Remote Sensing Applications: Society and Environment 42 (2026) 102087

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

journal homepage: www.elsevier.com/locate/rsase

Mangrove extent and above-ground biomass mapping in Sarawak, Malaysia using Sentinel data and machine learning

Abdul Qaiyum Alidin [a,c], Azlizam Aziz [a], Waseem Razzaq Khan [b], Shazali Johari [a,c,*]

[a] Department of Recreation and Ecotourism, Universiti Putra Malaysia, Faculty of Forestry and Environment, Serdang, Selangor, Malaysia
[b] Department of Forestry Science and Biodiversity, Universiti Putra Malaysia UPM, Faculty of Forestry and Environment, Serdang, 43400, Selangor, Malaysia
[c] i-CATS University College, Jalan Stampin Timur, 93350, Kuching, Sarawak, Malaysia

ARTICLE INFO

Keywords:
Mangrove forests
Above-Ground Biomass (AGB)
Remote sensing
Machine learning

ABSTRACT

Mangrove forests are critical blue carbon ecosystems yet remain underrepresented in national carbon accounting frameworks due to persistent data gaps at fine spatial scales. This study presents a regionally focused, operationally reproducible remote sensing and machine learning framework to map mangrove extent and estimate above-ground biomass (AGB) across western and central Sarawak, Malaysia. Open-access Sentinel-1 C-band SAR and Sentinel-2 multispectral data were fused with extensive field inventory data from 245 plots to develop classification and regression models using the eXtreme Gradient Boosting (XGBoost) algorithm. While L-band SAR offers known advantages for high-biomass estimation, such datasets were not freely or consistently available for the full temporal and spatial coverage required; the chosen C-band–optical fusion maximises accessibility, repeatability, and integration into regional monitoring workflows. The mangrove extent map, generated at 10 m resolution, achieved an overall accuracy of 88.8%, a Cohen's Kappa of 0.776, and an AUC of 0.960. The predicted mangrove area totalled 1078.7 km^2, with 83.8% exhibiting very low classification uncertainty. For AGB, model performance was quantified using held-out predictions to avoid optimistic bias; out-of-fold cross-validated predictions across all plots ($n = 245$) yielded $R^2 = 0.72$, RMSE = 28.59 Mg ha^{-1}, MAE = 20.80 Mg ha^{-1}, and Pearson's $r = 0.85$. Monte Carlo simulations ($n = 100$) confirmed the model's stability, yielding RMSE = 26.25 ± 1.33 Mg ha^{-1} and $R^2 = 0.783 \pm 0.023$ (evaluated on held-out predictions within each iteration). Species-level AGB analysis revealed the highest mean biomass in *Rhizophora mucronata* (102.1 Mg ha^{-1}) and the lowest in *Bruguiera parviflora* (74.5 Mg ha^{-1}). SHAP interpretation identified canopy height, NDVI, VH backscatter, and elevation as the most influential predictors. This accessible, fused, and interpretable modelling framework provides a scalable solution for national carbon inventories, including annual emission-factor refinement and spatial change-detection grids, while supporting targeted mangrove conservation and restoration under REDD+ and other climate initiatives.

* Corresponding author. Department of Recreation and Ecotourism, Universiti Putra Malaysia, Faculty of Forestry and Environment, Serdang, Selangor, Malaysia.
E-mail address: zali_johari@upm.edu.my (S. Johari).

https://doi.org/10.1016/j.rsase.2026.102087
Received 30 September 2025; Received in revised form 2 May 2026; Accepted 26 May 2026
Available online 27 May 2026
2352-9385/

1. Introduction

Mangrove forests are among the most ecologically and economically valuable coastal ecosystems, providing shoreline protection, sediment retention, water purification, habitat provision, and substantial carbon storage capacity (Sahputra et al., 2022; van Hespen et al., 2023; Khan et al., 2024). In Southeast Asia, mangroves are particularly extensive and biologically diverse, representing the largest regional share of global mangrove cover (Arjasakusuma et al., 2021; Ihinegbu et al., 2023). Malaysia, and Sarawak specifically, hosts some of the most structurally complex and carbon-rich mangrove systems in the region, making it a priority area for conservation and carbon monitoring (Abdul Aziz et al., 2015; Wolswijk et al., 2020). This regional importance is directly linked to blue carbon accounting because mangrove above-ground biomass (AGB), together with below-ground and soil carbon pools, contributes substantially to coastal carbon stocks and climate-mitigation planning (Aye et al., 2022; Monga et al., 2022). Mangroves across Southeast Asia continue to face pressure from aquaculture expansion, infrastructure development, and land-use conversion (Malik et al., 2022; Pesiu et al., 2022). In some regions, mangrove loss has exceeded 30% over recent decades, reducing coastal resilience and releasing previously stored carbon (Richards et al., 2020). Sarawak's mangrove belts remain relatively extensive, but increasing development pressure highlights the need for systematic monitoring to support land-use policy, conservation planning, and carbon reporting (Pandey et al., 2024). Traditional mangrove inventory remains valuable at plot scale, yet it is constrained by labour requirements, limited spatial coverage, and difficult access in wet intertidal environments (Chien et al., 2024; Sheehy et al., 2024; Soeprobowati et al., 2024). Remote sensing therefore offers a scalable route for repeated mangrove extent and AGB assessment across large coastal landscapes (Liu et al., 2024; Rahman et al., 2024).

Advances in remote sensing, particularly freely available Sentinel-1 synthetic aperture radar (SAR) and Sentinel-2 optical data, have improved mangrove monitoring by providing high spatial resolution and frequent revisit intervals (Li et al., 2020; Mehmood et al., 2024a). Sentinel-2 imagery provides spectral information for vegetation indices sensitive to canopy greenness, moisture condition, and biochemical properties (Anees et al., 2025a, 2025b). Optical observations alone, however, are often limited by persistent cloud cover in tropical regions. Sentinel-1 C-band SAR complements optical imagery by providing day–night, all-weather observations that are sensitive to canopy structure, surface moisture, and vegetation roughness (Ghorbanian et al., 2021; Cherian and Rajitha, 2024; Ghosh and Behera, 2021). When used together, Sentinel-1's structural sensitivity and Sentinel-2's spectral richness can improve mangrove extent mapping and biomass estimation over single-sensor approaches, with machine-learning models such as Random Forest and XGBoost commonly reporting strong classification and AGB prediction performance (Barenblitt et al., 2024; Hu et al., 2020; Huang et al., 2022). Ghorbanian et al. (2021), for instance, achieved high mangrove classification accuracy using multi-temporal Sentinel-1 and Sentinel-2 data, while the Global Mangrove Watch project used ALOS PALSAR and optical time series to generate global mangrove extent products (Bunting et al., 2018, 2022). Longer-wavelength L-band SAR remains advantageous for high-biomass mangroves because of deeper canopy penetration and reduced saturation, but freely available and temporally consistent L-band coverage was not available for the full spatial and temporal requirements of this study. The present work therefore develops an operational Sentinel-1 C-band and Sentinel-2 fusion framework that prioritises open access, reproducibility, and practical integration into regional monitoring systems.

Multi-temporal and multi-source remote sensing data are increasingly processed through cloud-based platforms such as Google Earth Engine, enabling efficient handling of large image collections for mangrove change detection and annual monitoring (Mehmood et al., 2024b, 2024c). Such workflows support reporting under REDD+, the Sustainable Development Goals, and national conservation programmes by identifying mangrove extent, degradation patterns, and restoration priorities (Shahzad et al., 2025). Beyond extent mapping, AGB estimation benefits from combining dual-polarisation Sentinel-1 backscatter, VV/VH ratios, backscatter transformations, and GLCM texture metrics with Sentinel-2 red-edge, NIR, SWIR, vegetation, and moisture-sensitive indices (Arthur Endsley et al., 2020; Sainuddin et al., 2024; Liu et al., 2024). Optical indices such as NDVI and red-edge vegetation indices correlate with canopy vigour but may plateau in dense, high-biomass stands (Anees et al., 2025c, 2025d). SAR predictors, particularly dual-polarised C-band and longer-wavelength L-band observations, add structural and moisture-related information that improves the representation of mangrove stand variability (Gegiuc et al., 2018; Karvonen et al., 2022). Lucas et al. (2020) showed that integrating Sentinel-1, Sentinel-2, and ALOS-2 improved biomass estimation across tall and scrub mangroves. Building on this evidence, the present study focuses on the practical value of a fully open Sentinel-1/Sentinel-2 framework rather than dependence on less consistently available L-band or airborne LiDAR data. This distinction is important for Sarawak, where operational carbon monitoring requires datasets that can be repeatedly acquired, processed, and updated at regional scale.

