Heliyon 9 (2023) e20745

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Estimating mangrove aboveground biomass in the Colombian Pacific coast: A multisensor and machine learning approach

John Josephraj Selvaraj [a,b,*], Bryan Ernesto Gallego Pérez [b]

[a] Universidad Nacional de Colombia, Tumaco Campus, Institute of Pacific Studies - Kilómetro 30-31, Cajapí Vía Nacional Tumaco-Pasto, Tumaco, 528514, Nariño, Colombia

[b] Universidad Nacional de Colombia, Palmira Campus, Faculty of Engineering and Administration, Department of Engineering, Research Group on Hydrobiological Resources, Cra 32 No. 12 - 00, Palmira, Código Postal 763533, Colombia

A B S T R A C T

The Colombian Pacific Coast is renowned for its exceptional biodiversity and hosts vital mangrove ecosystems that benefit local communities and contribute to climate change mitigation. Therefore, estimating mangrove aboveground biomass (AGB) in this region is crucial for planning and managing these coastal forest covers, ensuring the long-term sustainability of the essential environmental services provided by the Colombian Pacific Coast (CPC). This study employed a spatial estimation approach to assess mangrove AGB, evaluating various parametric and non-parametric models using a multisensor combination and machine learning on the Google Earth Engine (GEE) platform within the CPC. Synthetic aperture radar (SAR) satellite imagery (ALOS-2/PALSAR-2, SRTM, NASADEM, and ALOSDSM) and optical data (Landsat 8) were utilized to quantify mangrove AGB in 2022 across the four departments of the CPC. The Random Forest model exhibited superior predictive performance compared to the other models evaluated, achieving values of $R^2 = 0.783$, RMSE = 38.239 [Mg/ha], MAE = 27.409 [Mg/ha], and BIAS = 0.164. Our findings reveal that the mangrove AGB map for the CPC exhibits a mean ± standard deviation of 181.236 ± 28.939 [Mg/ha] across eight classes, ranging from 88.622 [Mg/ha] to 378.21 [Mg/ha]. This research provides valuable information to inform and strengthen various management strategies and decision-making processes for the mangrove forests of the CPC.

1. Introduction

A vital role in the carbon cycles of both land and sea is played by the mangrove forest biome, which is a type of forest cover [1]. Located in the intertidal zones of some tropical and semitropical regions, these coastal ecosystems provide numerous ecosystem services, such as climate change mitigation, and are among the primary ecosystems for organic carbon fixation and sequestration by area [2,3]. The unique ecological characteristics of these forests lie in their ability to efficiently use and sequester atmospheric carbon, which is due to their duality of terrestrial and marine ecosystem traits [2]. Land use changes caused by activities such as agricultural production, aquaculture, deforestation, and urban sprawl are primarily responsible for reducing the quality and quantity of notable benefits provided by these ecosystems, which are currently among the most vulnerable [4,5].

Measuring biomass in tropical forest ecosystems and its relationship with carbon sequestration is of great importance in the context of developing or improving national and regional policies to reduce greenhouse gas (GHG) emissions, as deforestation or potential deterioration of forest cover is responsible for approximately 10 % of global GHG emissions [6]. In situ, quantifying aboveground biomass (AGB) in mangroves is typically an accurate methodology, but it can be costly and challenging to apply over large areas. Consequently, a practical and economical alternative involves combining remote sensing with field measurements to quantify AGB.

* Corresponding author. Universidad Nacional de Colombia - Tumaco Campus, Institute of Pacific Studies - Kilómetro 30-31, Cajapí Vía Nacional Tumaco-Pasto, Tumaco, 528514, Nariño, Colombia.

E-mail address: jojselvaraj@unal.edu.co (J.J. Selvaraj).

https://doi.org/10.1016/j.heliyon.2023.e20745

Received 27 July 2023; Received in revised form 17 September 2023; Accepted 5 October 2023

Available online 19 October 2023

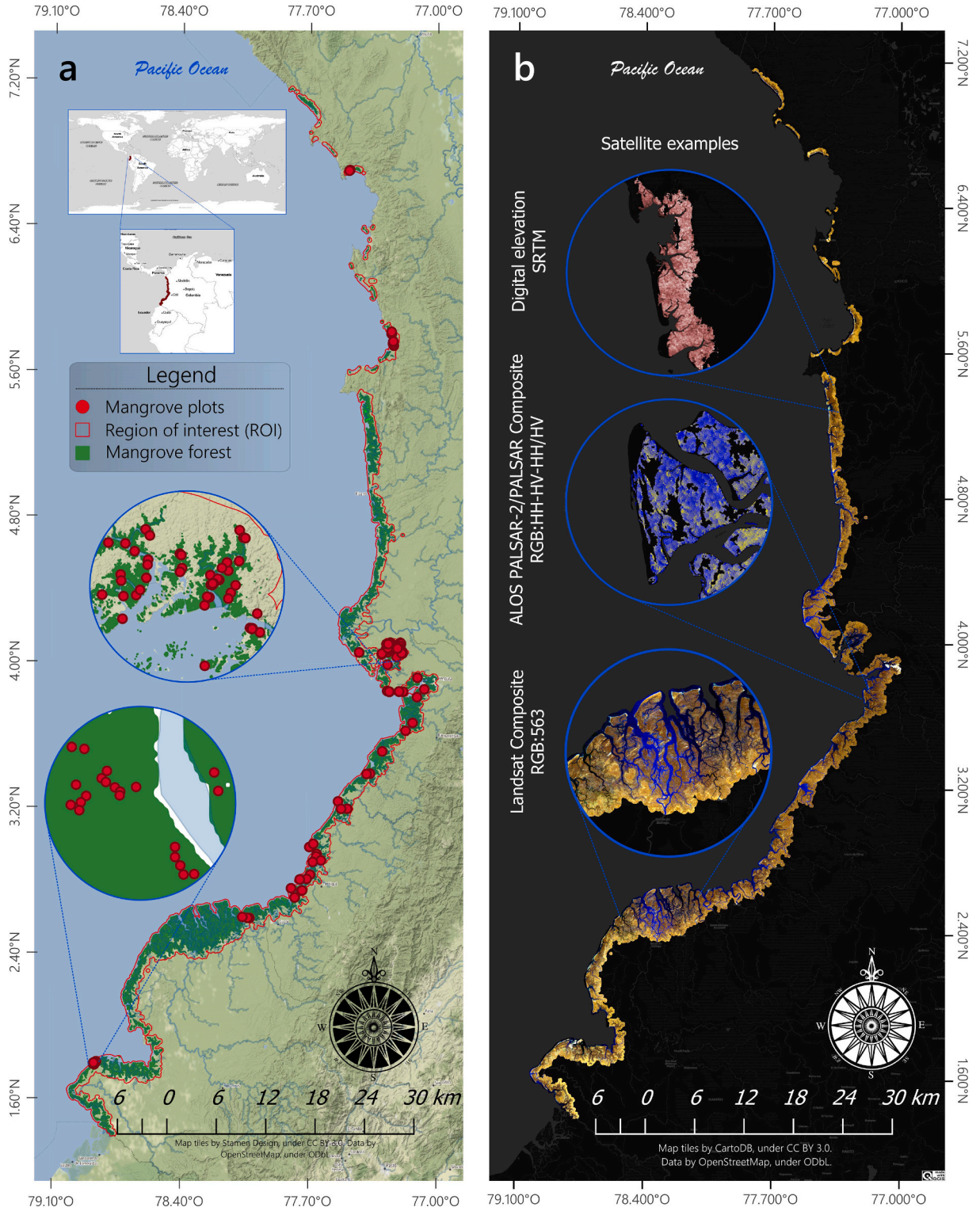

Fig. 1. (a) Location of the study area with the plots, within the Region Of Interest (ROI) and mangrove forest defined at [19] (b) Samples of satellite information used in different locations, and their false color composition.

Although remote sensing is not as accurate as field measurements, it is an effective option for estimating AGB at regional or continental scales, aiding in understanding and analyzing forests [6,7].

Remote sensing delivers abundant spectral and textural data at varying temporal and spatial resolutions, offering invaluable insights into studying the earth's surface [8]. This data, harvested by both active (such as synthetic aperture radar (SAR) and light localization (LiDAR)) and passive sensors (like optical imaging), is frequently used to analyze structural parameters, including height, and its correlation with mangrove ecosystem composition [9]. Several studies have demonstrated a substantial and consistent link between mangrove height and its biomass content [1,10,11]. Moreover, canopy height potently predicts biomass and carbon in forest ecosystems [12]. Likewise, a recent study indicated that taller mangroves sequester more carbon than their shorter counterparts [9]. In this investigation, a multi-sensor approach was adopted to gauge the Aboveground Biomass (AGB) in the CPC mangroves using satellite data from active sources like ALOS PALSAR/PALSAR2, Shuttle Radar Topography Mission (SRTM) V3, NASADEM, ALOS Global Digital Surface Model (ALOSDSM), and passive sources like Landsat 8 OLI/TIRS sensor. Integrating this satellite data with forest inventory plots forms a foundational base for estimating AGB in CPC mangroves.

Several methods for estimating AGB through remote sensing, including parametric models such as linear regression (LR), assume a direct relationship between predictive variables and results. However, more than conventional statistical methods might be needed to fully capture the potential correlation between biomass in a forest and remotely sensed data [10]. In contrast, non-parametric models such as Gradient Tree Boost (GTB) or Random Forest (RF) can integrate various variables following different statistical distributions. Additionally, some studies on forest cover suggest that machine-learning techniques can, in some cases, produce better AGB estimates

