 is when the same species was found in two out of the four models, whereas agreement level 3 shows that the

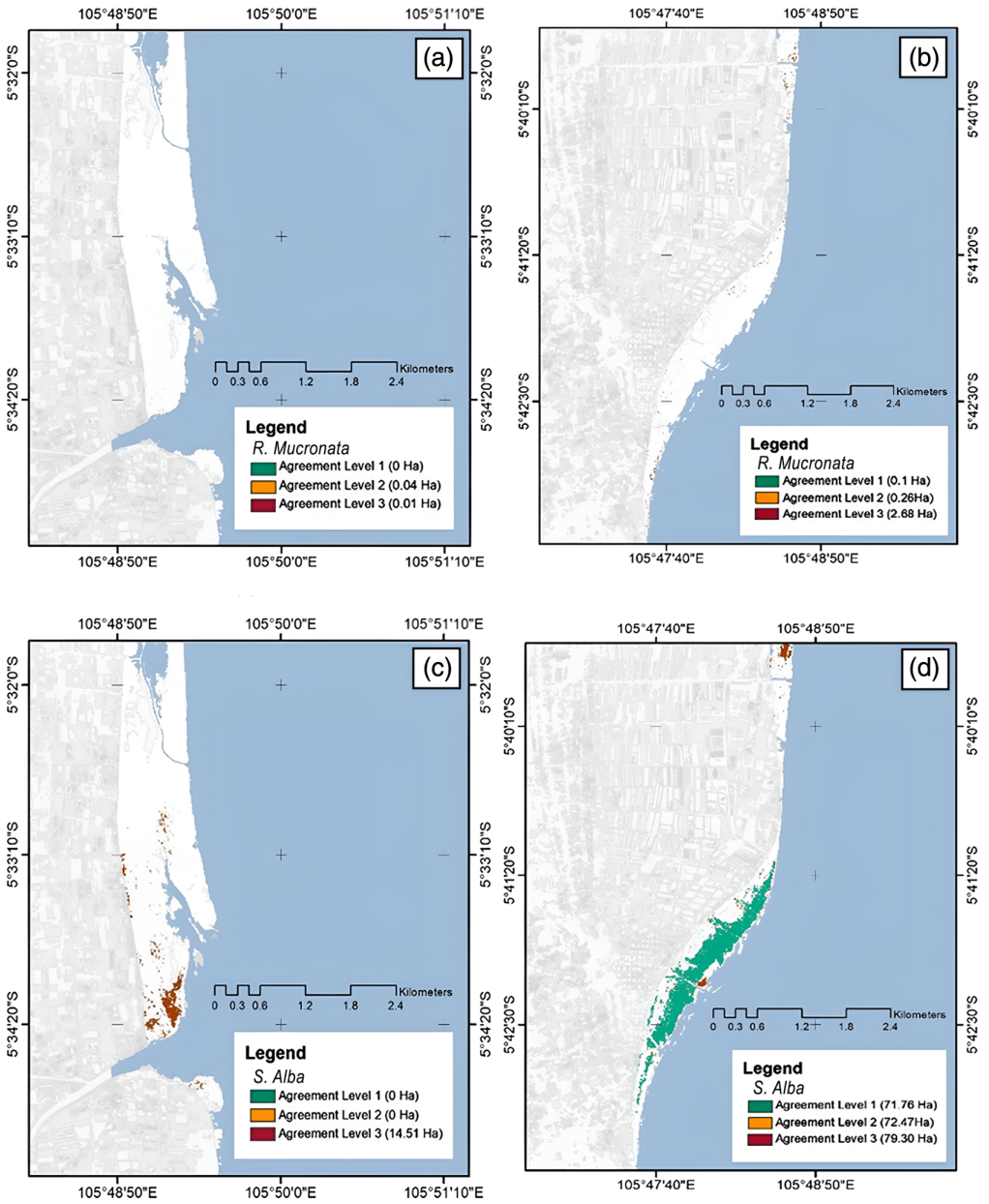


**Fig. 9** Confident level for mangrove species classification: agreement level of (a) *A. lanata* in Pasir Sakti, (b) of *A. lanata* in Ketapang, (c) *A. marina* in Pasir Sakti, and (d) *A. marina* in Ketapang.

species was only found in one model. *R. mucronata* and *S. alba* were not found at this location, but only *S. alba* did not have an area in agreement level 2. The results of agreement level 1 showed that *A. marina* at site 1 had the largest area compared with that of other species at 115.46 ha, with agreement level 2 at 134.37 ha and agreement level 3 at 201.11 ha. This analysis showed that site 1 was dominated by *A. marina*, the second site was mostly dominated by *R. mucronata*. The respective distributions at levels 1, 2, and 3 were 71.76, 72.47, and 79.30 ha, respectively.

## 4 Discussion

By combining the characteristics obtained from Sentinel-2A photographs with the findings of field reflectance measurements made using a field spectroradiometer, this study helped to identify the types of mangrove species. We assessed the potential of the RF algorithm for