Machine-learning methods are increasingly used to estimate AGB from remote-sensing predictors because they can model nonlinear relationships within high-dimensional datasets (Liu et al., 2021; Xiong et al., 2024; Muhammad et al., 2025). Ensemble models, including Random Forest and Gradient Boosting, have shown particular value in biomass modelling. XGBoost has demonstrated strong performance in mangrove environments, with Pham et al. (2020) showing that models trained on fused SAR-optical data outperformed alternative algorithms for AGB prediction. Its compatibility with feature-importance and SHAP-based interpretation also allows the ecological contribution of spectral, radar, texture, topographic, and structural predictors to be examined more transparently (Fan et al., 2024). In this study, XGBoost is not used only as a predictive algorithm; it is embedded within an interpretable modelling framework that combines C-band SAR, optical indices, texture features, ancillary variables, field-measured AGB, model uncertainty assessment, and SHAP-based explanation. This combination distinguishes the framework from Sentinel-based studies that focus mainly on classification, use limited field calibration, or report model accuracy without detailed predictor interpretation. High-resolution AGB maps generated through such integrated approaches are crucial for blue carbon accounting, monitoring, reporting, and verification under climate frameworks. In Sarawak, where mangroves overlap with expanding human development, fine-scale AGB information is particularly important for restoration planning, land-use decisions, and coastal resilience assessment.

Despite substantial progress in mangrove mapping and biomass estimation, important gaps remain in Malaysian Borneo. Sarawak contains extensive and structurally diverse mangrove systems, but local-scale AGB estimation using integrated Sentinel-1, Sentinel-2, field inventory, and machine learning remains limited. Many Southeast Asian studies focus on Indonesia, Vietnam, or Peninsular Malaysia, while global products may not adequately capture Sarawak's local variation in canopy structure, species composition, and biomass density. The contribution of this study lies in developing a regionally calibrated, open-data, and interpretable C-band–optical fusion framework for both mangrove extent mapping and AGB estimation in Sarawak. This study addresses the gap by: (1) producing a 10 m-resolution mangrove extent map for Sarawak using fused Sentinel-1 and Sentinel-2 data; (2) developing an XGBoost-based model to estimate AGB from multi-source satellite and ancillary predictors; and (3) evaluating model reliability through held-out prediction, uncertainty analysis, sensor-branch comparison, and SHAP-based interpretation. The results provide a practical baseline for blue carbon accounting and support evidence-based mangrove management in Sarawak.

2. Methods and materials

2.1. Study area

The study was conducted in the coastal and deltaic lowlands of western and central Sarawak, Malaysia, covering an area defined by the geographic extent of 109.54°E to 112.16°E longitude and 0.98°N to 2.86°N latitude. This region encompasses a network of estuarine and riverine landscapes where widespread mangrove forests are ecologically prominent (Fig. 1). Administratively, the area encompasses key coastal districts, including Kuching, Lundu, Bau, Samarahan, Asajaya, Simunjan, Pusa, Kabong, Sarikei, Sibu, Daro, Matu, Tanjung Manis, and Maradong. These districts span Sarawak's southern and central coastal zones and are recognized for their extensive mangrove belts, supported by sediment-rich river deltas and sheltered shorelines. (Khan et al., 2025). The terrain is

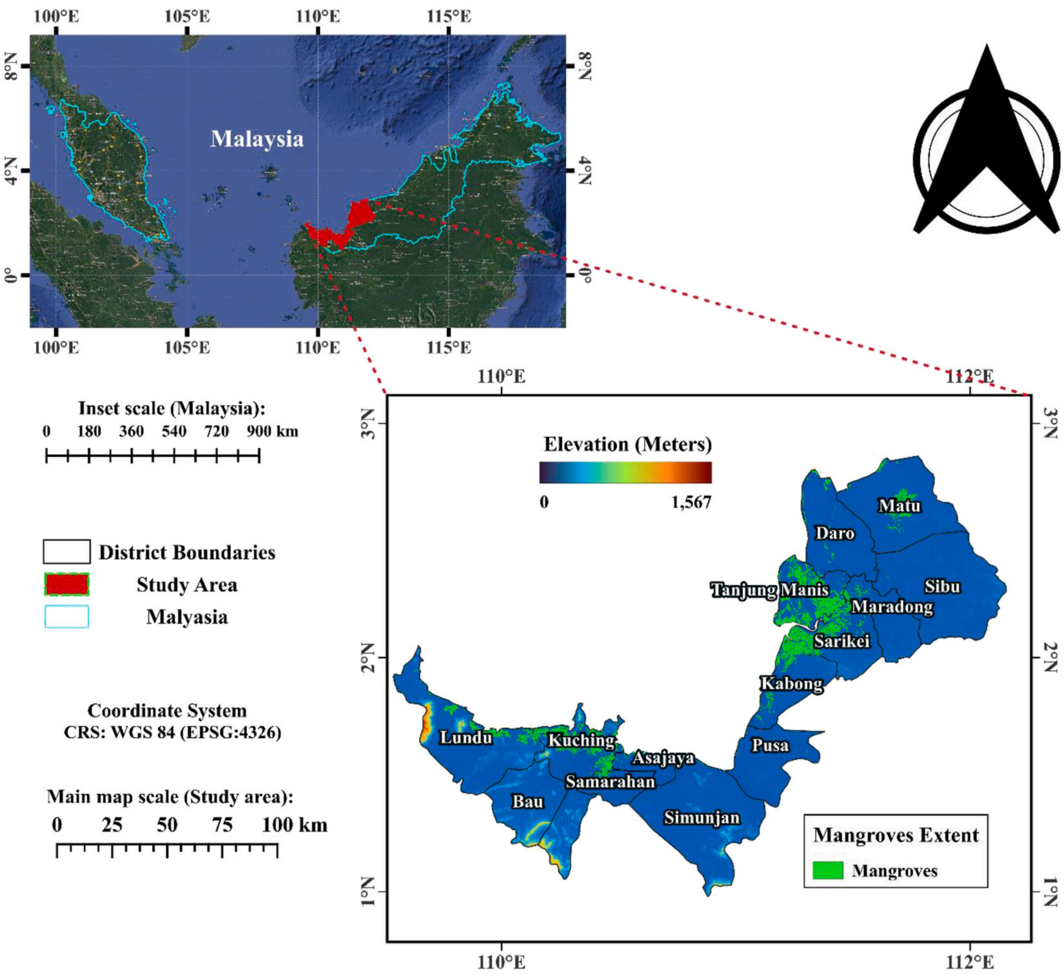

Fig. 1. Study area in western and central Sarawak, Malaysia. SRTM v3 elevation (30 m) with hillshade forms the base. Administrative district boundaries are overlaid, and mangrove extent is shown as a discrete class. District names indicate focal zones for mapping and biomass estimation. The inset locates the study area within Malaysia.

predominantly flat and low-lying, with mangroves restricted mainly to elevations below 20 m. Combined with high annual rainfall and tidal influence, these conditions provide an optimal habitat for diverse mangrove assemblages. (Kanniah et al., 2015; Phua et al., 2023). However, the region is increasingly exposed to land-use transformation, aquaculture development, and coastal infrastructure expansion, all of which threaten mangrove stability and their carbon storage potential. The selection of this area is grounded in both ecological relevance and conservation urgency. Sarawak contains the most significant proportion of Malaysia's mangrove forests, making it a priority zone for biomass monitoring and habitat protection (Alam et al., 2012; Lestari et al., 2024). By focusing on this region, the study aims to map mangrove extent and estimate above-ground biomass using a fused remote sensing approach, contributing to national efforts in blue carbon accounting and sustainable coastal ecosystem management.