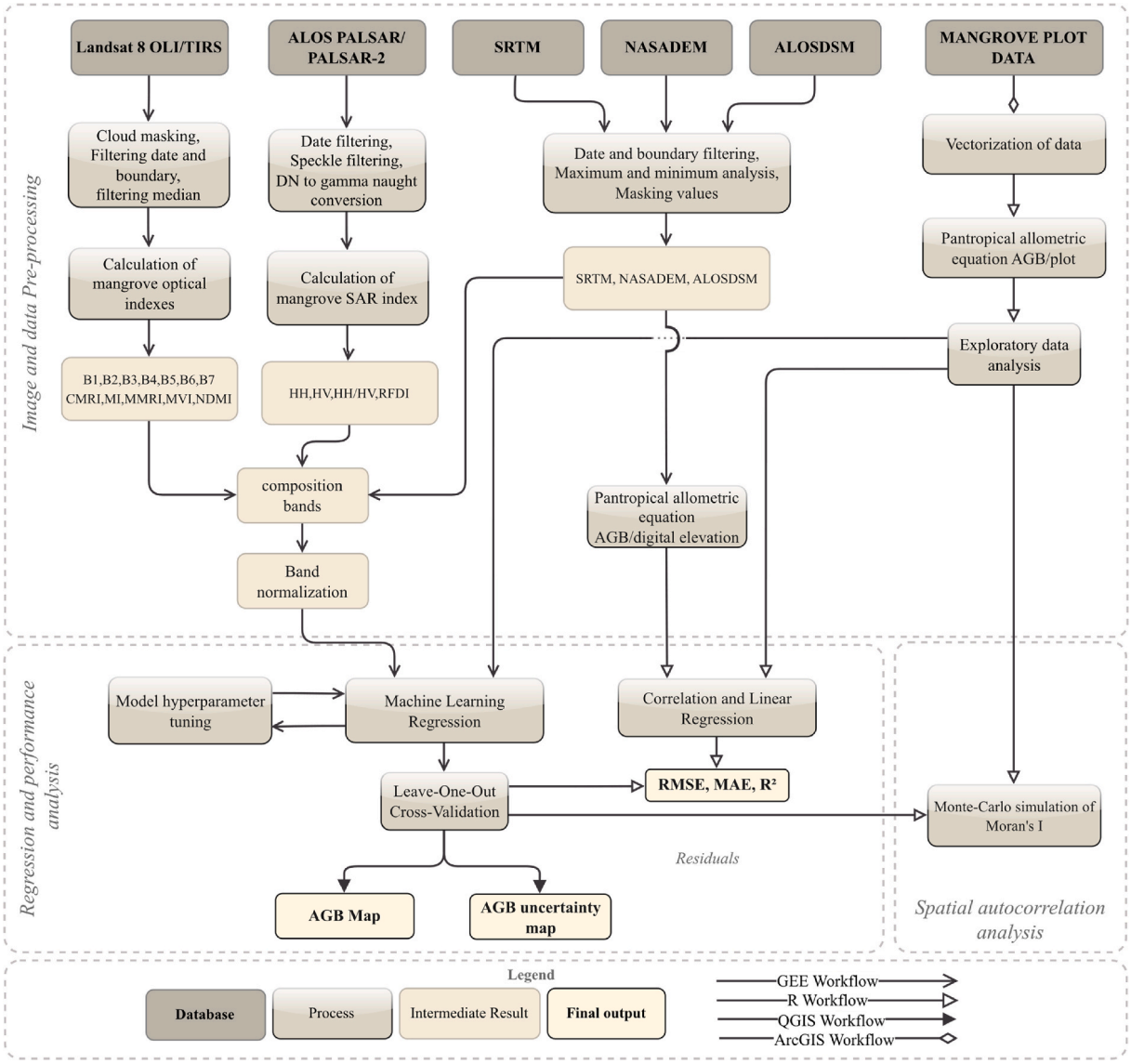

Fig. 2. General methodological scheme of the workflow.

than traditional techniques [13].

Most of Colombia's mangrove forest cover is on the Pacific Coast (CPC). However, most studies on mangroves have been conducted on the Caribbean Coast [14]. Furthermore, the CPC is part of the Chocó/Darién hotspot, which boasts one of the world's highest biodiversity records [5]. Therefore, it is crucial to conduct research that contributes to scientific knowledge, in this case, by focusing on the AGB associated with the mangrove cover of the CPC. In this research, we evaluated two remote sensing approaches applied on the geospatial analysis platform Google Earth Engine (GEE) [15]: first, a parametric estimation of mangrove AGB in the CPC using LR, and second, a non-parametric estimation of mangrove AGB in the CPC using two algorithms, RF and GTB. The specific objectives of this research were as follows: (1) to explore the potential presence of spatial autocorrelation of AGB in the mangrove plots of the CPC; (2) to evaluate and compare the accuracy and predictive ability of LR, RF, and GTB models in estimating mangrove AGB in the CPC, and (3) to estimate mangrove AGB in the CPC and the uncertainty associated with the model demonstrating the best estimation performance.

2. Description of the study area

The study area is located in western Colombia and encompasses four Colombian Pacific Coast (CPC) departments: Chocó, Valle del Cauca, Cauca, and Nariño. This coastal region extends from north to south, from Punta Ardita on the border with Panama (coordinates 7°12′ N and 77°52′ W) to the Mataje River on the border with Ecuador (coordinates 1°28′ N and 78°46′ W) [16]. The coastline features diverse geomorphology, including barrier islands, sandy beaches, cliffs, alluvial valleys, tidal flats, intertidal mudflats, and mangrove swamps [17]. The CPC is characterized by tropical rainforests with high biodiversity [18]. The region has an average annual rainfall ranging from 6821 mm to 7673 mm; however, in some areas, such as Alto Atrato, rainfall can exceed 10,000 mm, making it one of the rainiest places in the world. The average annual temperature is 25.7 °C and varies between 19 °C and 32 °C [16].

3. Methods

3.1. General methodological approach of the workflow

The main methodology of this research was developed through scripting in the GEE platform. Additionally, processes were used for some of the data in ArcGIS, QGIS and R. Fig. 2 summarizes the process through three general sections: (1), the preprocessing of spatial information that included the selection, masking, and temporal and spatial filtering of the satellite data. (2), the application of LR models for the digital elevation data (SRTM, NASADEM and ALOSDSM), and the application of RF and GTB models for a stack of 19 satellite bands from Landsat 8, ALOS PALSAR/PALSAR2, SRTM, NASADEM and ALOSDSM. In addition, in this section is the performance analysis of the different models, the cross-validation process along with the uncertainty analysis. (3) The spatial autocorrelation analysis for the AGB data of the plots by means of the global Moran index.

3.2. Preprocessing of mangrove plots

A total of 152 mangrove plots were included in this study for the region of interest (ROI), these plots are distributed throughout the CPC (Fig. 1a). Of the 152 plots, 59.2 % (90 plots) were derived from other studies, while the remaining 40.8 % (62 plots) are new plots from this study (Table 1). For the new plots collected between March and November 2022, the sampling protocol for structural parameters recommended by Ref. [20] was followed, where the plots were determined through random sampling in the departments of Valle del Cauca and Nariño (Table SM2).

A limitation of this study was the sample plot availability for the ROI. Although it is crucial to maintain a relationship between the sample size (plots) and the satellite imagery's pixel size, various logistical and resource challenges prevented adjusting the sample size or adding more plots to align precisely with the imagery's spatial resolution. Consequently, we used 152 samples with circular and square geometries, ranging from 0.015 to 0.190 ha (Table SM2). Given the limited field data available, this approach provided the best feasible approximation to assess mangrove AGB variability in the ROI.

Since canopy height presents a strong relationship with forest AGB and provides valuable information on vertical stand structures

Table 1
Database of mangrove plots by department and study.

Reference	Departament				Total_plots
	Chocó	Valle del Cauca	Cauca	Nariño	
[26]		11			11
[27]			4		4
[28]			12		12
[29]		44			44
[30]	10				10
[31]	2				2
[1]	4				4
[32]				3	3
This study (2022)		28		34	62
Total_plots	16	83	16	37	152

[21], the allometric relationship proposed by Ref. [11] was used to calculate AGB per plot. This pantropical allometric equation (Eq.1) is a function of mean stand height and was estimated with globally distributed field data. Moreover, it has been employed in several investigations involving mangroves [1,6,22].

$$AGB\,[Mg\,/\,ha] = 10.8 * Heigth(m) + 34.994\ (r = 0.774, P < 0.0001) \quad \text{Eq.1}$$

Subsequently, an exploratory data analysis (EDA) for AGB was conducted using a script in R, where different statistical parameters were calculated with the libraries (basic, nortest and boot) [23–25] (Figure SM1).

3.3. Spatial autocorrelation analysis

Following the recommendations of [33], this study evaluated the spatial autocorrelation for the AGB dataset of the plots (training data) and residuals of the selected model in Section 4. This analysis was implemented using the package 'spdep' for R software, applying a permutation test for Moran's I statistic with the function "moran.mc" and using 1000 simulations within the test [34] (Fig. 4a).

3.4. Satellite data

In this research, a modification of the multisensor approach for AGB estimation proposed by Ref. [8] was used in the mangroves of the CPC. The satellite information used comes from active and passive sensors, which are stored in the GEE catalog [35]. The first group belongs to the Synthetic Aperture Radar (SAR) record and is divided into two sections: (1) corresponds to the ALOS PALSAR/PALSAR2 sensors. This information has undergone several preprocessing steps, such as accurate radiometric and geometric corrections, which allow the integration of these data into the workflow [33,36,37]. (2) Three types of digital elevation data were used: (I) Shuttle Radar Topography Mission (SRTM) V3, featuring gap-filling preprocessing using open-source data, such as ASTER Global Digital Elevation Model (ASTER GDEM2), The Global Multi-resolution Terrain Elevation Data (GMTED2010), and National Elevation Dataset (NED) [38]; (II) NASADEM, which is a reprocessing of SRTM information with improvements in accuracy by including ancillary data from ASTER GDEM, The Geoscience Laser Altimeter System (GLAS) instrument on the Ice, Cloud, and Land Elevation Satellite (ICESat GLAS), and PRISM databases [39]; (III) ALOS Global Digital Surface Model (ALOSDSM), which has an improved outlier detection method and was generated from a 5 m resolution DSM [40].

For the second group, we worked with the Landsat 8 OLI/TIRS sensor. The surface reflectance optical information of Collection 2 has undergone orthorectification and atmospheric correction preprocessing that guarantee the quality of the information [41,42]. Table 2 shows the configuration of the satellite datasets used, and Fig. 1b shows samples of the information used.

3.4.1. ALOS PALSAR/PALSAR2 and SAR index preprocessing

The annual ALOS PALSAR/PALSAR2 mosaics were processed in four stages using a script in GEE. (1), Date and mean filtering (Table 2) was performed to reduce seasonal and phenological effects of vegetation cover on the images [33]. (2), A "Refined Lee" speckle reduction filter was applied [43]. (3), Digital number values were converted to zero gamma backscatter values γ_0 [dB], using Eq.2 [44]. (4), The Radar Forest Degradation Index (RFDI) was calculated through an expression [45] that allows to improve the

Table 2
Satellite data set configuration.