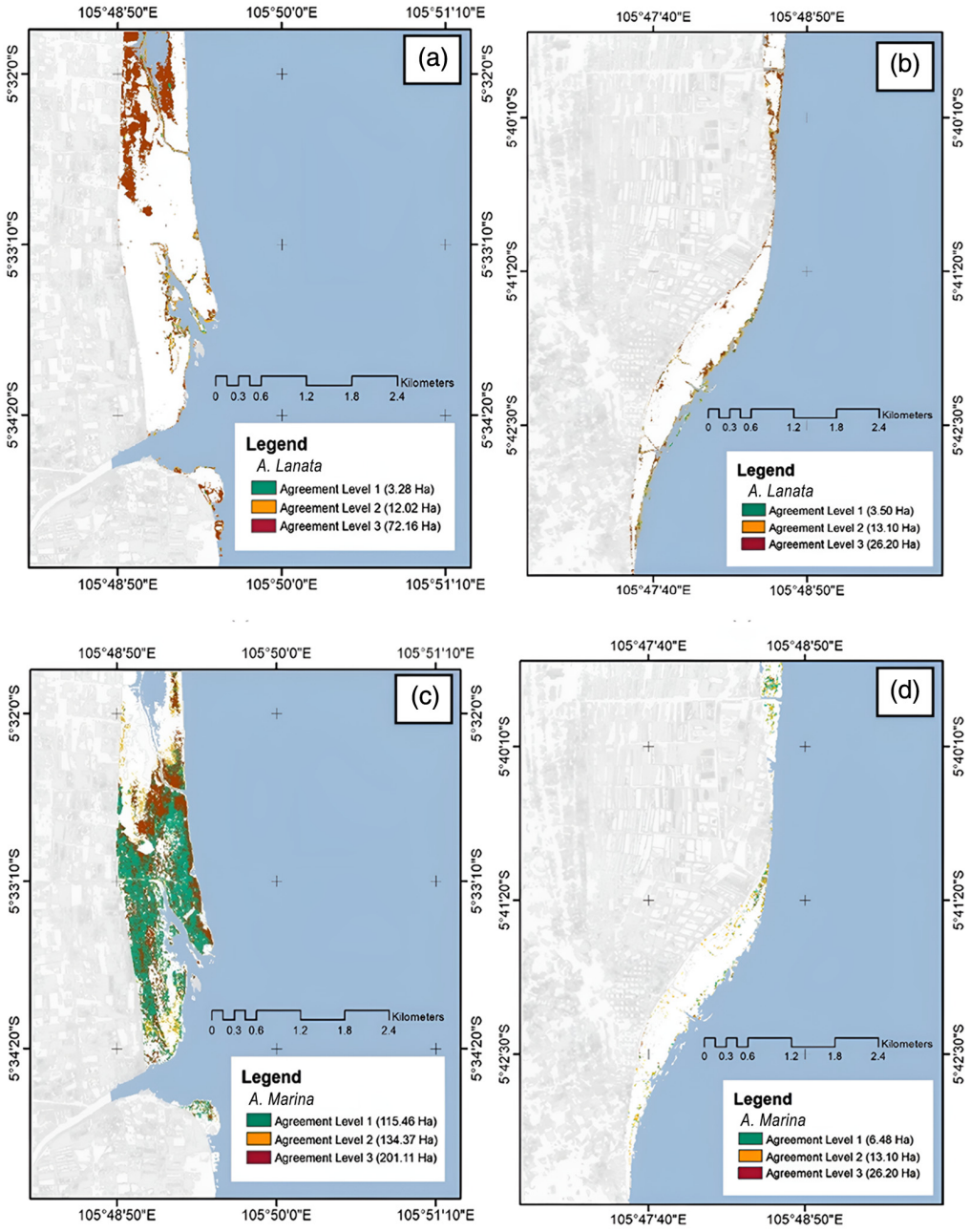


**Fig. 10** Confident level for mangrove species classification: agreement level of (a) *R. mucronata* in Pasir Sakti, (b) *R. mucronata* in Ketapang, (c) *S. alba* in Pasir Sakti, and (d) *S. alba* in Ketapang.

**Table 9** Species similarity based on agreement level analysis of each model.

| Species | Model 1 (%) | Model 2 (%) | Model 3 (%) | Model 4 (%) |
|---|---|---|---|---|
| *R. mucronata* | 15 | 3 | 72 | 10 |
| *S. alba* | 23 | 26 | 29 | 22 |
| *A. marina* | 32 | 23 | 25 | 20 |
| *A. lanata* | 25 | 9 | 41 | 25 |

**Table 10** Mangrove species area (ha) based on agreement level analysis.

| | Agreement level 1 | | Agreement level 2 | | Agreement level 3 | |
|---|---|---|---|---|---|---|
| Mangrove species | Location 1 | Location 2 | Location 1 | Location 2 | Location 1 | Location 2 |
| *A. lanata* | 3.28 | 3.51 | 12.02 | 13.10 | 72.16 | 26.20 |
| *A. marina* | 115.46 | 6.48 | 134.37 | 15.66 | 201.11 | 8.43 |
| *R. mucronata* | 0 | 71.76 | 0.04 | 72.47 | 0.01 | 79.30 |
| *S. alba* | 0 | 0.1 | 0 | 0.26 | 14.51 | 2.66 |

classification of mangrove species using the best model to improve accuracy. The following mangrove species were identified: *A. lanata*, *A. marina*, *R. mucronata*, and *S. alba.*

### 4.1 Spectral Signature Clustering of Mangrove Species

Spectral clustering of mangrove species signatures was accomplished by analyzing the spectral response patterns produced by mangrove plants at various wavelengths, especially in the spectral range that includes visible and NIR wavelengths. To determine the spectral separation between the various mangrove species, a canonic discriminant analysis was conducted within a reflectance range of 443 to 940 nm. Our findings demonstrated the association between the discriminant value and the group that had a correlation value of 0.896; since this value is very close to 1, a positive relationship is observed (the correlation magnitude is between 0 and 1)[15] (Table 11).

The Wilk's lambda value of 0.368 indicated a significant difference between at least one group or category among the wavelength and reflectance values. This is reinforced by the chi-square value of 23.23, which indicates that the difference was extremely significant, especially when compared with the general significance level (0.05).[15] The significance result of 0.000 indicates that the observed difference is statistically significant since this value is much lower than the preset significance level. Canonical correlation analysis revealed that different wavelengths have a differentiating impact on each other. As such, these findings provide strong evidence of the relationship between wavelength variables and reflectance values, with differences being identifiable through the canonical correlation analysis conducted.