2.2. Field data collection

2.2.1. Reference data for mangrove classification

Reference data for classifying mangrove and non-mangrove areas were collected between June and October 2023 and were used to support the study-year (2023) mapping framework, through a combination of field-based observations and expert interpretation of high-resolution satellite imagery. A total of 700 georeferenced points were compiled across representative coastal regions in Malaysia, capturing a wide range of mangrove ecosystems as well as adjacent land-cover types, including agricultural land, built-up areas, and open water bodies. Field surveys were conducted to gather in-situ information, while inaccessible locations were supplemented through visual interpretation of very high-resolution imagery in Google Earth. This dual approach ensured comprehensive coverage of both spatial and ecological gradients. All reference points were manually verified and stratified to maintain a balanced representation across major land-cover types and geographic gradients. The dataset was divided into training and validation subsets using a 70:30 split, ensuring proportional representation of each class (Anees et al., 2024b, 2025b). This stratified sampling design reduced classification bias and strengthened model calibration and accuracy assessment. The final reference dataset served as ground truth for training the supervised classification algorithm (XGBoost) and for validating the resulting 2023 mangrove extent map.

2.2.2. Field-based AGB measurements

AGB reference data were collected from plot-based field measurements within representative mangrove stands between June and October 2023 (dry season in Sarawak; n = 245 plots). Satellite predictors were subsequently extracted from Sentinel-1 and Sentinel-2 acquisitions temporally matched to the field campaign period (or the closest cloud-free observations within the compositing window) to minimise temporal mismatch between in-situ biomass estimates and remotely sensed variables (Table S8). Fixed-area plots of 30 × 30 m were established to ensure compatibility with the spatial resolution of S1 and S2 imagery (Ferreira and Lacerda, 2016; Martinez et al., 2024). The decision to use 30 m plots equivalent to a 3 × 3-pixel window was guided by several practical and methodological considerations (Cherian and Rajitha, 2024; Elbakidze et al., 2016). These include minimizing spatial mismatches due to handheld GNSS horizontal uncertainties of approximately 3–5 m, reducing spectral mixing and canopy edge effects, and achieving consistency with other datasets, such as GEDI and Landsat, which also operate at 30 m resolution. A 3 × 3-pixel window (30 × 30 m) centered on each field plot was used to extract and average spectral and SAR-based variables. This approach reduced residual errors caused by geolocation uncertainty, mixed pixels, and misalignment between field plots and satellite observations, thereby strengthening the empirical linkage between ground measurements and remotely sensed data (Allais et al., 2024; Nguyen et al., 2016). Field sampling was designed to capture representative mangrove structure and species composition; however, some accessibility-related bias is unavoidable in intertidal mangrove environments. Areas with deep mud, dense pneumatophore/root networks, unsafe tidal channels, and difficult boat or foot access were less likely to be sampled intensively. Consequently, highly degraded, recently disturbed, very low-stature, fragmented edge stands, and some remote interior patches may be under-represented relative to more accessible and structurally mature stands. This could slightly shift the field AGB distribution toward better-developed mangrove conditions and reduce the representation of low-biomass or disturbance-affected stands in model calibration. Therefore, the AGB map was interpreted as a regionally calibrated estimate based on accessible field plots, while recognising that prediction uncertainty may be higher in disturbed, fragmented, or poorly sampled stand conditions. This limitation was partly addressed by distributing plots across major mangrove taxa and coastal gradients, pairing field observations with spatially continuous Sentinel-1/Sentinel-2 predictors and evaluating model performance using held-out and out-of-fold predictions.

Within each plot, standard forest inventory protocols were followed, including measuring diameter at breast height (DBH), total tree height using a hypsometer (Oliveira et al., 2021), and species identification. AGB was estimated using species-specific mangrove allometric equations where available and mixed-species Malaysian mangrove equations otherwise. Two allometric forms were considered: a DBH-only model (Eq. (1)) and a combined DBH and height model (Eq. (2)) (Chave et al., 2014; Komiyama et al., 2005).

$$\mathrm{AGB} = \mathrm{a} \times D^{b} \tag{1}$$

$$AGB = a \times \left(D^{2} \times H\right)^{b} \tag{2}$$

where D is the diameter at breast height (cm), H is the total tree height (m), and a and b are empirical coefficients. Since both DBH and tree height were measured for all trees, the D^2H model was consistently applied. A formal multi-equation sensitivity analysis was not undertaken because consistently parameterised regional equations were not available for all recorded taxa and stand conditions. However, equation choice remains an important source of uncertainty in mangrove AGB estimation. Alternative DBH-only or regional mixed-species equations could shift plot-level AGB estimates, particularly in large-diameter, tall, or high-wood-density stands where biomass scales nonlinearly with stem size. Such shifts would be expected to affect the absolute magnitude of AGB more than the broad spatial pattern of relative biomass variation, because the same allometric framework was applied consistently across all plots. This model demonstrated superior predictive performance (R^2 = 0.63–0.96) in structurally diverse mangrove stands (Khan et al., 2025). When log-linear equations were used, back-transform bias was corrected using the smearing estimator (Zhang et al., 2025), and species- or genus-level wood density values were applied from curated databases where required (Mo et al., 2024). Biomass estimates were expressed in megagrams per hectare (Mg ha^{-1}) according to the plot area. Species information was operationalised by reporting species-wise plot AGB summaries (mean, standard deviation, range, and sample size) for the nine recorded taxa (Table S9). For the AGB modelling task, plot-level AGB values were paired with predictor variables averaged within the 3×3-pixel neighbourhood and were used for model development and out-of-sample evaluation (i.e., held-out test predictions and/or cross-validated out-of-fold predictions), ensuring that reported performance metrics were not derived from in-sample fitted values (Corte et al., 2020; Ojoatre et al., 2019).

2.3. Remote sensing datasets for mangrove biomass estimation

This study employed a combination of multisensor satellite observations and ground-based field measurements to estimate AGB in mangrove forests. The integration of optical, radar, and in-situ data was designed to enhance the accuracy and spatial representation of biomass estimations across the study area. The primary remote sensing datasets included Sentinel-1 (S1) and Sentinel-2 (S2) imagery, selected for their complementary spectral and structural information. The overall methodological framework is summarised in Fig. 2, outlining the sequential steps from data acquisition to biomass modelling. All predictor layers were harmonised to a standard 10 m grid in a projected coordinate reference system (WGS 84/UTM zone 50N; EPSG:32650) and cover the analysis window January–December 2023. Scene counts, quality filters, and compositing windows are reported for each sensor in Table S1–S3.