Satellite data	Features	Dataset
ALOS PALSAR/PALSAR-2	Core study year	2021
	Image Collection	JAXA/ALOS/PALSAR/YEARLY/SAR
	Bands used	HH, HV
	Pixel resolution/Range years composite	25 m/2015–2022
(SRTM) V3	Core study year	2000
	Image Collection	USGS/SRTMGL1_003
	Bands used	Elevation
	Pixel resolution/Range years composite	30 m/2000
NASADEM	Core study year	2000
	Image Collection	NASA/NASADEM_HGT/001
	Bands used	Elevation
	Pixel resolution/Range years composite	30 m/2000
ALOSDSM	Core study year	2011
	Image Collection	JAXA/ALOS/AW3D30/V3_2
	Bands used	DSM
	Pixel resolution/Range years composite	30 m/2006–2011
Landsat 8 OLI/TIRS (Surface Reflectance)	Core study year	2022
	Image Collection	LANDSAT/LC08/C02/T1_L2
	Bands used	B1,B2,B3,B4,B5,B6,B7
	Pixel resolution/Range years composite	30 m/2015–2023

analysis of the structure and quantification of biomass in forests [46] (Table 3). Finally, a 4-band compositing was performed (Fig. 2).

$$\gamma^{\circ} = 10\log_{10}(DN^2) - 83.0 \quad \text{Eq.2}$$

3.4.2. Preprocessing of landsat 8 imagery and mangrove indices

Landsat 8 images were processed using a script in GEE in four stages. (1), A cloud masking function and a scale factor were applied to the image collection [47]. (2), A boundary filter was employed with the ROI and a date filter (Table 2). (3), A median filter applied to all matched bands in the collection image set was used to obtain a single composite image to facilitate spectral analysis [48,49]. (4), Five specific indices for mangrove recognition were calculated through expressions [45] (Table 3). Finally, a 12-band composition was performed (Fig. 2).

3.4.3. Preprocessing of digital elevation images

SRTM, NASADEM and ALOSDSM digital elevation images were processed in 3 stages for each one. (1), Boundary filters for ROI and date were used (Table 2). (2), The height records of mangroves in the CPC were reviewed in the literature and reports of the tallest trees for the CPC were found; this allowed for establishing the limiting heights to apply masking of elevation values in the range of −2 to 60 m height in the images within the ROI, in order to decrease the possible noise introduced by elevation outliers [1,32]. (3), Finally a 3-band composite was performed (Fig. 2).

3.4.4. Composition and normalization of satellite data

After the satellite preprocessing, a composition was performed to create a single 19-band package in a collection of images. Then, following the recommendations of [55], since the satellite information presented different ranges of values, a min-max normalization method was applied to transform the range of values from zero to one to improve the generalization capacity of the machine learning models.

3.5. Predictive methods

3.5.1. Evaluation of model performance

[56] recommend evaluating the performance of estimation models for various models by employing the performance parameters: root mean square error (RMSE) (Eq.3), mean absolute error (MAE) (Eq.4), coefficient of determination (R^2) (Eq.5) and bias (Bias) (Eq.6).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad \text{Eq.3}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \quad \text{Eq.4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2} \quad \text{Eq.5}$$

$$Bias = \sum_{i=1}^{n}\frac{(\hat{y}_i - y_i)}{n} \quad \text{Eq.6}$$

Where $\hat{y}_i, y_i$ represent prediction and observation data for the nth plot respectively; n is the total number of plots to be validated and $\bar{y}_i$ is the mean of the biomass observations.

Table 3
Indices used for the CPC mangroves.

Index type	Name Index	Formula	Reference
Specific optical indices for mangroves	Modular Mangrove Recognition Index	$MMRI = \lvert MNDWI\rvert - \lvert NDVI\rvert / \lvert MNDWI\rvert + \lvert NDVI\rvert$	[50]
	Mangrove Vegetation Index	$MVI = \lvert NIR\rvert - \lvert GREEN\rvert / \lvert SWIR_1\rvert - \lvert GREEN\rvert$	[51]
	Mangrove Index	$MI = (NIR - SWIR)/(NIR * SWIR) * 10000$	[52]
	Normalized Difference Mangrove Index	$NDMI = (SWIR_2 - GREEN)/(SWIR_2 + GREEN)$	[53]
	Combined Mangrove Recognition Index	$CMRI = (NDVI - NDWI)$	[54]
SAR Index	Radar Forest Degradation Index	$RFDI = \sigma^{\circ}_{HH} - \sigma^{\circ}_{HV} / \sigma^{\circ}_{HH} + \sigma^{\circ}_{HV}$	[46]

Where: NIR = near infrared band, RED = red band, SWIR = shortwave infrared band, MIR = middle infrared band.
NDVI = Normalized Difference Vegetation Index, NDWI = Normalized Difference Water Index.
MNDWI = Modified Normalized Difference Water Index, σ°_{HH} , σ°_{HV} = Backscattering of dual polarizations.

3.5.2. Correlations and linear models

In this section, digital elevation data (SRTM, NASADEM and ALOSDSM) were used to calculate AGB in GEE with Eq.1. Then, an extraction (ee.Image.sampleRegions) per pixel/plot of the AGB value corresponding to each plot location in the respective images [57] was performed. Next, the respective correlation analyses and LR of each image were performed in R. Then, for the residuals of the LR, the assumptions of normality, homoscedasticity, and linearity [58] were evaluated in R with the package 'lmtest'. Finally, performance parameters were evaluated.

3.5.3. Machine learning regression models

The nonparametric RF algorithm is a widely used technique for forest AGB estimates [59]. RF operates by combining the predictions of the results of multiple random and independent decision trees using different subsets of data, which decreases the over-fitting in the final estimate [60,61]. The other GTB algorithm, employs a sequential decision tree construction technique where each new tree is trained to correct the prediction errors of the previous model [62]. In this method the final prediction is made by combining the predictions of individuals weighted by a learning factor related to the importance in the prediction [59]. Additionally, for both models in GEE using the function 'Classifier.explain (.)', the importance of the explanatory variables in the prediction processes was calculated. [63], mentions that the greater the contribution in the estimation results of a variable, the higher its importance score.

3.5.4. Hyperparameter tuning

Machine learning algorithms can be configured by modifying their hyperparameters, which can impact the algorithm's behavior. However, the manual selection of values for tuning hyperparameters is a time-consuming potentially biased, and susceptible to errors [64]. As a result, an initial script was coded in GEE to enable the iterative tuning of the number of decision trees (DT) for each model (RF and GTB) based on performance parameters. At the same time, the other hyperparameters were set to default values. The goal was to minimize computational costs by selecting the smallest number of trees that maximized the model performance (R^2) with the lowest possible errors (RMSE and MAE). Finally, in Fig. 3, the values of the performance parameters were normalized to facilitate the visualization of the tuning process.

3.5.5. Cross-validation and uncertainty mapping

The literature indicates that splitting data for training validation in sparse datasets can be inefficient [65]. Therefore, this study employed the leave-one-out cross-validation (LOO-CV) approach (with 152 folds), where one validation site is reserved in each iteration. LOO-CV was applied by modifying the script compiled for the GEE platform by Ref. [66]. This procedure was applied to both algorithms; using the average of the 152 AGB predictions, the respective AGB estimates were obtained for the proposed models. Subsequently, the performance parameters of each model were calculated.

Following the advice of Zhou et al. (2019), the 95 % confidence intervals (CIs) of the 152 respective predictions were calculated for each pixel, suggesting that the AGB values present a 95 % chance of being within those CIs (Eq.7). Eq.8 was then used to calculate the uncertainty of the AGB estimates.

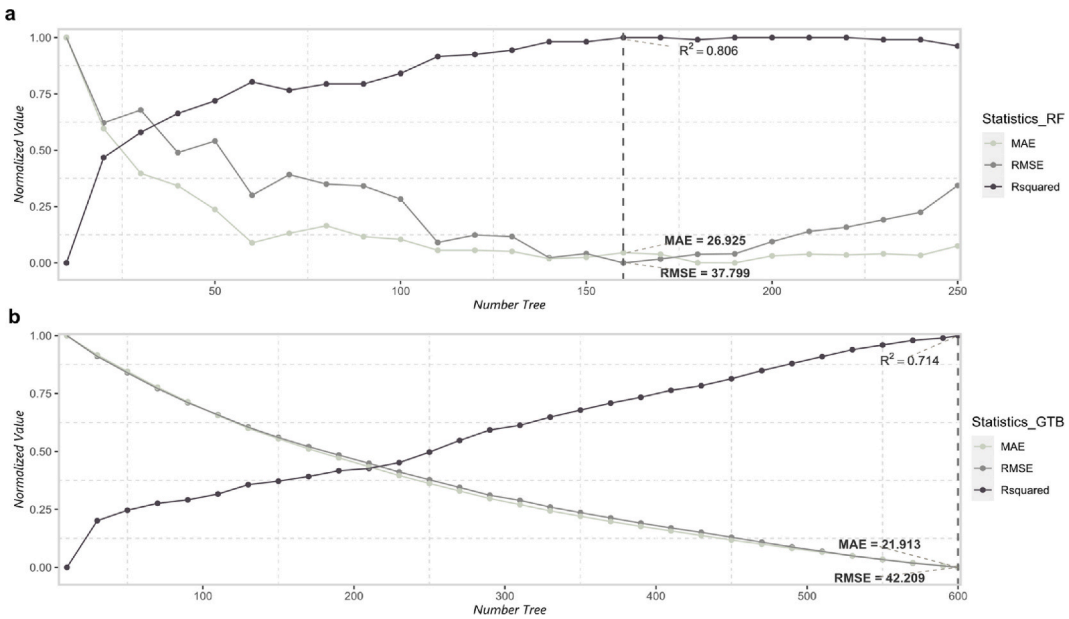

Fig. 3. (a), tuning of the DT for RF. (b), tuning of the DT for GTB

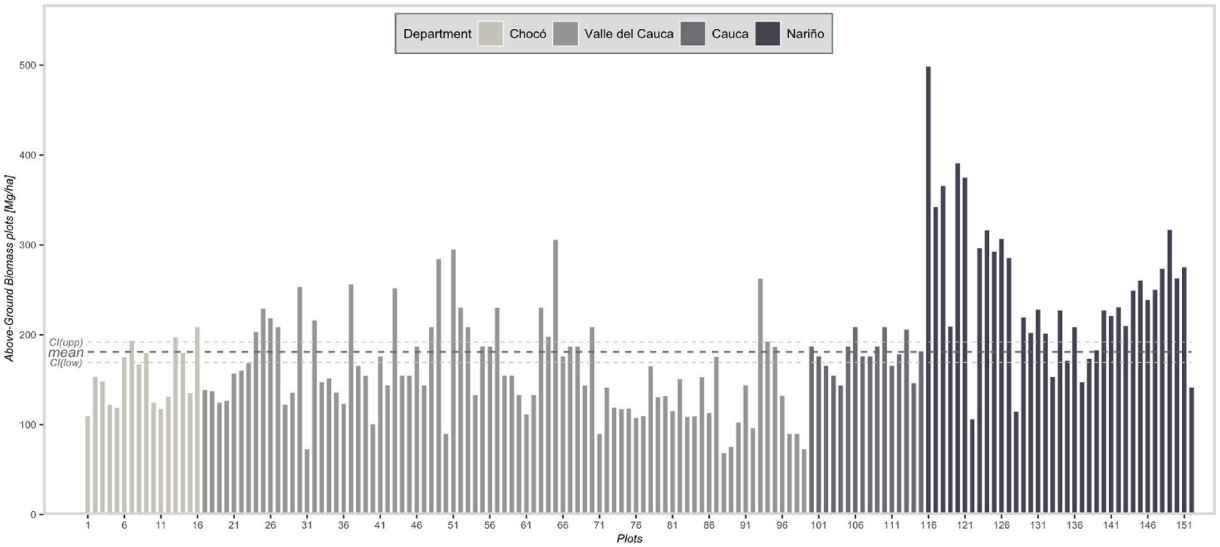

Fig. 4. Field plots profile of above-ground biomass in the study site.