### 4.2 Correlation of Sentinel-2A Image Reflectance and Spectral Field Measurements

Green, red, and NIR wavelengths in Sentinel-2A imagery and field measurements were correlated to determine how well the data varied from the results of both measurements. The statistical analysis results showed that in the green, red, and NIR bands, the $R^2$ value *s* were 0.69, 0.73, and 0.82, respectively (Fig. 11); the $R^2$ value of 0.69 in the green band indicated that ~69% of the variation in the field measurements could be explained by the reflectance value at the green wavelength in the Sentinel-2A imagery. This shows a positive relationship between the field and image data; however, this model could not explain ~31% of the variation. Similarly, the $R^2$ value of 0.73 in the red band indicated that ~73% of the variation in the field measurements could be explained by the reflectance values at the red wavelength in the Sentinel-2A imagery and showed a good relationship between the field and image data (slightly better than in the green band).[62] In the NIR band, the $R^2$ value (0.82) was the highest among the three spectral bands, suggesting that ~82% of the variation in the field measurements could be explained by the reflectance values at NIR wavelengths in the Sentinel-2A imagery. This relationship was stronger and

**Table 11** Canonical discriminant eigenvalue results.

| Function | Eigenvalue | % of variance | Cumulative % | Canonical correlation |
|---|---|---|---|---|
| 1 | 4.056a | 100 | 100 | 0.896 |

[a]First 1 canonical discriminant functions were used in the analysis.

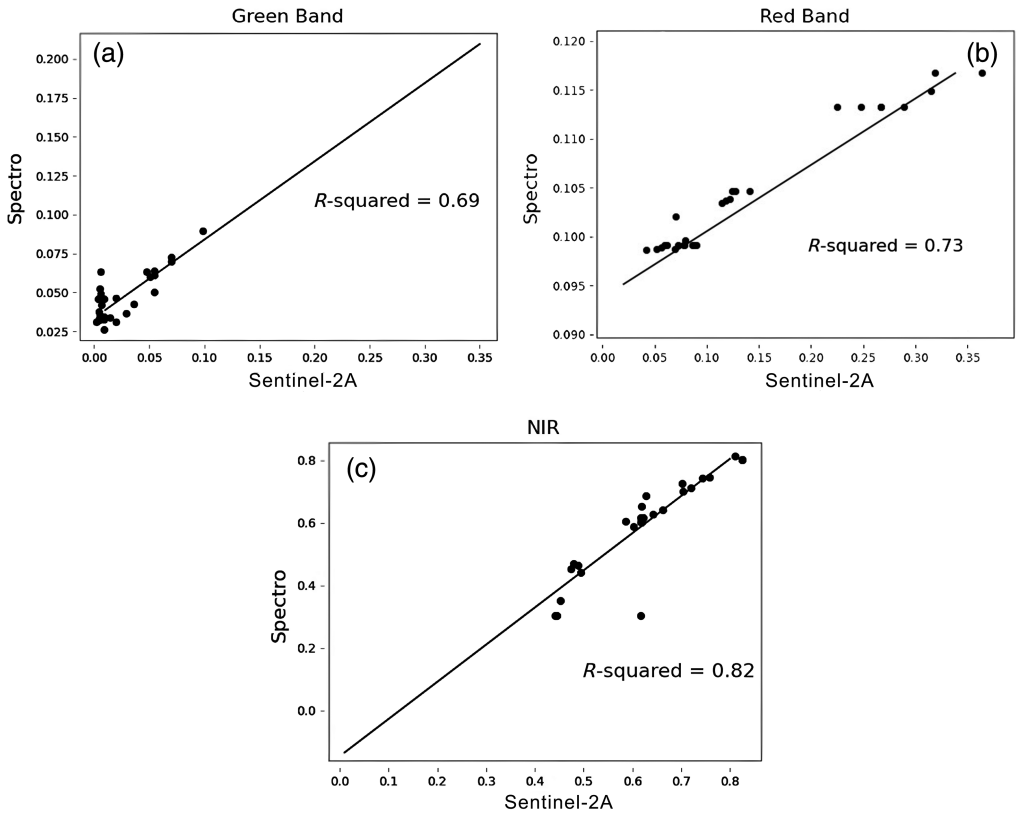


**Fig. 11** Correlation between Sentinel-2A image reflectance values and field spectroradiometer: *R* square value of the (a) green, (b) red, and (c) NIR band reflectance of the field spectroradiometer and Sentinel-2A.

more consistent than for the previous two spectral bands. The correlation between the Sentinel-2A wavelength and field spectroradiometer is illustrated in Fig. 11.

### 4.3 Variable Importance Analysis

The best parameter analysis uses variable importance to identify important variables as inputs for further biomass estimation.[85] Analysis of the most important variables was performed on all classification models. Details of the permutation graph are illustrated in Fig. 11.

All selected variables were tested and evaluated for their contribution to the classification model. Based on variable importance analysis, the highest PI value for model 1 was 0.102 (red band), whereas those for models 2 and 3, were 0.338 and 0.283, respectively (NIR band). In model 4, GNDVI had the highest importance score, 0.14972, indicating that GNDVI had the most significant impact on decision-making. NDMI had an importance score of 0.121066, indicating that this variable also had a considerable contribution to the analysis. Meanwhile, NDWI had a lower importance score of 0.054272, indicating that in this context, NDWI had a more limited impact compared with the other two variables. More details of the PI values are presented in Table 12 (Fig. 12).

Similar to findings of related mangrove-related research,[85,86] the vegetation index generated from mid-IR and NIR, and texture, which is derived from the red band, are the two most crucial factors.[87]

### 4.4 Best Model for Mangrove Species Classification

The best parameter analysis uses variable importance to identify important variables as input for mangrove species classification. Model 3 showed the highest accuracy compared with the other two models that used only parameters from Sentinel bands and ground reflectance measurements. The OA of model 3 was 81.25%, the UA was 81.68%, and the PA was 81.25% (Table 13). These findings suggest that adding parameters to the classification can improve the accuracy of mangrove species mapping.