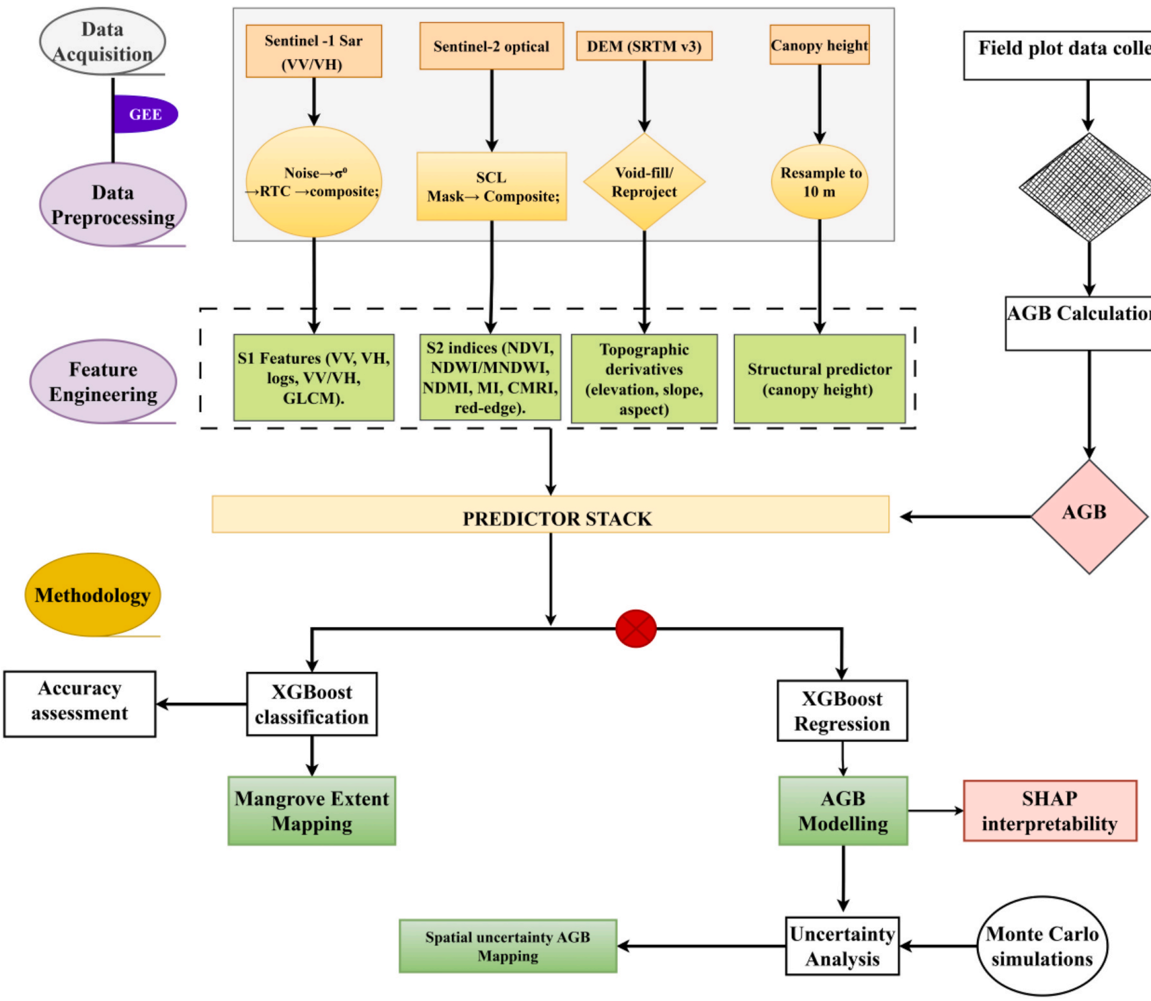

Fig. 2. Methodological workflow for mangrove extent mapping and above-ground biomass estimation in Sarawak using fused Sentinel-1 and Sentinel-2 data, field inventory, and XGBoost modelling.

2.3.1. Sentinel-2 optical data

Sentinel-2 Level-2A surface reflectance (COPERNICUS/S2_SR) was used to derive vegetation- and mangrove-sensitive spectral information. All Sentinel-2 scenes acquired from January to December 2023 were accessed in Google Earth Engine and filtered using a scene-level cloud-cover threshold of $<10\%$; residual cloud and shadow contamination was removed using the Scene Classification Layer (SCL), and a median composite was generated to represent 2023 surface conditions (Table S1). The sensor provides 13 bands at 10–60 m spanning the visible, NIR, red-edge, and SWIR regions (Table S2). These bands supported calculation of NDVI, MNDWI, NDMI, and additional mangrove-oriented indices (Table S4). Bands B2–B4 capture chlorophyll and suspended-sediment signals, B8/B8A characterise canopy structure and water content, and B11/B12 represent SWIR moisture sensitivity and lignocellulosic absorption. To characterise fine-scale spatial heterogeneity relevant to mangrove stands, grey-level co-occurrence matrix (GLCM) textures (contrast, entropy, homogeneity, dissimilarity, correlation, ASM, mean, and variance) were computed from NIR and red-edge bands using a 5×5 window, four directions, and 64 grey levels (Table S4).

2.3.2. Sentinel-1 SAR data

Sentinel-1 C-band SAR was incorporated to provide canopy-structure and moisture information under persistent cloud conditions. C-band backscatter is sensitive to mangrove architecture and surface wetness (Yilmaz et al., 2023) and can be transformed into indices useful for biomass and habitat mapping (Mohammadpour et al., 2022; Park and Guldmann, 2020). Sentinel-1 GRD imagery acquired in Interferometric Wide (IW) mode with dual polarisation (VV and VH) was used for the period January–December 2023 (Table S1; Table S3). To maintain consistent viewing geometry across the annual time series, the analysis was restricted to descending-orbit acquisitions (Table S1). After standard radiometric processing to analysis-ready backscatter (σ^0), a multi-temporal median composite was generated for 2023, from which VV, VH, the VV/VH ratio, and log-transformed backscatter were retained (Table S1; Table S4). Grey-level co-occurrence matrix (GLCM) textures (contrast, entropy, homogeneity, dissimilarity, correlation, ASM, mean, and variance) were derived from VV and VH to characterise canopy roughness and spatial heterogeneity using a 5×5 window, four directions, and 64 grey levels (Jia et al., 2023; Yilmaz et al., 2023; Table S4). Although L-band SAR (ALOS-2 PALSAR) can offer deeper penetration and reduced saturation in very dense mangroves, continuous coverage matching the study-year window was not available; Sentinel-1 was therefore adopted for its open access, revisit capability, and operational scalability. Limitations in high-biomass zones were mitigated through subsequent fusion with Sentinel-2 spectral predictors, canopy-height products, and topographic variables. The workflow can also be extended where higher-resolution data are available; for example, PlanetScope imagery could be incorporated as an optional refinement layer in very high-biomass or structurally heterogeneous mangrove stands to improve the representation of narrow belts, fragmented edges, canopy gaps, and fine-scale texture variation (Hemati et al., 2025). Such an extension would complement, rather than replace, the Sentinel-1/Sentinel-2 baseline framework used in this study.

2.4. Pre-processing of satellite data

Pre-processing was standardised across optical and SAR inputs to generate co-registered, analysis-ready predictor layers. Sentinel-2 cloud and shadow were removed using the Scene Classification Layer (la Cecilia et al., 2023; Nevavuori et al., 2020), and a median composite was generated to represent 2023 surface conditions (Mehmood et al., 2025a, 2025b). Sentinel-1 GRD IW-mode imagery (VV, VH; January–December 2023) was processed to analysis-ready backscatter (σ^0) and terrain-corrected using SRTM v3 ($\sim$30 m) (Potin et al., 2016; Filipponi, 2018; Ali et al., 2025; Sheehy et al., 2024). All SAR layers were then co-registered to the Sentinel-2 10 m grid, which served as the spatial reference because it defines the target mapping resolution for the fused predictor stack. Co-registration accuracy was evaluated using stable linear features (coastline/river margins) across 50 tie locations; the mean horizontal displacement decreased from 11.8 ± 4.6 m before alignment to 3.2 ± 1.7 m after co-registration. A full-year multi-temporal median composite and a 3×3 median filter was applied to reduce speckle (Dagne et al., 2023). From this stack, VV, VH, log (VV), log (VH), and VV/VH were retained, and GLCM textures on VV/VH were computed using a 5×5 window, four directions, and 64 grey levels (Sainuddin et al., 2024). Ancillary layers included SRTM v3 elevation, slope, and aspect, and Global Forest Canopy Height (GFCH) canopy height (10 m spatial resolution); all layers were resampled and harmonised to 10 m using bilinear (continuous) or nearest-neighbour (categorical) schemes (Hussain et al., 2025). Details for each sensor and derived predictors are listed in Table S1–S4.

2.5. Ancillary structural and topographic data

Structural and topographic predictors were added to strengthen the biomass model. GFCH, 10 m; derived from GEDI was used as a single canopy-height predictor; no airborne LiDAR was collected. GFCH was reprojected to WGS 84/UTM 50N (EPSG:32650) and harmonised to the 10 m grid (Neuenschwander et al., 2020). SRTM v3 (1-arc-sec, $\sim$30 m) provided elevation; slope and aspect were derived at native resolution, reprojected, and snapped to the Sentinel-2 10 m grid. Elevation and slope were resampled bilinearly; aspect used nearest neighbour to respect its circular nature. This keeps the DEM's 30 m information while enabling pixel-wise stacking with the 10 m S1/S2 predictors. Distances to coastline and main channels were computed from OpenStreetMap (Geofabrik Malaysia, 2023 extract) and HydroRIVERS v1.0 and rasterised to the standard 10 m grid. All ancillary layers share the same CRS and analysis window (January–December 2023).