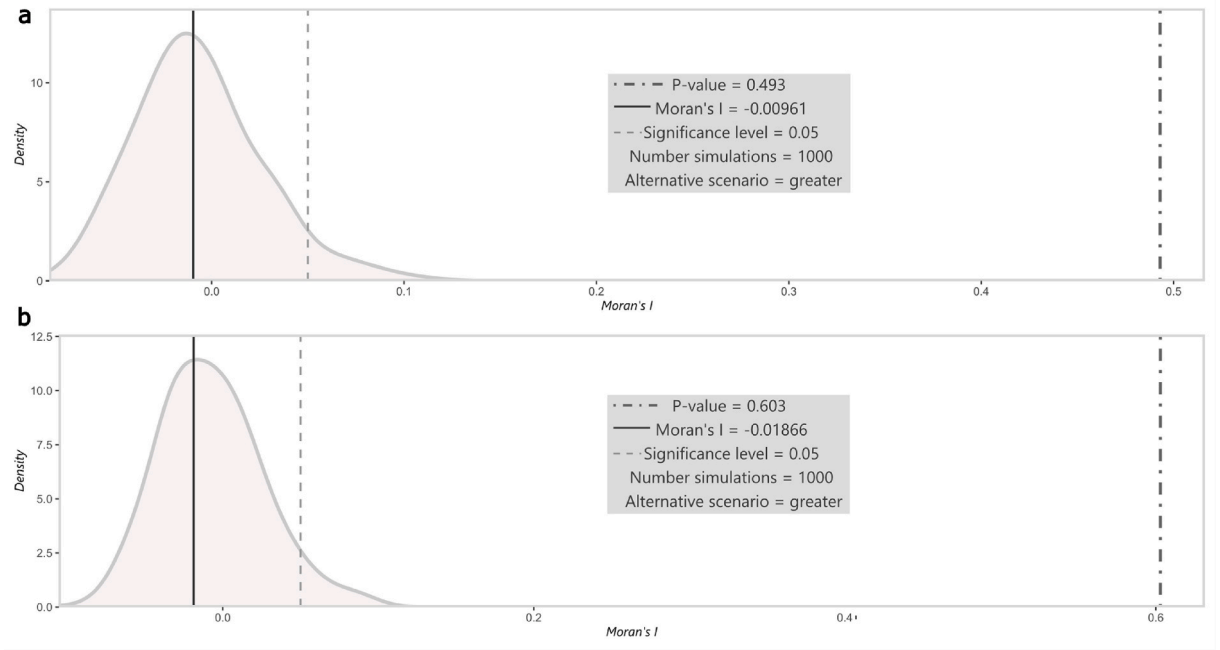

Fig. 5. Monte-Carlo simulations.

$$CIs = \overline{P} \pm 1.96 \frac{S}{\sqrt{n}} \quad \text{Eq.7}$$

$$U = \frac{CI_{upper} - CI_{lower}}{\overline{P}} \quad \text{Eq.8}$$

where $\overline{P}$ is the mean AGB of the n predictions, in this study n = 152. S is the standard deviation of the n predictions.

4. Results

4.1. EDA and spatial autocorrelation

The EDA revealed that the mean AGB of the 152 mangrove plots was 180.965 Mg/ha (CI_{upper}:191.870/CI_{lower}:169.255). Fig. 4 displays the distribution of the plots by department where heterogeneous values of AGB are observed. However, although it can be identified that the highest values of AGB per plot are located in the department of Nariño.

The Monte-Carlo simulation results for the Moran's I statistic of −0.00961 and the P-value>0.05 (Fig. 5a) suggest a slight negative spatial autocorrelation; however, the P-value = 0.493 indicates that there is insufficient statistical evidence to reject the null hypothesis of no significant spatial autocorrelation for the AGB plots in the CPC. This scenario was, similarly observed for the residuals of the model selected in section 4.2 (Fig. 5b).

4.2. Modeling, correlation, comparison and evaluation results

In Fig. 6 a, b, and c, the AGB dispersion models of the CPC mangroves for the three digital elevation data are observed; they presented low correlation coefficients: R_{SRTM} = 0.305, $R_{NASADEM}$ = 0.323 and $R_{ALOSDSM}$ = 0.174. In addition, normality tests for the residuals of the LR models indicate that they are not normally distributed. Also, in two of the models (SRTM and NASADEM), evidence of autocorrelation was found in their residuals (Table SM1). In Contrast, in Fig. 6 d and e, in the dispersion models for the RF and GTB estimates of AGB of the CPC mangroves for the 19 explanatory variables, high and similar correlation coefficients were presented: R_{RF} = 0.885 y R_{GTB} = 0.820.

Then, the yield parameters in the estimation of the 5 AGB models for the CPC mangroves were compared (Table 4) and the RF model presents the highest R^2 = 0.783, with the lowest error and deviation values; RMSE = 38.239 [Mg/ha], MAE = 27.409 [Mg/ha] and BIAS = 0.164.

4.2.1. Variable importance

Fig. 7 displays 19 multispectral bands selected by the RF algorithm. From left to right 11 bands exceeded 50 % of their relative importance, meaning they were more sensitive to mangrove AGB. Additionally, a heterogeneous representation of Landsat 8 optical

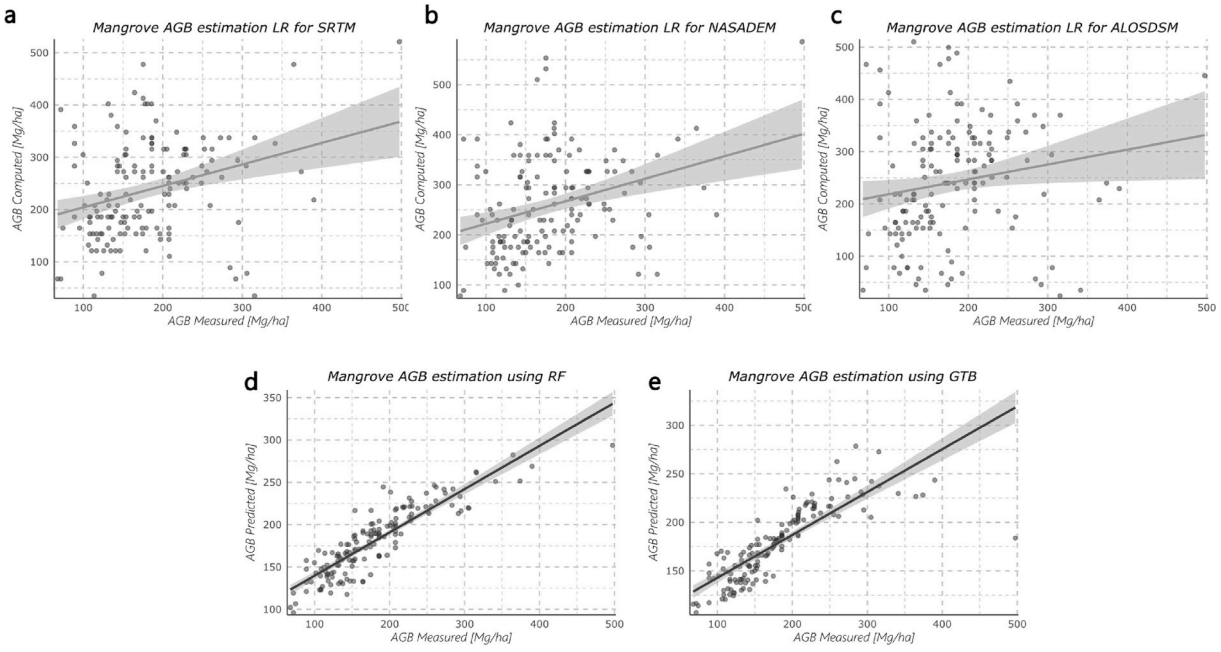

Fig. 6. Comparison of scatter plots of mangrove AGB estimation models in the CPC.

Table 4
Parameters observed for the estimation models.

Parameters	LR_{SRTM}	$LR_{NASADEM}$	$LR_{ALOSDSM}$	RF[a]	GTB
R^2	0.093	0.104	0.030	0.783	0.673
RMSE	112.074	124.905	134.975	38.239	43.7454
MAE	84.33	95.346	105.333	27.409	24.5247
Bias	56.435	77.608	60.556	0.164	−2.385

[a] The best performance.

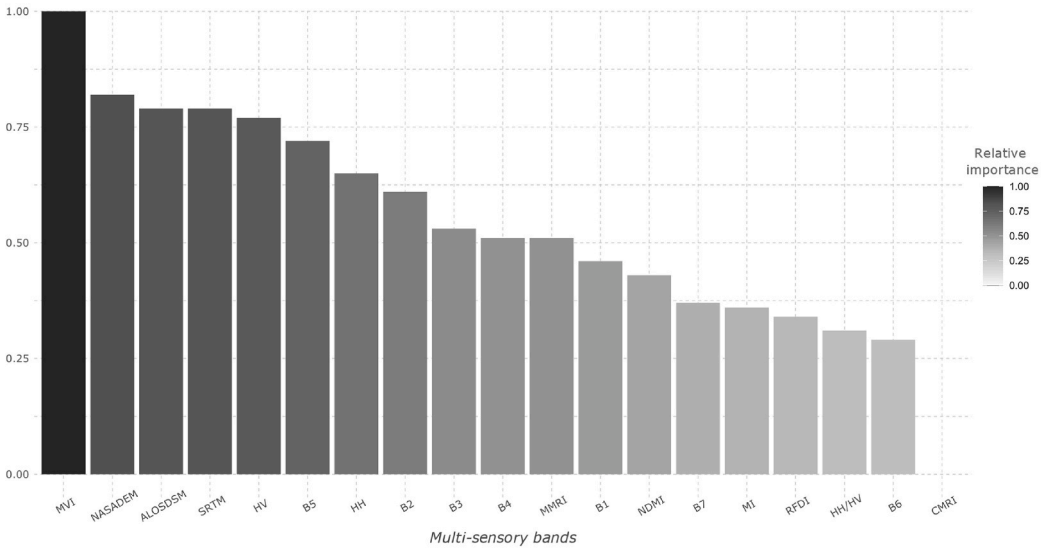

Fig. 7. Relative importance of the 19 optical and radar sensor variables in this study.

bands (B5: near infrared, B2: blue, B3: green, B4: red), ALOS PALSAR/PALSAR2 dual polarimetric SAR data (HV and HH) and digital elevation (NASADEM, SRTM and ALOSDSM) is observed among the main importance values. Interestingly, we found that the MVI and MMRI indices present a strong and medium sensitivity with AGB in the study area respectively. The remaining eight bands present values below 50 % of relative importance where the CMRI index stands out with a weak importance with the AGB of the mangrove in the study area.