**Table 12** PI values obtained in each model.

| Model 1 | | Model 2 | | Model 3 | | Model 4 | |
|---|---|---|---|---|---|---|---|
| Band | PI | Band | PI | Band | PI | | |
| Green (Sentinel-2A) | 0.005 | Green (Spectro) | 0.033 | Green (Spectro) | 0.007 | NDWI | 0.054 |
| NIR (Sentinel-2A) | 0.060 | Green (Sentinel-2A) | 0.042 | NIR (Spectro) | 0.007 | NDMI | 0.121 |
| Red (Sentinel-2A) | 0.102 | NIR (Spectro) | 0.059 | NDWI | 0.009 | GNDVI | 0.149 |
| | | Red (Spectro) | 0.113 | Green (Sentinel-2A) | 0.020 | | |
| | | Red (Sentinel-2A) | 0.187 | NDMI | 0.020 | | |
| | | NIR (Sentinel-2A) | 0.338 | GNDVI | 0.036 | | |
| | | | | Red (Spectro) | 0.075 | | |
| | | | | Red (Sentinel-2A) | 0.137 | | |
| | | | | NIR (Sentinel-2A) | 0.283 | | |

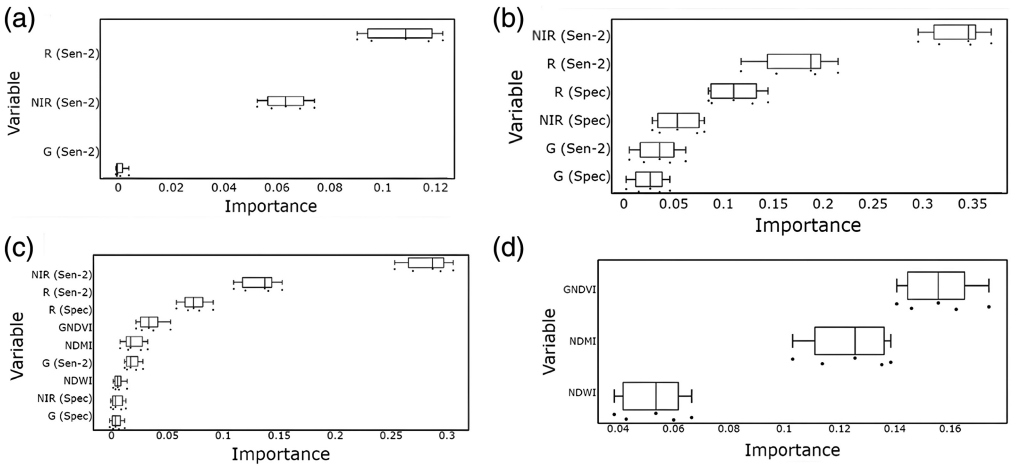


**Fig. 12** Variable importance plots are shown for (a) model 1, (b) model 2, (c) model 3, and (d) model 4.

**Table 13** Classification accuracy results of each model.

| Model | OA (%) | UA (%) | PA (%) |
|---|---|---|---|
| Model 1 (red, green, and NIR) | 70.69 | 70.69 | 69.26 |
| Model 2 (red, NIR, green, and field spectral data) | 76.81 | 77.59 | 76.80 |
| Model 3 (red, NIR, green, field spectral data, GNDVI, NDWI, and NDMI) | 81.25 | 81.68 | 81.25 |
| Model 4 (GNDVI, NDWI, and NDMI) | 79.17 | 79.17 | 79.91 |

Abbreviations: GNDVI, green normalized difference vegetation index, NDWI, normalized difference water index, NDMI, normalized difference moisture index.

Although mangrove mapping can be done well with RF algorithms, errors often occur when classifying objects. This is common in the classification of other vegetation, bare land, or water bodies.[88] The decrease in accuracy can be caused by several factors, such as phenological similarities, incorrect input parameters in modeling, and the level of heterogeneity of the objects being described.[78,89]

### 4.5 Spatial Distribution of Mangrove Species

Based on the findings of the image analysis and field reflection measurements, the mangrove species identified in South Lampung Beach were *R. mucronata*, *S. alba*, *A. lanata*, and *A. marina*. *R. mucronata* species in the best model (model 3) has an area of 109.72 ha with a percentage of 16.69%; *S. alba* has an area of 17.83 ha with a percentage of 2.71%; *A. lanata* has an area of 222.98 ha with a percentage of 33.91%; and *A. marina* has an area of 307.06 ha with a percentage of 46.69%. The total area of mangroves is 657.59 ha. Classification results show that it has the highest area, whereas *S. alba* species has the lowest area.

An RF algorithm with vegetation index parameters and spectral reflectance was used to map mangrove species. GNDVI has been used as a parameter to represent the phenology of mangrove species.[30,39,90] The RF was used to demonstrate the feasibility of mangrove species classification. The reflectance patterns of *R. mucronata*, *S. alba, A. lanata*, and *A. marina* were similar at a wavelength of 325 to 475 nm. An increase in reflectance occurred at wavelengths of 525 to 575 nm, whereas a significant increase was observed at wavelengths of 720 to 1075 nm. The existence of irregular patterns in object reflections can be caused by several factors, such as disturbances that occur owing to variable cloud cover, fluctuations in light sources, and weather conditions during sampling in the field. This interference can be referred to as "noise" and cannot be used for analyzing the spectral characteristics of vegetation reflectance; therefore, the data from the noise are ignored.[80]