2.6. Machine learning-based mangrove extent classification

Mangrove extent classification was implemented using the XGBoost algorithm, leveraging a subset of informative predictors derived from Sentinel-1 and Sentinel-2 imagery. The reference dataset (n = 700 points) was stratified by class and split into 70% for model development and 30% as an independent hold-out test set. Hyperparameter tuning was performed on the 70% subset using repeated stratified five-fold cross-validation (two repetitions), with ROC–AUC as the optimisation criterion. The final model was refit on the full 70% training subset and evaluated on the independent 30% hold-out set using confusion-matrix metrics (overall accuracy, kappa, precision, recall/sensitivity, specificity, balanced accuracy) and ROC–AUC. Predictors comprised Sentinel-2 spectral bands and derived indices (e.g., vegetation and moisture indices), Sentinel-1 backscatter metrics (VV, VH, VV/VH and log-transforms), and texture features (GLCM) derived from both sensors (Table S1–S4).

Predictor selection and interpretability were supported using SHAP analysis, which enabled identification of the most influential variables contributing to the classification outcome (Mehmood et al., 2025b) (Fig. S1). A conservative set of hyperparameters was adopted to enhance model generalisation and prevent overfitting. The final XGBoost configuration included 100 boosting rounds (nrounds = 100), maximum tree depth = 3 (max_depth = 3), learning rate = 0.1 (eta = 0.1), gamma = 1, colsample_bytree = 0.8, subsample = 0.8, and minimum child weight = 5. Fixed random seeds were used to ensure model reproducibility.

To assess classification uncertainty, an ensemble modelling approach was applied by training ten XGBoost models on bootstrapped samples of the training data. Predictions were aggregated to compute the mean probability and standard deviation for each pixel. The standard deviation was used as a proxy for uncertainty and reclassified into four ordinal levels: very low (<0.1), low (0.1–0.2), moderate (0.2–0.3), and high (≥0.3). The spatial distribution and proportional coverage of these uncertainty levels were quantified to assess confidence across the mapped mangrove extent.

2.7. AGB prediction using remote sensing and XGBoost

AGB was predicted with XGBoost regression using the fused Sentinel-1/2 predictor stack, including SAR backscatter, texture metrics, spectral bands, vegetation indices, and ancillary predictors. XGBoost trains an additive ensemble of decision trees; each tree fits the residuals of the current model, and the ensemble is regularised by shrinkage, tree-depth limits, subsampling, and L1/L2 penalties (Mehmood et al., 2025c). A squared-error objective was used, and model performance was evaluated only from held-out predictions rather than from in-sample fitted values. Boruta guided predictor parsimony by retaining features with importance consistently above shadow controls, after which the model was refit on the retained predictor set (Agarwal et al., 2023). Outliers in plot-level AGB were screened using the IQR rule.

Model evaluation followed two complementary held-out protocols to avoid ambiguity between the independent test set and the full-sample cross-validated predictions. First, the dataset was split into fixed 80% training and 20% independent test subsets using fixed random seeds. The 20% test subset was kept completely unseen during feature selection, hyperparameter tuning, and early stopping, and was used only once for final independent test-set evaluation. Second, spatially blocked out-of-fold predictions were generated across the full dataset (n = 245) to provide one held-out prediction for every field plot and to support the observed–predicted scatterplot. Thus, Fig. 6 is based on full-sample out-of-fold predictions, whereas the fixed independent 20% test-set metrics are reported separately as a final benchmark.

Model training was conducted using the squared-error objective (`reg:squarederror`) and RMSE as the evaluation metric. Key hyperparameters included maximum tree depth of 6, learning rate (`eta`) of 0.1, `subsample = 0.8`, and `colsample_bytree = 0.8`. Additional tuned parameters included `learning_rate`, `max_depth`, `n_estimators`, `subsample`, `colsample_bytree`, `min_child_weight`, λ, and α. Hyperparameter tuning and early stopping were performed within the training subset only, using five-fold cross-validation with early stopping after 10 rounds to identify the optimal number of boosting rounds (Islam et al., 2023).

To reduce spatial leakage among nearby plots, cross-validation was implemented using spatially blocked folds. Plot centroids were grouped into spatial blocks before fold assignment, ensuring that neighbouring plots within the same local cluster were retained within the same fold rather than being split between training and validation subsets. This provided a stricter assessment of model transferability than ordinary random plot-level cross-validation. In the full-sample out-of-fold procedure, each plot was predicted by a model trained without the spatial block containing that plot. These predictions were then combined across folds to generate held-out predictions for all 245 plots. Model performance was evaluated using RMSE, mean absolute error (MAE), coefficient of determination (R^2), and Pearson's correlation coefficient.

Prediction consistency and model reliability were assessed through 100 Monte Carlo simulations using bootstrapped samples from the training data (Sarker, 2010). In each iteration, an XGBoost model was trained, tuned through cross-validation, and evaluated on a held-out test subset. The resulting distributions of RMSE and R^2 values provided insight into variability in model performance across repeated resampling. Mean and standard deviation were calculated for each metric, and their distributions were visualised using histograms to examine the stability of AGB predictions across simulations.

Residual spatial dependence was examined as an additional diagnostic of spatial robustness. Global Moran's I was used to test spatial autocorrelation in observed AGB and in model residuals. Residuals were defined as $e_i = y_i - \hat{y}_i$, where y_i is the observed AGB and $\hat{y}_i$ is the held-out predicted AGB for plot *i*. A row-standardised k-nearest-neighbour weights matrix (k = 8) was built from plot centroids. Significance was assessed using 999 random permutations under the null hypothesis of spatial randomness. As a reliability check, the analysis was repeated using a fixed distance band that produced a similar neighbour count. Computations were performed in R using the `sf` and `spdep` packages.

2.8. Model evaluation and performance metrics

Model performance for both mangrove extent classification and AGB prediction was evaluated using a combination of statistical and classification-based metrics in accordance with established recommendations (Anees et al., 2024a). For classification, performance metrics included overall accuracy, the kappa coefficient, sensitivity (also known as recall), specificity, precision, balanced accuracy, and the area under the receiver operating characteristic curve (AUC). The classification threshold was optimized using the "closest-to-top-left" method on the ROC curve. These metrics were computed on an independent validation set, enabling an unbiased assessment of the classifier's ability to differentiate between mangrove and non-mangrove classes. For AGB estimation, model evaluation employed RMSE, mean absolute error (MAE), coefficient of determination (R^2), bias, and relative root mean square error (rRMSE), following recommended practices (Chuine and Beaubien, 2001; Rodigheri et al., 2023). Pearson's correlation coefficient (r) was used to quantify the linear association between predicted and observed AGB. For clarity, the mathematical definitions of RMSE, rRMSE (mean-normalised and range-normalised forms), MAE, bias, and Pearson's r are provided in Supplementary S1.

SHAP (SHapley Additive exPlanations) was applied to assess the influence of individual predictors on AGB predictions generated by the XGBoost model. Rooted in cooperative game theory, the SHAP framework assigns a contribution value (ϕ) to each feature for every prediction, thereby providing both global and local interpretability (Mehmood et al., 2025b).

$$\varphi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} [f(S \cup \{J\}) - F(S)] \tag{3}$$

where N is the set of all features, S is a subset of features excluding j, and f(S) is the model prediction based on features in S. The shapviz package in R was employed to compute and visualize SHAP values based on the final trained model and its corresponding predictor matrix. A combined importance plot was generated to display the average absolute SHAP values for each variable, alongside the distribution of their effects across all observations. This dual-format visualization facilitated both the ranking of predictor importance and the assessment of variability in their contributions to AGB estimates. For local interpretation, a SHAP waterfall plot was constructed for a single observation (Mehmood et al., 2025b), illustrating the additive contributions of each predictor in relation to the model's baseline prediction. This visualization clarified the extent to which individual features increased or decreased the predicted AGB value for that specific case.