4.2.2. Generation of mangrove AGB maps and their uncertainty in the study area

The final raster layer was exported in GeoTIFF format and then overlaid for each department of the CPC (Fig. 8). The CPC mangrove AGB map is shown in 8 classes, ranging from 88.622 Mg/ha to 378.21 Mg/ha, with a mean ± standard deviation of 181.236 ± 28.939 Mg/ha (Table 5). The primary trend in the study area for mangrove AGB values is heterogeneous, as different AGB classes are mixed throughout the CPC. However, some discernible patterns of distinct groupings exist in northern Nariño La Tola and El Charco municipalities and in the District of San Andres de Tumaco in southern Nariño, where AGB values range from roughly 127–246 Mg/ha. To the north, also in Nariño (Santa Bárbara) and part of Cauca in the municipalities of Guapi, Timbiquí, and López de Micay, higher AGB values are found, oscillating approximately between 246 and 370 Mg/ha.

For the department of Valle del Cauca, from Ají Island to the south of Buenaventura Bay, a heterogeneous trend is observed, with AGB values ranging from 127 to 276 Mg/ha. To the north of Buenaventura Bay, from Cangrejo Island through the Santa Clara estuary in La Bocana to El Tigre Cove, slightly lower AGB values are observed, ranging from 127 to 186 Mg/ha. Further north, around the influence from Bahía Málaga to the Pital estuary in Puerto España, AGB values range from 121 to 211 Mg/ha. Finally, in the department of Chocó, AGB values ranging from 89 to 292 Mg/ha are reported for the El Litoral del San Juan municipalities, Bajo Baudó, Nuquí, and Jurado.

An uncertainty map was developed to estimate mangrove AGB in the CPC. The result was exported in GeoTIFF format and overlaid for each department of the CPC (Fig. 9). High uncertainty values were found in the southern area of Tumaco Bay in Nariño, as well as in some municipalities of Cauca and on Partera Island (Cauca). Uncertainty values gradually decrease in the department of Valle del Cauca until reaching the southern part of Buenaventura District Bay. Buenaventura Bay and Malaga Bay present a mix of low and high uncertainties. In the department of Chocó, a heterogeneous mixture with a tendency towards low uncertainty values was found in some municipalities (see Fig. 10).

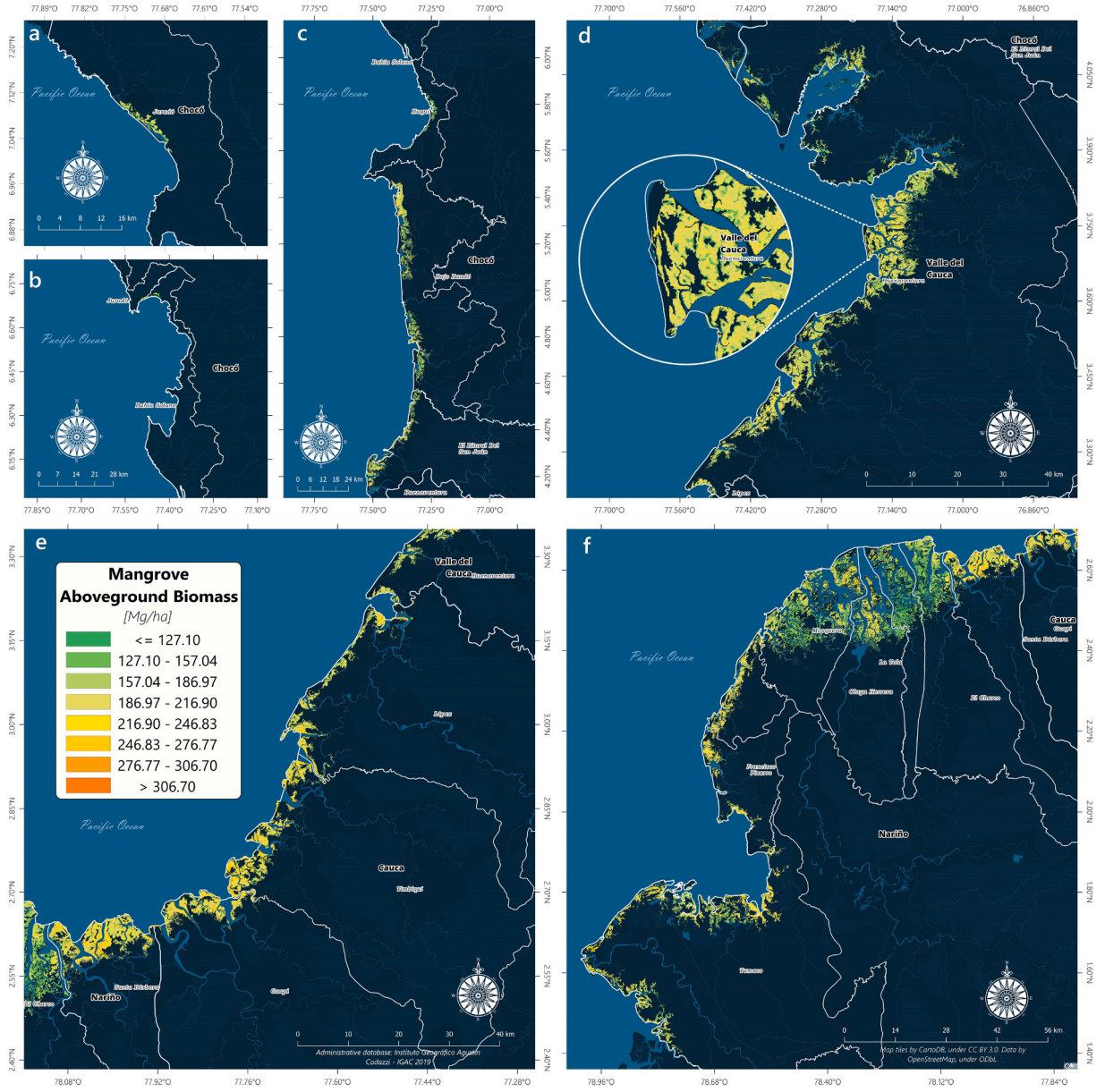

Fig. 8. a, b y c Chocó; d, Valle del Cauca; e, Cauca; y f, Nariño.
Predicted map of AGB [Mg/ha] in the study site derived from biomass RF model.

Table 5
RF estimation statistics for AGB [Mg/ha] in CPC mangroves.

Site	Mean ± stdev	Median	Variance	Minimum	Maximum
CPC	181.236 ± 28.939	185.007	837.457	88.622	378.21
Chocó	178.241 ± 25.799	180.501	665.603	88.622	292.926
Valle del Cauca	182.220 ± 23.045	184.495	531.085	95.219	294.801
Cauca	199.165 ± 25.257	197.502	637.927	104.625	340.118
Nariño	178.143 ± 31.657	180.998	1002.156	91.938	378.21

5. Discussion

The findings of this research for the AGB of the plots are above the pantropical average of 139.50 Mg/ha and the national average of

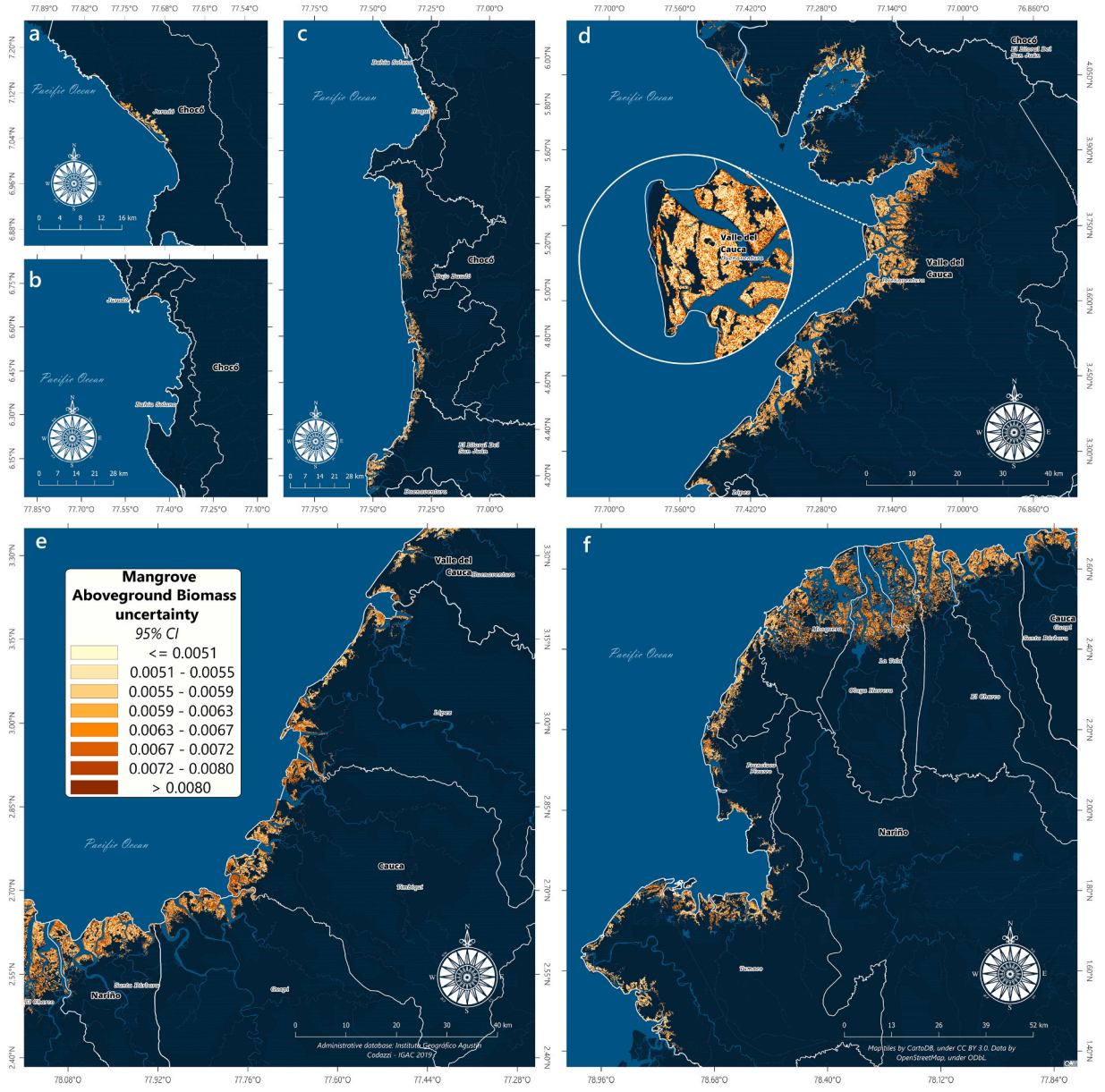

Fig. 9. a, b y c Chocó; d, Valle del Cauca; e, Cauca; y f, Nariño.
Uncertainty map of above-ground biomass in the study site derived from biomass RF model.