The accuracy test results showed an OA of 84.51% with 71 field-observation sample points. Mangrove management, conservation, and restoration depend on accurate mapping of the quality, distribution, and number of species.[91] In this study, we combined recursive feature elimination and deep learning methods with ensemble RF, XGBoost, LightGBM, CatBoost, and AdaBoost Mcnemar tests. Our findings revealed a significant difference in the classification of mangrove species. SVM classification produced better accuracy than decision tree classification as it can minimize errors in image interpretation, with OA values reaching 95% (kappa = 0.86) and 93% (kappa = 0.82).[26] We observed better accuracy with the RF model when used with Sentinel-2 in distinguishing the three dominant species.[38] Using Worldview imagery to classify mangrove species, it was discovered that the RF technique was more accurate and effective than SVM with an OA of 95.89% and a kappa coefficient of 0.95. As there are uncertain data related to the distribution and extent of mangroves, particularly in Asia, conducting assessments, and modeling related to mangrove ecosystem services is required.[44]

This mapping is scenario-based, resulting in a well-developed quantification.[92] Extraction of mangrove information for high accuracy, via optimizing the images used by integrating the results of spectroradiometer measurements and medium resolution, remains essential.[93] In this regard, further research related to the analysis of ecosystem services in mangrove ecosystems is warranted to assess the balance of ecosystem services between beneficiaries (humans) and mangrove resources. This analysis involves several aspects related to mangrove management, which is often known as community-based management.[94]

## 5 Conclusion

Combining spectral imagery and reflectance measurements using an RF algorithm can improve the accuracy value of the classification of mangrove species. The most common mangrove species classified are *R. mucronata*, *S. alba, A. lanata*, and *A. marina*. Reflectance measurement results using a field spectroradiometer for mangrove species, *A. lanata* shows reflectance values between 0.014 and 0.768, *R. mucronata* between 0.002 and 0.493, and *A. marina* between 0.002 and 0.758 and *S. alba* between 0.006 and 0.833. PI that affects the classification model are the red band, NIR band, and GNDVI where the most PI in model 3 is 0.283. Overall, the highest level of agreement in the analysis results for mangrove species was found in model 3. Model 3 is the

best parameter for RF classification, which showed the best mapping accuracy, with the OA, UA, PA, and kappa value being 81.25%, 81.68%, 81.25%, and 0.80, respectively. Future research for improving accuracy value can be used multispatial, multispectral imagery, and adding field data for training samples and check points.

## Disclosures

The authors declare no conflicts of interest.

## Code and Data Availability

The Sentinel Data Hub of the European Space Agency generated the Sentinel-2 pictures utilized in this study accessible to the general public (https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-1c/algorithm-overview). Researchers can access and download Sentinel-2 imagery for a specified coastal region by creating an account on the data hub.

## Acknowledgments

This study was funded by a grant administered from the Directorate of Research, Technology, and Community Service (DRTPM) of the Ministry of Education, Culture, Research, and Technology (Grant No. 160/ IT1.B07.1/SPP-LPPM/IV/2023), to the Lembaga Penelitian dan Pengabdian Masyarakat of the Bandung Institute of Technology, which helped with the administrative and financial processes in research activities. The authors would like to thank the research assistants and assistant surveyors who have helped with obtaining and measuring field data. We want to express our sincere gratitude to the editors and reviewers of the journal for their valuable contributions in helping us finalize this article. The suggestions and reviews provided have been invaluable for improving the quality and completeness of our writing.

## References

1. S. Darmawan et al., "Development of aboveground mangrove forests' biomass dataset for Southeast Asia based on ALOS-PALSAR 25-m mosaic," *J. Appl. Remote Sens.* **13**(04), 1 (2019).
2. R. Morocho et al., "Mangrove forests in Ecuador: a two-decade analysis," *Forests* **13**(5), 656 (2022).
3. C. Giri, "Recent advancement in mangrove forests mapping and monitoring of the world using Earth observation satellite data," *Remote Sens.* **13**(4), 563 (2021).
4. R. Zhang et al., "A comparison of Gaofen-2 and Sentinel-2 imagery for mapping mangrove forests using object-oriented analysis and random forest," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens*. **14**, 4185–4193 (2021).
5. D. Wang et al., "Artificial mangrove species mapping using Pléiades-1: an evaluation of pixel-based and object-based classifications with selected machine learning algorithms," *Remote Sens.* **10**(2), 294 (2018).
6. N. B. Toosi et al., "Land cover classification in mangrove ecosystems based on VHR satellite data and machine learning-an upscaling approach," *Remote Sens.* **12**(17), 2684 (2020).
7. D. Lomeo and M. Singh, "Cloud-based monitoring and evaluation of the spatial-temporal distribution of Southeast Asia's mangroves using deep learning," *Remote Sens.* **14**(10), 2291 (2022).
8. E. McLeod et al., "A blueprint for blue carbon: toward an improved understanding of the role of vegetated coastal habitats in sequestering $CO_2$," *Front. Ecol. Environ.* **9**(10), 552–560 (2011).
9. L. Pendleton et al., "Estimating global 'blue carbon' emissions from conversion and degradation of vegetated coastal ecosystems," *PLoS One* **7**(9), e43542 (2012).
10. N. R. A. Jachowski et al., "Mangrove biomass estimation in Southwest Thailand using machine learning," *Appl. Geogr.* **45**, 311–321 (2013).
11. G. Wolswijk et al., "Can mangrove silviculture be carbon neutral?" *Remote Sens.* **14**(12), 2920 (2022).
12. D. M. Alongi et al., "Indonesia's blue carbon: a globally significant and vulnerable sink for seagrass and mangrove carbon," *Wetl. Ecol. Manag.* **24**(1), 3–13 (2016).
13. D. C. Donato et al., "Mangroves among the most carbon-rich forests in the tropics," *Nat. Geosci.* **4**(5), 293–297 (2011).
14. R. Yanti, "Komposisi Jenis Dan Kerapatan Mangrove Dikawasan Hutan Mangrove Desa Sriminosari Kecamatan Labuhan Maringgaikabupaten Lampung Timur," Uin Raden Intan Lampung, http://repository.radenintan.ac.id/id/eprint/16560 (2021).
15. J. Kamaruza and K. Ibrahim, "Imaging spectrometry on mangrove species identification and mapping in Malaysia," *WSEAS Trans. Biol. Biomed.* **8**(8), 118–126 (2007).