3. Results

3.1. Observed AGB patterns across mangrove species and plots

AGB estimates from 245 mangrove plots revealed substantial variability, both across species and spatially among plots. While several plots comprised mixed-species assemblages, each plot was assigned to the species with the largest basal area for the purpose of species-level AGB reporting. This approach, widely applied in mangrove biomass studies, enables meaningful interspecific comparisons while retaining the influence of structural diversity within the overall analysis. Species-level summaries showed that *Rhizophora mucronata* recorded the highest mean AGB (102.1 Mg ha^{-1}), followed by *Sonneratia alba* (100.1 Mg ha^{-1}). In contrast, *Bruguiera parviflora* had the lowest mean AGB (74.5 Mg ha^{-1}) (Table 1). These interspecific differences reflect underlying variations in wood density, tree height, and site-specific growing conditions. In addition to species-level variation, the ranked distribution of observed

Table 1
Species-wise summary statistics of observed above-ground biomass (AGB) and stand structure from 245 mangrove plots (30 × 30 m) across Sarawak, Malaysia. AGB is reported in Mg ha^{-1}. Mean tree count, mean DBH (cm), and mean height (m) are averaged across plots assigned to each dominant taxon (dominant basal area). Wood density (g cm^{-3}) denotes the species- or genus-level value used in the allometric calculations.

Scientific name	Genus	Plot count	Mean AGB (Mg ha^{-1})	SD	Min	Max	Mean tree count (per 30 × 30 m plot)	Mean DBH (cm)	Mean height (m)	Wood density (g cm^{-3})
Rhizophora mucronata	Rhizophora	37	102.07	56.26	23.00	220.00	61.0	15.26	15.57	0.85
Sonneratia alba	Sonneratia	25	100.14	58.05	24.00	230.00	83.0	15.44	18.52	0.50
Rhizophora apiculata	Rhizophora	49	97.78	42.96	31.73	210.14	64.8	14.91	14.99	0.85
Sonneratia caseolaris	Sonneratia	20	93.71	51.64	23.26	165.00	75.2	15.58	18.21	0.50
Bruguiera gymnorhiza	Bruguiera	29	91.88	61.13	21.00	212.00	76.8	13.96	14.71	0.74
Avicennia marina	Avicennia	25	89.96	57.38	30.00	224.00	69.4	14.42	14.05	0.70
Avicennia officinalis	Avicennia	12	86.94	39.98	33.00	148.00	65.3	14.69	13.76	0.70
Avicennia alba	Avicennia	29	78.06	46.00	22.00	237.46	66.6	14.31	13.64	0.70
Bruguiera parviflora	Bruguiera	19	74.54	43.96	20.03	151.31	85.8	11.85	13.65	0.74

AGB across all plots illustrates a gradual biomass gradient, with a small number of high-biomass outliers exceeding 200 Mg ha^{-1} (Fig. 3). This plot-level variation emphasizes the structural heterogeneity within the sampled mangrove stands and supports the diversity captured during field sampling.

3.2. Mangrove extent classification and prediction confidence

The XGBoost model produced accurate and reliable predictions of mangrove extent across the study region. Classification performance was strong, with an overall accuracy of 88.8% and a Cohen's Kappa of 0.776, indicating substantial agreement beyond random chance. Sensitivity for the mangrove class reached 92.8%, while specificity was 84.7%, reflecting the model's balanced ability to detect both mangrove and non-mangrove classes correctly. The precision for mangrove classification was 86.5%, resulting in a balanced accuracy of 88.7% (Table S4). ROC analysis further confirmed the model's discriminative power, yielding an AUC of 0.963 (Fig. 4). The optimal probability threshold for distinguishing mangrove from non-mangrove pixels was determined to be 0.728, using the “closest-to-top-left” criterion (Fig. S3). At this threshold, the model achieved a sensitivity of 87.6% and a specificity of 92.0%, indicating a well-calibrated decision boundary (Fig. 5).

Applying this threshold across the full study area yielded a predicted mangrove extent of 1078.70 km^2. The optimal threshold was $\tau = 0.728$ from the ROC closest-to-(0,1) criterion; at this operating point, sensitivity was 87.6% and specifically 92.0%. The classification map effectively delineated mangrove and non-mangrove zones, while a complementary probability map highlighted confidence gradients in predicted mangrove presence (Fig. 6). Areas with high classification probability corresponded closely to dense and well-established mangrove stands, whereas lower probabilities were associated with fragmented or transitional regions. Model confidence was further assessed using ensemble predictions from fifty (50) bootstrap iterations. Standard deviation values derived from these ensemble outputs were used to quantify classification uncertainty. Most of the predicted mangrove area (83.8%, or 904.25 km^2) exhibited very low uncertainty, reflecting consistent agreement across models. An additional 6.96% (75.09 km^2) and 7.29% (78.65 km^2) were classified as low and moderate uncertainty, respectively. Only 1.92% of the area (20.71 km^2) exhibited high uncertainty, primarily located along classification boundaries or in spectrally ambiguous zones. A stratified random sample of 225 reference points interpreted from very-high-resolution imagery yielded overall accuracy = 0.929, $\kappa = 0.674$, MCC = 0.679, and balanced accuracy = 0.810 (Table S11). The confusion matrix (Table S13) shows mangrove: PA = 0.645, UA = 0.800 (F1 = 0.714) and non-mangrove: PA = 0.974, UA = 0.945 (F1 = 0.959). Using a mapped-stratum estimator, the area-adjusted OA = 0.942 (95% CI 0.911–0.974) (Table S14). Errors are spatially concentrated along class boundaries and fragmented patches, consistent with the probability map and the operating point $\tau = 0.728$, and their distribution is illustrated in Fig. S5.

Overall accuracy was 88.8% ($\kappa = 0.776$) with AUC = 0.960. A 50-run bootstrap uncertainty map showed 83.8% of mapped mangroves in the very-low-uncertainty class, with high-uncertainty pixels concentrated along class boundaries and fragmented zones. These patterns indicate spatially coherent predictions with quantified uncertainty at the chosen operating point ($\tau = 0.728$).

3.3. AGB prediction and model interpretation

The XGBoost regression model demonstrated reliable performance in predicting AGB across the study area. Model accuracy was quantified using held-out predictions to avoid optimistic bias, specifically through the spatially blocked out-of-fold prediction scheme described in Section 2.7. The out-of-fold evaluation included all 245 plots, but each plot was predicted only by a model that did not

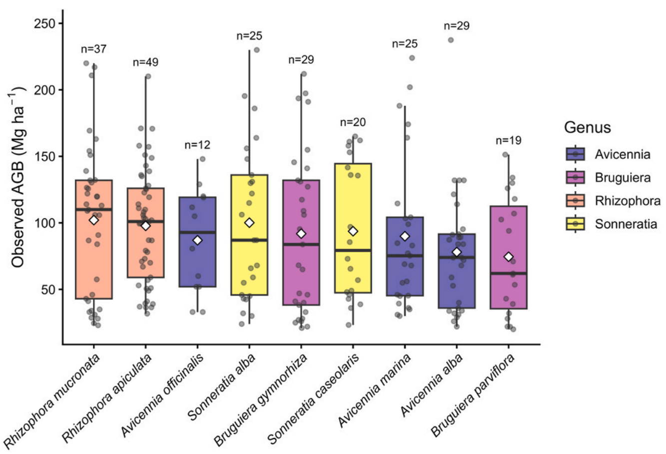

Fig. 3. Species-wise distribution of observed above-ground biomass (AGB) across 245 mangrove plots (30 × 30 m) in Sarawak, Malaysia. Boxplots show the median and interquartile range (IQR) of plot-level AGB for each species; whiskers extend to 1.5 × IQR. Points represent individual plots, and sample sizes (n) are shown above each species. Boxes are coloured by genus, and the white diamond denotes the mean. AGB is reported in Mg ha^{-1}.No significant spatial autocorrelation was detected in observed AGB or in model residuals (global Moran's I; 999 permutations); details are provided in Supplementary text S2 and Fig. S6.

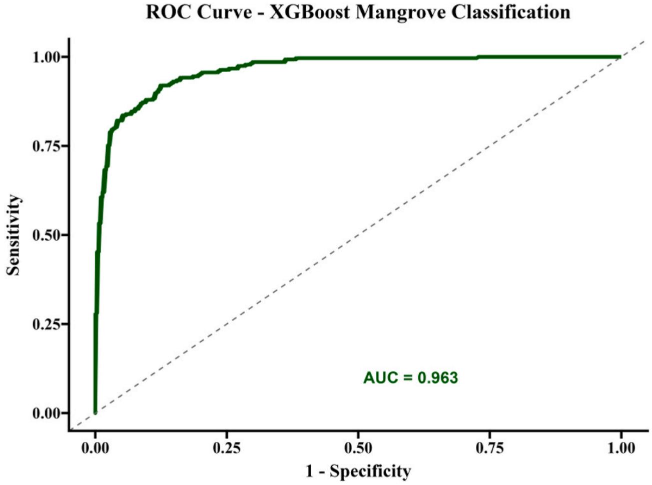

Fig. 4. ROC curve for mangrove extent classification using the XGBoost model. The model achieved a high AUC of 0.963, indicating excellent discriminatory ability between mangrove and non-mangrove classes. The curve illustrates the trade-off between sensitivity and specificity across classification thresholds, with the dashed line representing random performance.