102.93 Mg/ha reported by Refs. [11,67], respectively. Also, the different AGB values found in the mangrove plots of the CPC may be caused by environmental and ecological factors like water salinity, soil depth, substrate type, sediment regime, wave exposure, presence of rivers or estuaries, temperature, humidity, and rainfall patterns, among others, which affect and shape the structure, shape, composition, and health of mangrove tree species [68].

Likewise, this variability of AGB in the mangrove plots of the CPC is a sample of these ecosystem's complexity and spatial heterogeneity. It underlines the need to obtain a representative sample of plots to obtain accurate estimates of AGB in the CPC. However, a significant limitation found in this study, as pointed out by Ref. [69], is that for Colombia, there is a scarcity of field studies in mangrove forests that makes spatial estimates difficult; in addition, sampling in mangroves is complex due to the particular conditions (Fig. 8) of this type of ecosystem, (e.g., tidal regime, muddy flats, root structure) and the costs associated with sampling large areas [70]. Public order problems and conflicts in the CPC [5] also limited sample collection in certain mangrove areas. Therefore, to minimize this obstacle in this research, combining data collected in the field with data from secondary studies was used to have the most extensive possible spatial coverage of AGB plots in the CPC to train and validate the information obtained by remote sensing.

[71], mention that the spatial autocorrelation assessment allows for identifying statistically significant clustering patterns of high

Fig. 10. Samples of the plots in the departments of Nariño and Valle del Cauca.

and low values. Moran's test was used to determine whether observations from AGB plots closer together are more similar than observations farther apart. The results found in this study in the Monte-Carlo simulations for the Moran's I statistic are very close to zero, and the P-value>0.05, suggesting that the alternative hypothesis of the existence of significant spatial autocorrelation should be rejected. However, it is important to clarify that this does not necessarily mean no spatial autocorrelation in the data. Although the hypothesis test could not find any evidence of spatial autocorrelation, it may exist. In addition, the Monte-Carlo simulation increases the confidence in the results of the hypothesis testing [72].

The inferences drawn in this analysis for the RL models indicate that the residuals of the linear models were not normally distributed (Table SM1.). In addition, two of the SRTM and NASADEM models found evidence of autocorrelation in the residuals. In all three cases of RL, the R^2 is low, indicating that models based on digital height alone explain little of the AGB variability in the CPC. The RMSE and MAE are also high, indicating that the model predictions have a high average error. This may be partly due to the mangrove AGB in the CPC's high complexity and variability, making it difficult to analyze using traditional parametric methods in this situation. [33], points out that nonparametric approaches may become more effective in terms of accuracy and results by better capturing complex and nonlinear relationships between variables and their spatial and temporal variation.

The RF model performed the best in making predictions and outperformed the other models evaluated, where parametric and nonparametric methods were used to retrieve the amount of AGB in the mangroves of the CPC (Table 4). This result suggests that spatializing mangrove AGB from multisensor data, together with the machine learning approach, may eventually improve the accuracy of spatial AGB estimation. This inference is supported by the studies of [8,10,73], which suggest that the use of nonparametric machine learning algorithms such as RF and GTB can establish complex nonlinear relationships between information from remote sensing in a flexible manner (altimetric, radar, and optical data) with national AGB inventory data. Furthermore, these studies also indicate that these models can provide good results in the estimation of AGB for mangroves even when the data have an indeterminate distribution.

The result obtained for the mangrove AGB of the CPC in this study, with a mean ± stdev of AGB of 181.236 ± 28.939 [Mg/ha], is above the mean value reported for the Colombian Pacific of 165.47 Mg/ha, which was calculated from altimetric data such as SRTM and LiDAR data from ICESAT/Glas and in situ measurements to create global AGB maps for mangroves [1]. Furthermore, it falls below the mean value reported for mangroves of 229.91 ± 89.7 Mg/ha of AGB in the tropical rainforests of Chocó, Colombia, where the authors employed in situ inventory data and LiDAR information along with an RF machine learning model to convert measurements into aerial biomass for the entire region from the coast to the highest elevations of the Andean forests [18].

The modeling of AGB values in the study area was highly heterogeneous, although regional patterns of low and high clusters were

observed in different sectors for the four departments. In northern Nariño and most of Cauca, for example, high AGB values were found in the municipalities of Santa Bárbara, Guapi, Timbiquí, and López de Micay. This is consistent with the analyses conducted by Ref. [74], and [1], which indicate those locations with exceptionally high mangrove AGB presence. However, it's worth noting that for these locations, the training samples [27,28,32] that informed the selected model span different dates and places over roughly a 12-year period (Table 1 and Table SM2). This demonstrates that both older and more recent data exhibit notably high AGB values, which is evident in the prediction model. In Valle del Cauca, a heterogeneous distribution was observed, but the highest values were found in the area from Isla el Ají to the south of Buenaventura Bay. In Chocó, values ranging from 89 to 292 Mg/ha were reported. These regional heterogeneous and clustering patterns could be related to different environmental factors, such as soil type, hydrology, or topography, suggesting, as one would expect, that there are significant differences in the mangrove ecosystems of the CPC [68,75].

Radar-optical-biomass models produce spatial uncertainty associated with AGB estimation. In this regard, each CPC department has developed an uncertainty map (Fig. 7). The analyses indicate that high uncertainty values are found in the southern area of Tumaco Bay in Nariño and some municipalities of Cauca. This is in line with the research of [18], for mangrove AGB of the CPC, where the existence of higher uncertainty associated with high AGB values is mentioned. On the other hand, it is observed that uncertainty values gradually decrease in the department of Valle del Cauca until reaching the southern part of the Buenaventura District Bay. In other words, the AGB estimates for mangroves in these areas are more accurate due to the greater availability of information from plots. In the department of Chocó, a heterogeneous mix with a tendency to low uncertainty values was found in some municipalities. The development of an uncertainty map for mangrove AGB estimation in the CPC is a valuable tool to identify areas where the precision of the estimation is lower; therefore, more research and data collection efforts are required, this will improve the precision of mangrove AGB estimates.

The present study is novel in the use of multi-sensor data and machine learning approaches to estimate aboveground biomass (AGB) in mangroves. This method offers several advantages, such as increased accuracy, cost-effectiveness, and reduced fieldwork. By integrating information from multiple sensors, we were able to assess and spatialize the variability in the AGB of different mangrove species. In addition, machine learning algorithms allowed us to model complex relationships between variables and improve the accuracy of predictions. This new approach for the study area contributes to the estimation and management of mangrove AGB, providing more reliable and efficient methods for the monitoring and conservation of these important ecosystems.

We faced several obstacles during the research, such as limited data availability and difficulties in sampling mangroves. To address these challenges, we recommend that future research efforts focus on improving data collection and integration. Remote sensing and exploration of drone technology can bring added benefits. In addition, we suggest developing more comprehensive and integrated models that consider the multiple structural factors that influence AGB, such as the normal diameter and density of mangrove species present. Addressing these challenges and improving methods for estimating AGB will help us better understand and manage mangroves, which play a crucial role in mitigating climate change, protecting coastal communities from natural disasters and sustaining local livelihoods. Our study contributes to this broader context by providing accurate and reliable estimates of mangrove AGB, which can contribute to policy and management decisions to conserve these valuable ecosystems.

6. Conclusions

The spatial autocorrelation analysis of this study found that the AGB information on the plots used did not provide sufficient support to claim that the presence of autocorrelation is significant in the study area. The mapping associated with this research exposed the heterogeneous variations present in mangrove AGB for the four departments of the CPC. Our results suggest that the methodology for developing mangrove AGB mapping from the multi-sensor combination (active and passive sensors) together with machine learning using the GEE platform, improves the accuracy of AGB estimation for mangroves compared to traditional parametric models. Accordingly, the RF model was the most accurate. This approach, employed for the first time for Colombia in this study, has already been implemented in other places such as China [8,10] and the Philippines [73] for different vegetation cover types from mangroves to subtropical forests.

The uncertainty of mangrove AGB was relatively low suggesting that the approach is an interesting tool that can be modified and applied in other studies for mangroves elsewhere. The mapping generated in this study is not only an important tool to inform decision-making and the strengthening of different use, management, and conservation strategies, but also generates updated information that allows for further analysis of different types of factors influencing mangrove dynamics in the region.

This study has demonstrated the potential of using multi-sensor data and machine learning approaches to estimate aboveground biomass (AGB) in mangroves. However, it is important to recognize the limitations encountered that introduce uncertainty, such as the lack of high-resolution data and difficulties in sampling mangroves. To address these limitations, future research could focus on improving plot sampling methods and exploring alternative machine learning techniques. The results of this study have important implications for mangrove conservation and management efforts in the study area. Given the critical role mangroves play in climate change mitigation, coastal protection, and the livelihoods of local communities, it is crucial that researchers, policy makers and practitioners prioritize mangrove conservation. We call on all stakeholders to use the results of this study to inform and improve AGB mangrove mapping and management efforts, and to work towards preserving these important ecosystems for future generations.

Author contributions

John J. Selvaraj: Conceived and designed the analysis; Wrote the paper. Bryan E. Gallego: Conceived and designed the analysis; Performed the analysis; Wrote the paper.

Funding statement

This work was financed by: General System Royalties in Colombia (FCTel-SGR): BPIN:2020000100054 and Universidad Nacional de Colombia HERMES: 45980.