16. F. Kasim, "Characteristics of *Avicennia lanata* (Ridley) species for mangrove restoration on the coast of North Gorontalo," *Tomini J. Aquat. Sci.* **3**(1), 8–20 (2022).
17. R. Djamaluddin, *Mangrove: Biologi, Ekologi, Rehabilitasi, dan Konservasi*, Unsrat Press (2018).
18. Shinta et al, "Identification of mangrove type in mangrove ecosystem area in Pangandaran regency," *J. Akuat.* 9–18 (2022).
19. A. Anand et al., "Use of hyperion for mangrove forest carbon stock assessment in Bhitarkanika forest reserve: a contribution towards blue carbon initiative," *Remote Sens.* **12**(4), 597 (2020).
20. E. L. Bullock et al., "Temporal patterns in species zonation in a mangrove forest in the Mekong Delta, Vietnam, using a time series of Landsat imagery," *Cont. Shelf Res.* **147**, 144–154 (2017).
21. N. R. Chaube et al., "Mangrove species discrimination and health assessment using AVIRIS-NG hyperspectral data," *Curr. Sci.* **116**(7), 1136–1142 (2019).
22. A. Ghorbanian et al., "Mangrove ecosystem mapping using Sentinel-1 and Sentinel-2 satellite images and random forest algorithm in Google Earth Engine," *Remote Sens.* **13**(13), 2565 (2021).
23. S. Darmawan et al., "Identification before-after forest fire and prediction of mangrove forest based on Markov-cellular automata in part of Sembilang National Park, Banyuasin, South Sumatra, Indonesia," *Remote Sens.* **12**(22), 3700 (2020).
24. F. Taureau et al., "Mapping the mangrove forest canopy using spectral unmixing of very high spatial resolution satellite images," *Remote Sens.* **11**(3), 367 (2019).
25. R. P. Parman et al., "Geostatistical analysis of mangrove ecosystem health: mapping and modelling of sampling uncertainty using kriging," *Forests* **13**(8), 1185 (2022).
26. S. Firmansyah, J. L. Gaol, and S. B. Susilo, "Perbandingan Klasifikasi SVM dan Decision Tree untuk Pemetaan Mangrove Berbasis Objek Menggunakan Citra Satelit Sentinel-2B di Gili Sulat, Lombok Timur," *J. Nat. Res. Environ. Manage.* **9**(3), 746–757 (2019).
27. N. H. Quang et al., "Multi-decadal changes in mangrove extent, age and species in the Red River estuaries of Vietnam," *Remote Sens.* **12**(14), 2289 (2020).
28. B. W. Heumann, "An object-based classification of mangroves using a hybrid decision tree—support vector machine approach," *Remote Sens.* **3**(11), 2440–2460 (2011).
29. Z. Li et al., "Remote estimation of mangrove aboveground carbon stock at the species level using a low-cost unmanned aerial vehicle system," *Remote Sens.* **11**(9), 1018 (2019).
30. X. Zhang et al., "Mapping mangrove forests using multi-tidal remotely-sensed data and a decision-tree-based procedure," *Int. J. Appl. Earth Obs. Geoinf.* **62**, 201–214 (2017).
31. I. Jamaluddin et al., "Two decades mangroves loss monitoring using random forest and Landsat data in East Luwu, Indonesia," *Geomatics* **2**(3), 282–296 (2022).
32. S. M. Ghosh, M. D. Behera, and S. Paramanik, "Canopy height estimation using sentinel series images through machine learning models in a mangrove forest," *Remote Sens.* **12**(9), 1519 (2020).
33. T. D. Pham et al., "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 Sens.* **12**(8), 1334 (2020).
34. J. Miao et al., "Mapping seasonal leaf nutrients of mangrove with Sentinel-2 images and XGBoost method," *Remote Sens.* **14**(15), 3679 (2022).
35. Y. Li et al., "Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms," *Sci. Rep.* **10**(1), 9952 (2020).
36. S. Tian et al., "Random forest classification of wetland landcovers from multi-sensor data in the arid region of Xinjiang, China," *Remote Sens.* **8**(11), 954 (2016).
37. V. F. Rodriguez-Galiano et al., "Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture," *Remote Sens. Environ.* **121**, 93–107 (2012).
38. M. D. Behera et al., "Species-level classification and mapping of a mangrove forest using random forest—utilisation of AVIRIS-NG and Sentinel data," *Remote Sens.* **13**(11), 2027 (2021).
39. H. Li et al., "Incorporating the plant phenological trajectory into mangrove species mapping with dense time series Sentinel-2 Imagery and the Google Earth Engine platform," *Remote Sens.* **11**(21), 2479 (2019).
40. L. Wang et al., "A review of remote sensing for mangrove forests: 1956–2018," *Remote Sens. Environ.* **231**, 111223 (2019).
41. M. K. Heenkenda et al., "Mangrove species identification: comparing worldview-2 with aerial photographs," *Remote Sens.* **6**(7), 6064–6088 (2014).
42. S. Paramanik et al., "Species-level classification of mangrove forest using AVIRIS-NG hyperspectral imagery," *Remote Sens. Lett.* **14**(5), 522–533 (2023).
43. J. M. Kovacs, J. Wang, and F. Flores-Verdugo, "Mapping mangrove leaf area index at the species level using IKONOS and LAI-2000 sensors for the Agua Brava Lagoon, Mexican Pacific," *Estuar. Coast. Shelf Sci.* **62**, 377–384 (2005).
44. Y. Jiang et al., "High-resolution mangrove forests classification with machine learning using worldview and UAV hyperspectral data," *Remote Sens.* **13**(8), 1529 (2021).