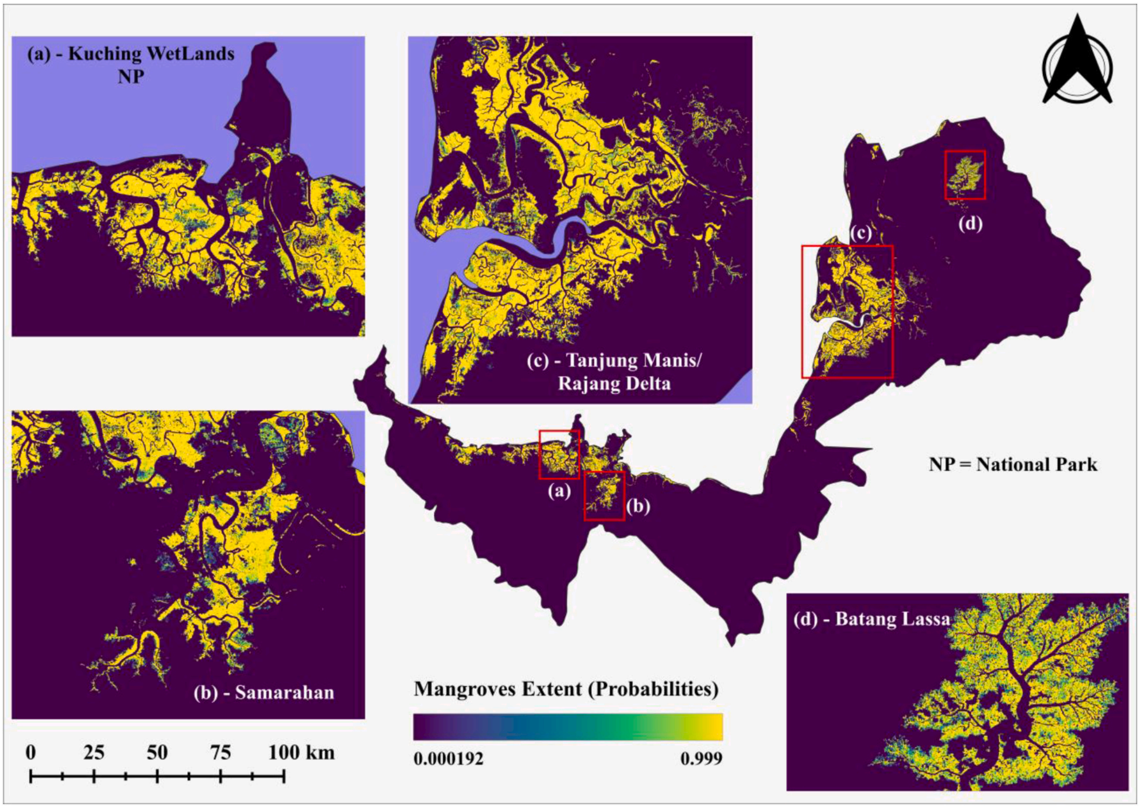

Fig. 5. Spatial prediction of mangrove extent and associated classification probabilities across selected districts in Sarawak, Malaysia. Insets (a) Kuching Wetlands NP, (b) Samarahan, (c) Tanjung Manis/Rajang Delta, and (d) Batang Lassa illustrate varying patterns of mangrove distribution and classification confidence. The color gradient represents the probability of mangrove presence, with darker green indicating higher confidence. Core mangrove zones exhibit high classification certainty, while transition zones show reduced model confidence.

include the spatial block containing that plot during training. Therefore, Fig. 6 scatterplot represents full-sample held-out prediction performance rather than in-sample fitted values. The out-of-fold evaluation yielded an RMSE of 28.59 Mg ha^{-1} and an MAE of 20.80 Mg ha^{-1}. The coefficient of determination (R^2) was 0.72, indicating that approximately 72% of the variance in observed AGB was

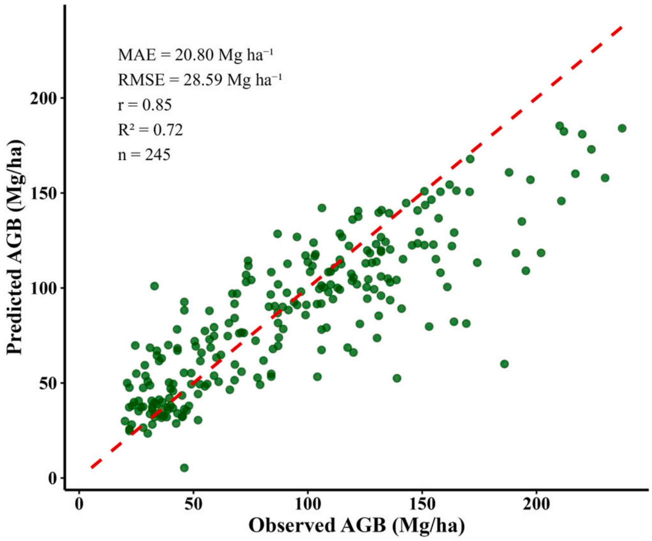

Fig. 6. Scatterplot of observed versus predicted above-ground biomass (AGB) values using the XGBoost regression model. The red dashed line represents the 1:1 reference line, indicating perfect agreement between observed and predicted values. Green dots represent individual plots (n = 245). Predicted values are spatially blocked out-of-fold predictions, meaning that each plot was predicted by a model trained without the spatial block containing that plot. The metrics shown in this figure therefore describe full-sample out-of-fold performance and are distinct from the fixed independent 20% test-set metrics reported separately in the text and supplementary benchmarking table.

explained under held-out prediction. Pearson's correlation coefficient (r) between observed and predicted AGB values was 0.85, signifying a strong positive linear relationship (Fig. 6; n = 245).

The fixed independent 20% test set was used as a separate final benchmark and was not used during feature selection, hyperparameter tuning, or early stopping. This independent test-set evaluation yielded RMSE = 25.53 Mg ha^{-1}, MAE = 19.39 Mg ha^{-1}, R^2 = 0.79, and Pearson's r = 0.90. These values are reported separately from the full-sample out-of-fold metrics because the two evaluations serve different purposes: the independent test set provides a single withheld-subset benchmark, whereas the out-of-fold evaluation provides one held-out prediction for every field plot and supports the observed–predicted scatterplot.

The observed versus predicted AGB scatterplot exhibited a clear alignment along the 1:1 line, with most predictions concentrated around moderate to high AGB values. Some degree of underestimation was observed in plots with very high AGB, a common occurrence in remote sensing-based biomass estimation due to signal saturation and structural complexity in dense stands.

In addition to continuous prediction values, AGB was classified into five discrete biomass classes to facilitate interpretation and mapping. The classes were defined using rounded, reader-friendly thresholds, ranging from very low biomass (<50 Mg ha^{-1}) to very high biomass (≥200 Mg ha^{-1}) (Table S6). Most of the mapped mangrove area (51.42%) fell within the 150–200 Mg ha^{-1} class, followed by 23.84% in the 100–150 Mg ha^{-1} class. Areas with ≥200 Mg ha^{-1} accounted for 21.32% of mangrove extent (Table S6). The classified AGB map (Fig. 7) highlights pronounced spatial heterogeneity, with higher biomass concentrations mainly occurring in the central and northern portions of the study area, while lower biomass classes are more common near urban fringes and fragmented coastal zones.