Declaration of competing interest

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

Acknowledgements

Data for this study were sourced from NASA, USGS, JAXA, and GEE. Special acknowledgment to the Colombian National Natural Park for Collection Authorization No. 013 of 2021 was pivotal for our fieldwork. We appreciate the contributions of Professor Devin Routh from the University of Zurich and researcher Zulmary Valoyes Cardozo from the Instituto De Investigaciones Ambientales Del Pacífico "John Von Neumann." We are grateful to the students (Ms. Kelly Tatiana G.T., Mr. Cesar Augusto M.F., Ms. Sofia Vargas D., Ms. Laura Lozano A., and Ms. Adriana Martinez A.) as well as Dr. Angela I. Guzman, faculty member at Universidad Nacional de Colombia, and the experienced guides who supported our extensive data collection efforts. Lastly, we thank the reviewers for their invaluable feedback and timely suggestions, which significantly improved our manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e20745.

References

[1] M. Simard, L. Fatoyinbo, C. Smetanka, V.H. Rivera-Monroy, E. Castañeda-Moya, N. Thomas, T. Van der Stocken, Mangrove canopy height globally related to precipitation, temperature and cyclone frequency, Nat. Geosci. 12 (2019) 40–45, https://doi.org/10.1038/s41561-018-0279-1.
[2] D.M. Alongi, S.K. Mukhopadhyay, Contribution of mangroves to coastal carbon cycling in low latitude seas, Agric. For. Meteorol. 213 (2015) 266–272, https://doi.org/10.1016/j.agrformet.2014.10.005.
[3] T.D. Pham, K. Yoshino, D.T. Bui, Biomass estimation of Sonneratia caseolaris (l.) Engler at a coastal area of Hai Phong city (Vietnam) using ALOS-2 PALSAR imagery and GIS-based multi-layer perceptron neural networks, GIScience Remote Sens. 54 (2017) 329–353, https://doi.org/10.1080/15481603.2016.1269869.
[4] J.D.T. De Alban, J. Jamaludin, D. Wong De Wen, M.M. Than, E.L. Webb, Improved estimates of mangrove cover and change reveal catastrophic deforestation in Myanmar, Environ. Res. Lett. 15 (2020), 34034, https://doi.org/10.1088/1748-9326/ab666d.
[5] A.M. Rojas, C.A. Ruiz–Agudelo, M.C. Diazgranados, H. Polanco, R. Anderson, Approach to an integral valuation of mangrove's ecosystem services in a marine protected area. Colombian Pacific region, Journal of Environmental Economics and Policy 8 (2019) 322–342, https://doi.org/10.1080/21606544.2019.1584127.
[6] C.J. Wong, D. James, N.A. Besar, K.U. Kamlun, J. Tangah, S. Tsuyuki, M.H. Phua, Estimating Mangrove above-ground biomass loss due to deforestation in Malaysian Northern Borneo between 2000 and 2015 using SRTM and landsat images, Forests 11 (2020), https://doi.org/10.3390/F11091018.
[7] S. Abbas, M.S. Wong, J. Wu, N. Shahzad, S. Muhammad Irteza, Approaches of satellite remote sensing for the assessment of above-ground biomass across tropical forests: pan-tropical to national scales, Rem. Sens. 12 (2020) 3351, https://doi.org/10.3390/rs12203351.
[8] H. Su, W. Shen, J. Wang, A. Ali, M. Li, Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests, For. Ecosyst. 7 (2020) 64, https://doi.org/10.1186/s40663-020-00276-7.
[9] A. Aslan, M.O. Aljahdali, Characterizing global patterns of mangrove canopy height and aboveground biomass derived from SRTM data, Forests 13 (2022) 1545, https://doi.org/10.3390/f13101545.
[10] Y. Li, M. Li, C. Li, Z. Liu, Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms, Sci. Rep. 10 (2020) 9952, https://doi.org/10.1038/s41598-020-67024-3.
[11] P. Saenger, S.C. Snedaker, Pantropical trends in mangrove above-ground biomass and annual litterfall, Oecologia 96 (1993) 293–299, https://doi.org/10.1007/BF00317496.
[12] D. Lagomasino, T. Fatoyinbo, S. Lee, E. Feliciano, C. Trettin, M. Simard, A comparison of mangrove canopy height using multiple independent measurements from land, air, and space, Rem. Sens. 8 (2016) 327, https://doi.org/10.3390/rs8040327.
[13] A. Novo-Fernández, M. Barrio-Anta, C. Recondo, A. Cámara-Obregón, C.A. López-Sánchez, Integration of national forest inventory and nationwide airborne laser scanning data to improve forest yield predictions in north-western Spain, Rem. Sens. 11 (2019) 1693, https://doi.org/10.3390/rs11141693.
[14] G.A. Castellanos-Galindo, L.C. Kluger, M.A. Camargo, J. Cantera, J.E. Mancera Pineda, J.F. Blanco-Libreros, M. Wolff, Mangrove research in Colombia: temporal trends, geographical coverage and research gaps, Estuarine, Coastal and Shelf Science 248 (2021), 106799, https://doi.org/10.1016/j.ecss.2020.106799.
[15] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, R. Moore, Google earth engine: planetary-scale geospatial analysis for everyone, Rem. Sens. Environ. 202 (2017) 18–27, https://doi.org/10.1016/j.rse.2017.06.031.
[16] B. Posada, W. Henao, G. Guzmán, Diagnóstico de la erosión y sedimentación en la zona costera del Pacífico colombiano, INVEMAR, INVEMAR, vol. 17, Serie Publicaciones Especiales No., Santa Marta, 2009.
[17] B.E. Gallego Perez, J.J. Selvaraj, Evaluation of Coastal Vulnerability for the District of Buenaventura, Colombia: A Geospatial Approach, Remote Sensing Applications, Society and Environment, 2019, p. 16, https://doi.org/10.1016/j.rsase.2019.100263.
[18] V. Meyer, S. Saatchi, A. Ferraz, L. Xu, A. Duque, M. García, J. Chave, Forest degradation and biomass loss along the Chocó region of Colombia, Carbon Bal. Manag. 14 (2019) 1–15, https://doi.org/10.1186/s13021-019-0117-9.
[19] J.J. Selvaraj, B.E. Gallego Pérez, An enhanced approach to mangrove forest analysis in the Colombian Pacific coast using optical and SAR data in Google Earth Engine, Remote Sens. Appl.: Society and Environment 30 (2023), 100938, https://doi.org/10.1016/j.rsase.2023.100938.
[20] J.B. Kauffman, D. Donato, Protocols for the Measurement, Monitoring and Reporting of Structure, Biomass and Carbon Stocks in Mangrove Forests, Center for International Forestry Research (CIFOR), 2012, https://doi.org/10.17528/cifor/003749.