45. A. D. Rahmandhana, M. Kamal, and P. Wicaksono, "Spectral reflectance-based mangrove species mapping from worldview-2 imagery of Karimunjawa and Kemujan Island, Central Java Province, Indonesia," *Remote Sens.* **14**(1), 183 (2022).
46. T. T. Ajithkumar, T. Thangaradjou, and L. Kannan, "Spectral reflectance properties of mangrove species of the Muthupettai Mangrove Environment, Tamil Nadu," *J. Environ. Biol.* **29**(5), 785–788 (2008).
47. M. F. Ghazali, "Spektral Analisis Untuk Pengamatan Keragaman Vegetasi Mangrove di Labuhan Maringgai, Lampung Timur. (Studi Pendahuluan)," *Semin. Nas. Konserv.* **2020**, 11 (2020).
48. European Space Agency, "Algorithm overview–sentinel-2 MSI level-1C," Sentinel Online Technical Guides, https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-1c/algorithm-overview (2022).
49. M. Kamal et al., "Combining field and image spectral reflectance for mangrove species identification and mapping using WorldView-2 image," *Proc. SPIE* **10790**, 107901P (2018).
50. Badan Informasi Geospasial, "Peraturan Kepala Badan Informasi Geospasial Nomor 3 Tahun 2014 Tentang Pedoman Teknis Pengumpulan Dan Pengolahan Data Geospasial Mangrove," https://jdih.big.go.id/id/home (2014).
51. A. Tridawati et al., "Mapping the distribution of coffee plantations from multi-resolution, multi-temporal, and multi-sensor data using a random forest algorithm," *Remote Sens.* **12**(23), 3933 (2020).
52. Y. Ding et al., "A field-data-aided comparison of three 10 m land cover products in Southeast Asia," *Remote Sens.* **14**(19), 5053 (2022).
53. T. Noi Phan, V. Kuch, and L. W. Lehnert, "Land cover classification using Google Earth Engine and random forest classifier-the role of image composition," *Remote Sens.* **12**(15), 2411 (2020).
54. Y. Gao et al., "Consistency analysis and accuracy assessment of three global 30-m land-cover products over the European Union using the LUCAS dataset," *Remote Sens.* **12**(21), 3479 (2020).
55. European Space Agency, "Sentinel-2 MSI Level-1C algorithms and products," https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-1c-algorithms-products (2022).
56. F. Muchsin, L. Fibriawati, and K. A. Pradhono, "Model Koreksi Atmosfer Citra Landsat-7 (Atmospheric Correction Models of Landsat-7 Imagery)," *J. Penginderaan Jauh dan Pengolah. Data Citra Digit.* **14**(2) (2018).
57. M. Kamal et al., "The effect of field spectral reflectance measurement distance to the spectral reflectance of *Rhizophora stylosa*," *IOP Conf. Ser.: Earth Environ. Sci.* **98** 012059 (2017).
58. A. W. Zulfa et al., "Spectral signature analysis to determine mangrove species delineation structured by anthropogenic effects," *Ecol. Indic.* **130**, 108148 (2021).
59. S. Padma and S. Sanjeevi, "Spectral correlation and Jeffries-Matusita based matching algorithm for improved information extraction from hyperspectral images," in *37th Asian Conf. Remote Sensing, ACRS 2016*, Vol. 1(i), pp. 629–638 (2016).
60. N. Simarmata et al., "Evaluation of multi-temporal imagery and vegetation indices for mangrove mapping (case study of the East Coast of Lampung)," *IOP Conf. Ser. Earth Environ. Sci.* **1083**(1), 012067 (2022).
61. Y. Riko, A. I. Meha, and S. Y. J. Prasetyo, "Perubahan Konversi Lahan Menggunakan NDVI, EVI, SAVI dan PCA pada Citra Landsat 8 (Studi Kasus : Kota Salatiga)," *Indones. J. Comput. Model.* **1**, 25–30 (2019).
62. J. Xue and B. Su, "Significant remote sensing vegetation indices: a review of developments and applications," *J. Sens.* **2017**, 1–17 (2017).
63. H. Nguyen et al., "Classification methods for mapping mangrove extents and drivers of change in Thanh Hoa Province, Vietnam during 2005–2018," *For. Soc.* **4**, 225–242 (2020).
64. A. D. Rahmawati et al., "Vegetation-water-built up index combined: algorithm indices combination for characterization and distribution of mangrove forest through Google earth engine," *CELEBES Agric.* **3**(1), 20–42 (2022).
65. A. Hardianto et al., "PEMANFAATAN Citra Landsat 8 Dalam Mengidentifikasi Nilai Indeks Kerapatan Vegetasi (NDVI) Tahun 2013 dan 2019 (Area Studi: Kota Bandar Lampung)," *J. Geosains dan Remote Sens.* **2**(1), 8–15 (2021).
66. B. Alsaaideh et al., "Mangrove forests mapping in the southern part of Japan using Landsat ETM+ with DEM," *J. Geogr. Inf. Syst.* **05**(04), 369–377 (2013).
67. A. K. Taloor, D. S. Manhas, and G. C. Kothyari, "Retrieval of land surface temperature, normalized difference moisture index, normalized difference water index of the Ravi basin using Landsat data," *Appl. Comput. Geosci.* **9**, 100051 (2021).
68. Y.A. Singgalen and D. Manongga, "Monitoring of mangrove ecotourism area using NDVI, NDWI, and CMRI in Dodola Island, Morotai Island regency, Indonesia monitoring," *J. Ilmu dan Teknol. Kelaut. Trop.* **14**, 95–108 (2022).
69. A. D. Purwanto et al., "Decision tree and random forest classification algorithms for mangrove forest mapping in Sembilang National Park, Indonesia," *Remote Sens.* **15**(1), 16 (2023).
70. L. Breiman, "Random forests," *Mach. Learn.* **45**(1), 5–32 (2001).

71. S. I. Elmahdy et al., "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," *Front. Environ. Sci.* **8**(July), 1–23 (2020).
72. T. M. Berhane et al., "Decision-tree, rule-based, and random forest classification of high-resolution multispectral imagery for wetland mapping and inventory," *Remote Sens.* **10**(4), 580 (2018).
73. M. Belgiu and L. Drăguţ, "Random forest in remote sensing: a review of applications and future directions," *ISPRS J. Photogramm. Remote Sens.* **114**, 24–31 (2016).
74. V. F. Rodriguez-Galiano et al., "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS J. Photogramm. Remote Sens.* **67**, 93–104 (2012).
75. W. Deng et al., "A data mining based system for transaction fraud detection," in *IEEE Int. Conf. Consum. Electron. and Comput. Eng. (ICCECE)*, pp. 542–545 (2021).
76. C. Alkahfi et al., "Variable importance Kesehatan dan Pendidikan dalam Pembentukan IPM dengan algoritme machine learning," *J. Sains dan Inf.* **8**(2), 77–85 (2022).
77. T. T. Dan et al., "Mapping and change analysis in mangrove forest by using Landsat imagery," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **III–8**, 109–116 (2016).
78. J. Aviña-Hernández et al., "Predictive performance of random forest on the identification of mangrove species in arid environments," *Ecol. Inf.* **75**, 102040 (2023).
79. I. Ariawan et al., "Identification mangrove species use forest random algorithm," *J. Kemaritiman: Indones. J. Maritime* **2**(2), 118–128 (2021).
80. N. Anggraini, S. W. Adawiah, and A. D. Purwanto, "Analisis Spektral Reflektan Mangrove di Segara Anakan Dengan Menggunakan Data Penginderaan Jauh," in *Prosiding Pertemuan Ilmiah Tahunan XX dan Kongres VI Masyarakat Ahli Penginderaan Jauh Indonesia (MAPIN)*, pp. 598–605 (2015).
81. A. Alwidakdo, Z. Azham, and D. L. Kamarubayana, "Studi Pertumbuhan Mangrove Pada Kegiatan Rehabilitasi Hutan Mangrove Di Desa Tanjung Limau Kecamatan Muara Badak Kabupaten Kutai Kartanegara," *J. Agrifor.* **13**(1), 11–18 (2014).
82. N. Sahani and T. Ghosh, "GIS-based spatial prediction of recreational trail susceptibility in protected area of Sikkim Himalaya using logistic regression, decision tree and random forest model," *Ecol. Inf.* **64**, 101352 (2021).
83. L. Lymburner et al., "Mapping the multi-decadal mangrove dynamics of the Australian coastline," *Remote Sens. Environ.* **238**, 111185 (2020).
84. A. D. Sakti, W. Takeuchi, and K. Wikantika, "Development of global cropland agreement level analysis by integrating pixel similarity of recent global land cover datasets," *J. Environ. Prot.* **8**(12), 1509–1529 (2017).
85. Y. Zhu et al., "Retrieval of mangrove aboveground biomass at the individual species level with worldview-2 images," *Remote Sens.* **7**(9), 12192–12214 (2015).
86. L. T. H. Pham and L. Brabyn, "Monitoring mangrove biomass change in Vietnam using SPOT images and an object-based approach combined with machine learning algorithms," *ISPRS J. Photogramm. Remote Sens.* **128**, 86–97 (2017).
87. Y. Zhu et al., "Exploring the potential of worldview-2 red-edge band-based vegetation indices for estimation of mangrove leaf area index with machine learning algorithms," *Remote Sens.* **9**(10), 1060 (2017).
88. R. Jhonnerie et al., "Random forest classification for mangrove land cover mapping using Landsat 5 TM and ALOS PALSAR imageries," *Procedia Environ. Sci.* **24**, 215–221 (2015).
89. U. Nguyen et al., "Mapping vegetation types in semi-arid riparian regions using random forest and object-based image approach: a case study of the Colorado River Ecosystem, Grand Canyon, Arizona," *Ecol. Inf.* **50**, 43–50 (2019).
90. P. I. Macreadie et al., "Quantifying and modelling the carbon sequestration capacity of seagrass meadows—a critical assessment," *Mar. Pollut. Bull.* **83**(2), 430–439 (2014).
91. B. Fu et al., "Comparison of RFE-DL and stacking ensemble learning algorithms for classifying mangrove species on UAV multispectral images," *Int. J. Appl. Earth Obs. Geoinf.* **112**, 102890 (2022).
92. R. Dasgupta, S. Hashimoto, and O. Saito, "Envisioning the future of mangroves through mapping and modeling of mangrove ecosystem services," in *Assessing, Mapping and Modelling of Mangrove Ecosystem Services in the Asia-Pacific Region*, R. Dasgupta, S. Hashimoto, and O. Saito, Eds., pp. 1–12, Springer Nature Singapore, Singapore (2022).
93. B. Zhu, J. Liao, and G. Shen, "Spatio-temporal simulation of mangrove forests under different scenarios: a case study of mangrove protected areas, Hainan Island, China," *Remote Sens.* **13**(20), 4059 (2021).
94. S. Handayani et al., "Pemetaan Jasa Ekosistem Mangrove pada Wilayah Rehabilitasi di Pesisir Sayung, Kabupaten Demak (mapping of mangrove ecosystem services in rehabilitation areas of Sayung Coastal Zone, Demak District)," *Jurnal Ilmu Pertanian Indonesia* **25**(4), 574–583 (2020).


**Nirmawana Simarmata** is currently pursuing his doctoral degree in geodesy and geomatics engineering at ITB. His research interests include spatial modeling for blue carbon estimation in coastal areas based on machine learning.

**Soni Darmawan** is a lecturer in geomatics engineering at the National Institute of Technology. He received a doctorate degree in geodesy and geomatics engineering in 2011. His research interests are in remote sensing applications for environment and disaster and global geodatabase.

Biographies of the other authors are not available.