[21] L. Yang, S. Liang, Y. Zhang, A new method for generating a global forest aboveground biomass map from multiple high-level satellite products and ancillary information, IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 13 (2020) 2587–2597, https://doi.org/10.1109/JSTARS.2020.2987951.
[22] T.E. Fatoyinbo, M. Simard, Height and biomass of mangroves in Africa from ICESat/GLAS and SRTM, Int. J. Rem. Sens. 34 (2013) 668–681, https://doi.org/10.1080/01431161.2012.712224.
[23] A. Canty, B.D. Ripley, Boot: Bootstrap R (S-Plus) Functions, 2022.
[24] J. Gross, U. Ligges, Nortest: Tests for Normality. R Package Version 1, 2015, .0-4. https://CRAN.R-project.org/package=nortest.
[25] R Core Team, R.A. Language, And Environment for Statistical Computing, 2023. https://www.R-project.org/.
[26] C.V.C. Invemar, Monitoreo de los manglares del Valle del Cauca y la fauna asociada con énfasis en las aves y especies de importancia económica como la piangua y el cangrejo azul, Instituto de Investigaciones marinas y costeras José Benito Vives de Andreis, Buenaventura, Valle Del Cauca, 2007. http://koha.ideam.gov.co/cgi-bin/koha/opac-detail.pl?biblionumber=4471562.
[27] A. López Rodríguez, P.C. Sierra-Correa, M. Rodríguez Peláez, J.C. Hernández Ortiz, C. Muñoz, C. Satizabal, J. Zamudio, G. Almario, J. Bolaños, L.M. Prieto, Ordenamiento ambiental de los manglares del municipio de López de Micay, departamento del Cauca (Pacífico colombiano), INVEMAR No 34, Bogotá D.C., Colombia, 2009.
[28] P.C. Sierra-Correa, A. Sánchez, A. López Rodríguez, J.C. Rodríguez Peláez, C. Muñoz, C. Satizabal, A. Moreno, G. Almario, F. Bedoya, M. Hernández-Ortiz, J. Bolaños, L.M. Prieto, Ordenamiento ambiental de los manglares del municipio de Timbiquí, departamento del Cauca (Pacífico colombiano), Instituto de Investigaciones Marinas y Costeras - INVEMAR, 2009. http://hdl.handle.net/1834/6644.
[29] A. Monsalve, G. Ramírez, Caracterización de la estructura y contenido de carbono de los bosques de manglar en el área de jurisdicción del Consejo comunitario La Plata, Bahía Málaga, Valle del Cauca, Medellín, Colombia, 2015.
[30] IIAP, Data Not yet Published), Instituto De Investigaciones Ambientales Del Pacífico "John Von Neumann, Chocó-Colombia, 2016. https://iiap.org.co/.
[31] IIAP, Caracterización Ecológica De Los Manglares De Bahía Cupica, Bahía Solano (Chocó), Instituto De Investigaciones Ambientales Del Pacífico "John Von Neumann, Chocó-Colombia, 2017. https://iiap.org.co/.
[32] G.A. Castellanos-Galindo, E. Casella, H. Tavera, L.A. Zapata Padilla, M. Simard, Structural characteristics of the tallest mangrove forests of the American continent: a comparison of ground-based, drone and radar measurements, Frontiers in Forests and Global Change 4 (2021) 1–11, https://doi.org/10.3389/ffgc.2021.732468.
[33] S. Saatchi, SAR Methods for Mapping and Monitoring Forest Biomass, 2019, https://doi.org/10.25966/hbm1-ej07.
[34] R.S. Bivand, D.W.S. Wong, Comparing implementations of global and local indicators of spatial association, Test 27 (2018) 716–748, https://doi.org/10.1007/s11749-018-0599-x.
[35] GEE, Earth Engine Data Catalog, Google Developers, 2022. https://developers.google.com/earth-engine/datasets/catalog. (Accessed 21 March 2023).
[36] M. Shimada, On the ALOS/PALSAR operational and interferometric aspects, J. Geodetic Soc. Jpn. 56 (2010) 13–39.
[37] M. Shimada, T. Itoh, T. Motooka, M. Watanabe, T. Shiraishi, R. Thapa, R. Lucas, New global forest/non-forest maps from ALOS PALSAR data (2007-2010), Rem. Sens. Environ. 155 (2014) 13–31, https://doi.org/10.1016/j.rse.2014.04.014.
[38] T.G. Farr, P.A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, D. Alsdorf, The Shuttle radar topography mission, Rev. Geophys. 45 (2007) RG2004, https://doi.org/10.1029/2005RG000183.
[39] S.M. Buckley, P.S. Agram, J.E. Belz, R.E. Crippen, E.M. Gurrola, NASADEM, National Aeronautics and Space Administration (NASA), 2020. https://lpdaac.usgs.gov/documents/592/NASADEM_User_Guide_V1.pdf.
[40] T. Tadono, H. Nagai, H. Ishida, F. Oda, S. Naito, K. Minakawa, H. Iwamoto, Generation of the 30 M-mesh global digital surface model by alos prism, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLI- B4 (2016) 157–162, https://doi.org/10.5194/isprsarchives-XLI-B4-157-2016.
[41] O. Mutanga, L. Kumar, Google earth engine applications, Rem. Sens. 11 (2019) 11–14, https://doi.org/10.3390/rs11050591.
[42] E. Vermote, C. Justice, M. Claverie, B. Franch, Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product, Rem. Sens. Environ. 185 (2016) 46–56, https://doi.org/10.1016/j.rse.2016.04.008.
[43] G. Lemoine, Refined Lee GEE Implementation, 2018. https://code.earthengine.google.com/5d1ed0a0f0417f098fdfd2fa137c3d0c. (Accessed 9 March 2022).
[44] JAXA, Global 25 M Resolution PALSAR-2 Mosaic (Ver.2.1.0) Dataset Description, Japan Aerospace Exploration Agency (JAXA) Earth Observation Research Center (EORC), Japan Aerospace Exploration Agency, 2022.
[45] GEE, ee.Image.Expression, Google Developers, 2022. https://developers.google.com/earth-engine/apidocs/ee-image-expression. (Accessed 22 March 2023).
[46] D. Datta, M. Dey, S. Neogy, T. Basu Roy, D. Dutta, A. Kundu, G. Nandi, Assessing vegetation fragmentation and plantation efficiency in an intertidal mudflat of Eastern India using Radar Forest Degradation Index and spatial metrics, Geocarto Int. (2021) 1–22, https://doi.org/10.1080/10106049.2021.2017014.
[47] USGS, Landsat 8-9 Collection 2 (C2) Level 2 Science Product (L2SP) Guide, U.S. Geological Survey, Sioux Falls, South Dakota, 2022. https://www.usgs.gov/media/files/landsat-8-9-collection-2-level-2-science-product-guide.
[48] T.N. Phan, V. Kuch, L.W. Lehnert, Land cover classification using google earth engine and random forest classifier—the role of image composition, Rem. Sens. 12 (2020) 2411, https://doi.org/10.3390/rs12152411.
[49] GEE, ee.ImageCollection.Median, 2022. https://developers.google.com/earth-engine/apidocs/ee-imagecollection-median. (Accessed 22 March 2023).
[50] C. Diniz, L. Cortinhas, G. Nerino, J. Rodrigues, L. Sadeck, M. Adami, P.W.M. Souza-Filho, Brazilian mangrove status: three decades of satellite data analysis, Rem. Sens. 11 (2019), https://doi.org/10.3390/rs11070808.
[51] A.B. Baloloy, A.C. Blanco, R.R.C. Raymund Rhommel, K. Nadaoka, Development and application of a new mangrove vegetation index (MVI) for rapid and accurate mangrove mapping, ISPRS J. Photogrammetry Remote Sens. 166 (2020) 95–117, https://doi.org/10.1016/j.isprsjprs.2020.06.001.
[52] G. Winarso, A.D. Purwanto, D. Yuwono, New mangrove index as degradation/health indicator using remote sensing data: segara anakan and alas purwo case study, 12th Biennial Conference of Pan Ocean Remote Sensing Conference (2014) (2009) 309–316.
[53] T. Shi, J. Liu, Z. Hu, H. Liu, J. Wang, G. Wu, New spectral metrics for mangrove forest identification, Remote Sensing Letters 7 (2016) 885–894, https://doi.org/10.1080/2150704X.2016.1195935.
[54] K. Gupta, A. Mukhopadhyay, S. Giri, A. Chanda, S. Datta Majumdar, S. Samanta, D. Mitra, R.N. Samal, A.K. Pattnaik, S. Hazra, An index for discrimination of mangroves from non-mangroves using LANDSAT 8 OLI imagery, MethodsX 5 (2018) 1129–1139, https://doi.org/10.1016/j.mex.2018.09.011.
[55] Y. Fan, X. Ding, J. Wu, J. Ge, Y. Li, High Spatial-Resolution Classification of Urban Surfaces Using a Deep Learning Method, Building and Environment, vol. 200, 2021, 107949, https://doi.org/10.1016/j.buildenv.2021.107949.
[56] Y. Zhou, J. Xue, S. Chen, Y. Zhou, Z. Liang, N. Wang, Z. Shi, Fine-resolution mapping of soil total nitrogen across China based on weighted model averaging, Rem. Sens. 12 (2019) 85, https://doi.org/10.3390/rs12010085.
[57] GEE, ee.Image.sampleRegions, Google Developers-Google Earth Engine, 2022. https://developers.google.com/earth-engine/apidocs/ee-image-sampleregions. (Accessed 2 July 2022).
[58] A. Zeileis, T. Hothorn, Diagnostic checking in regression relationships, R. News 2 (2002) 7–10.
[59] T.D. Pham, N. Yokoya, J. Xia, N.T. Ha, N.N. Le, T.T.T. Nguyen, T.H. Dao, T.T.P. Vu, T.D. Pham, W. Takeuchi, 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, Rem. Sens. 12 (2020) 1–24, https://doi.org/10.3390/RS12081334.
[60] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[61] S. Praticò, F. Solano, S. Di Fazio, G. Modica, Machine learning classification of mediterranean forest habitats in google earth engine based on seasonal sentinel-2 time-series and input image composition optimisation, Rem. Sens. 13 (2021) 1–28, https://doi.org/10.3390/rs13040586.
[62] J.H. Friedman, Stochastic gradient boosting, Comput. Stat. Data Anal. 38 (2002) 367–378, https://doi.org/10.1016/S0167-9473(01)00065-2.
[63] Y. Yang, D. Yang, X. Wang, Z. Zhang, Z. Nawaz, Testing accuracy of land cover classification algorithms in the qilian mountains based on gee cloud platform, Rem. Sens. 13 (2021), https://doi.org/10.3390/rs13245064.

[64] B. Bischl, M. Binder, M. Lang, T. Pielok, J. Richter, S. Coors, J. Thomas, T. Ullmann, M. Becker, A. Boulesteix, D. Deng, M. Lindauer, Hyperparameter optimization: foundations, algorithms, best practices, and open challenges, WIREs Data Min & Knowl 13 (2023), https://doi.org/10.1002/widm.1484.
[65] B. Kempen, D.J. Brus, G.B.M. Heuvelink, Validation, in: Soil Organic Carbon Mapping Cookbook, Food and Agriculture Organization of the United Nations, Rome, 2018, pp. 109–131. https://fao-gsp.github.io/SOC-Mapping-Cookbook/index.html.
[66] D. Routh. Implementation of k-fold cross-validation for google earth engine, Google Developers-Google Earth Engine, 2017. https://code.earthengine.google.com/e0bd2992c08ec1a26c31a9fdea5cfc59 (Accessed 20 February 2023).
[67] J.M. Bolivar, V.H. Gutierrez-Velez, C.A. Sierra, Carbon stocks in aboveground biomass for Colombian mangroves with associated uncertainties, Regional Studies in Marine Science 18 (2018) 145–155, https://doi.org/10.1016/j.rsma.2017.12.011.
[68] T.A. Worthington, P.S.E. zu Ermgassen, D.A. Friess, K.W. Krauss, C.E. Lovelock, J. Thorley, R. Tingey, C.D. Woodroffe, P. Bunting, N. Cormier, D. Lagomasino, R. Lucas, N.J. Murray, W.J. Sutherland, M. Spalding, A global biophysical typology of mangroves and its relevance for ecosystem structure and deforestation, Sci. Rep. 10 (2020), 14652, https://doi.org/10.1038/s41598-020-71194-5.
[69] J.F. Blanco-Libreros, S.R. López-Rodríguez, A.M. Valencia-Palacios, G.F. Perez-Vega, R. Álvarez-León, Mangroves from rainy to desert climates: baseline data to assess future changes and drivers in Colombia, Front. For. Glob. Change. 5 (2022), 772271, https://doi.org/10.3389/ffgc.2022.772271.
[70] R.P. Parman, N. Kamarudin, F.H. Ibrahim, A.A. Nuruddin, H. Omar, Z. Abdul Wahab, Geostatistical analysis of mangrove ecosystem health: mapping and modelling of sampling uncertainty using kriging, Forests 13 (2022) 1185, https://doi.org/10.3390/f13081185.
[71] J. Jossart, S.J. Theuerkauf, L.C. Wickliffe, J.A. Morris Jr., Applications of spatial autocorrelation analyses for marine aquaculture siting, Front. Mar. Sci. 6 (2020) 806, https://doi.org/10.3389/fmars.2019.00806.
[72] Manny Gimond, A Basic Introduction to Moran's I Analysis in R, 2019. https://mgimond.github.io/simple_moransI_example/. (Accessed 11 April 2023).
[73] J.A.A. Castillo, A.A. Apan, T.N. Maraseni, S.G. Salmo, Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the Philippines using Sentinel imagery, ISPRS J. Photogrammetry Remote Sens. 134 (2017) 70–85, https://doi.org/10.1016/j.isprsjprs.2017.10.016.
[74] A.S. Rovai, P. Riul, R.R. Twilley, E. Castañeda-Moya, V.H. Rivera-Monroy, A.A. Williams, M. Simard, M. Cifuentes-Jara, R.R. Lewis, S. Crooks, P.A. Horta, Y. Schaeffer-Novelli, G. Cintrón, M. Pozo-Cajas, P.R. Pagliosa, Scaling mangrove aboveground biomass from site-level to continental-scale, Global Ecol. Biogeogr. 25 (2016) 286–298, https://doi.org/10.1111/geb.12409.
[75] M.L. Palacios Peñaranda, J.R. Cantera Kintz, E.J. Peña Salamanca, Carbon stocks in mangrove forests of the Colombian Pacific, Estuarine, Coastal and Shelf Science 227 (2019), 106299, https://doi.org/10.1016/j.ecss.2019.106299.