Feature-group experiments were run to see which sensor branch carries more biomass information. XGBoost was re-trained with (i) Sentinel-1 backscatter plus vegetation/moisture indices and ancillary variables and (ii) Sentinel-2 bands plus the same indices and ancillary variables, using the same 80/20 split and tuning settings. The S1-based stack performed slightly better (RMSE = 27.11 Mg ha^{-1}; MAE = 21.18 Mg ha^{-1}; R^2 = 0.715) than the S2-based stack (RMSE = 29.43 Mg ha^{-1}; MAE = 22.28 Mg ha^{-1}; R^2 = 0.697), indicating that C-band backscatter, when combined with NDVI/NDMI/SAVI and terrain–height variables, captures mangrove structure reasonably well (Table S15). After spatial pairing (≤25 m), n = 212-point pairs were retained. C-band AGB showed strong concordance with the GEDI reference (R^2 = 0.74, 95% CI 0.67–0.80; r = 0.86, 0.82–0.89; CCC = 0.85). Absolute errors were RMSE = 30.5 Mg ha^{-1} (27.5–33.3) and MAE = 24.1 Mg ha^{-1} (21.6–26.7). Mean difference (GEDI − C-band) was 2.2 Mg ha^{-1} (−1.9 to 6.1), indicating no systematic offset. The OLS slope was 1.00 (95% CI 0.92–1.07) with an intercept of 2.14 Mg ha^{-1} (−5.68 to 10.30), consistent with minimal scaling bias. Bland–Altman limits of agreement were centered near zero and showed no pronounced heteroscedasticity across the biomass range, corroborating the 1:1 scatter diagnostic (Fig. 8).

Alternative regressors were benchmarked Random Forest and support-vector regression with an RBF kernel under the same train–test split, cross-validation folds, and tuning/early-stopping settings as the XGBoost model. Both produced RMSE and R^2 values comparable to XGBoost and did not consistently reduce error across resamples. A pared predictor set delivered similar performance, indicating that results are not driven by weakly informative features. The remaining error likely reflects an intrinsic floor set by 30 × 30 m plot–10 m pixel incongruity, within-plot structural heterogeneity, uncertainty in species-specific allometry, and partial C-band saturation at very high biomass. Benchmark outcomes are summarised in Table S19.

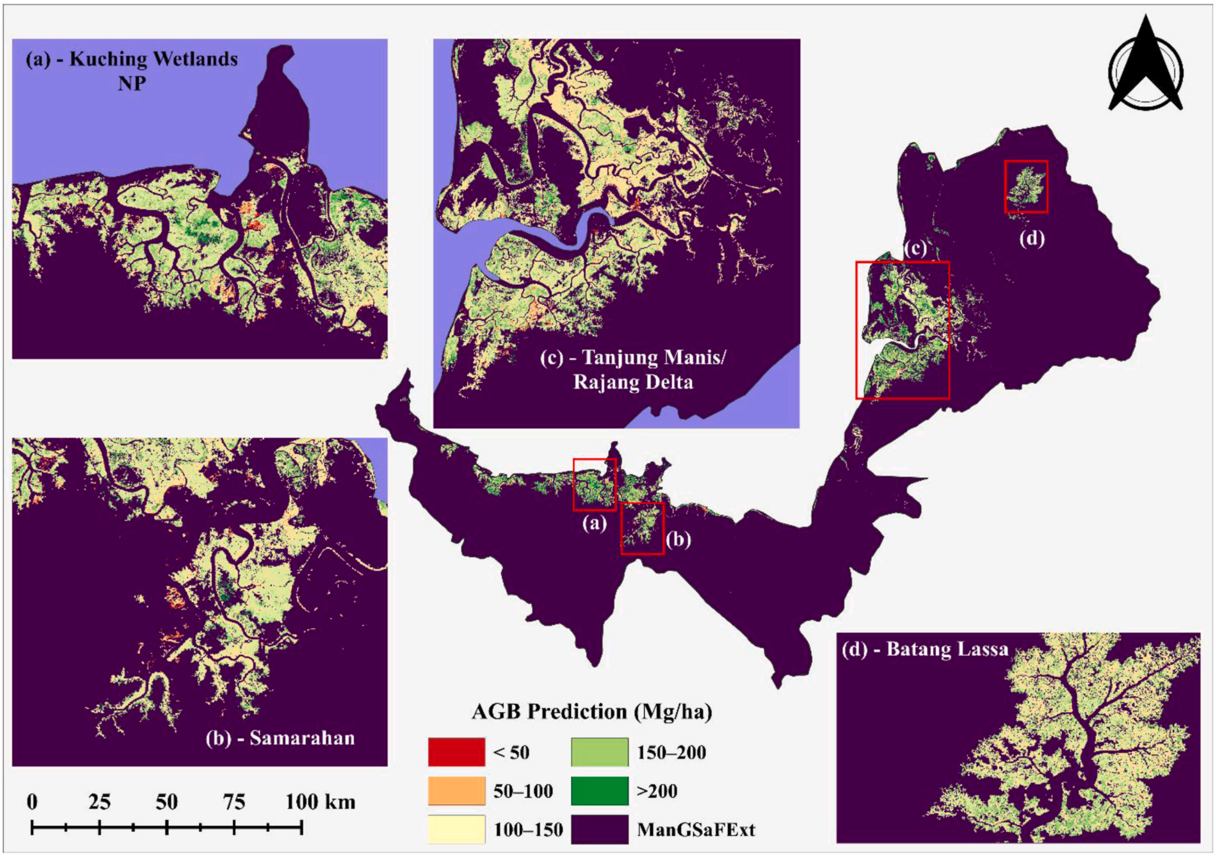

Fig. 7. Spatial distribution of predicted above-ground biomass (AGB) across mangrove areas in Sarawak, Malaysia. The map displays five biomass classes using rounded thresholds: <50, 50–100, 100–150, 150–200, and ≥200 Mg ha^{-1} (Table S6). Insets show localized patterns of biomass variation in (a) Kuching Wetlands NP, (b) Samarahan, (c) Tanjung Manis/Rajang Delta, and (d) Batang Lassa. Higher biomass concentrations are prevalent in central and northern coastal mangroves, while lower biomass values occur near urban fringes and fragmented zones. The classification is based on the XGBoost regression output using fused Sentinel-1 and Sentinel-2 predictors.

3.4. SHAP-based interpretation of the fused AGB model and sensors

Model interpretation was conducted using SHAP values, enabling transparent assessment of predictor contributions. The summary plot (Fig. 9A) ranked predictor importance based on mean absolute SHAP values. Canopy height (CHT) emerged as the most influential predictor, followed by elevation, VH backscatter difference (VHdB), NDVI, and NDMI. These variables consistently contributed higher predictive power, reflecting the structural and ecological characteristics that govern biomass accumulation in mangrove forests. A single-observation SHAP waterfall plot (Fig. 9B) further illustrates the additive contributions of key features to an individual AGB prediction. Positive contributions from high canopy height, NDVI, and s2B4 reflect their association with increased biomass, while slope and lower NDMI values slightly reduced the predicted value, consistent with ecological expectations.

SHAP values also revealed nonlinear and interacting effects among predictors. For instance, high NDVI and NDMI values positively influenced predictions only in combination with sufficient canopy height and radar backscatter intensity. Among the Sentinel-2 bands, B4, B3, and B12 contributed more meaningfully than the others, suggesting their relevance in capturing vegetation signals. Terrain variables such as slope and aspect had a moderate influence, primarily acting as secondary modifiers in areas with topographic variation. These results highlight the importance of integrating spectral, structural, and topographic predictors for satellite-based AGB modeling in mangrove systems. The model's ability to generalize across the field plots, supported by SHAP-based interpretability, affirms the effectiveness of the fused Sentinel-1 and Sentinel-2 approach under machine learning frameworks.

To examine the relative contribution of each sensor family, SHAP analyses were also generated for the two reduced models (S1 + indices + ancillary and S2 + indices + ancillary). Both models preserved CHT and elevation as dominant predictors, but the S1 model assigned higher influence to VH-based features and NDMI, whereas the S2 model increased the importance of NDVI and red/SWIR bands (s2B4, s2B12). This pattern agrees with the groupwise modelling results (Table S14), where the fused predictor stack outperformed the two single-sensor stacks, indicating that radar and optical information are complementary rather than redundant (Fig. 10). These SHAP results provide a practical guide for similar tropical mangrove applications: where full predictor stacks are not feasible, priority should be given to structural layers such as canopy height and elevation/DEM-derived variables, together with VH

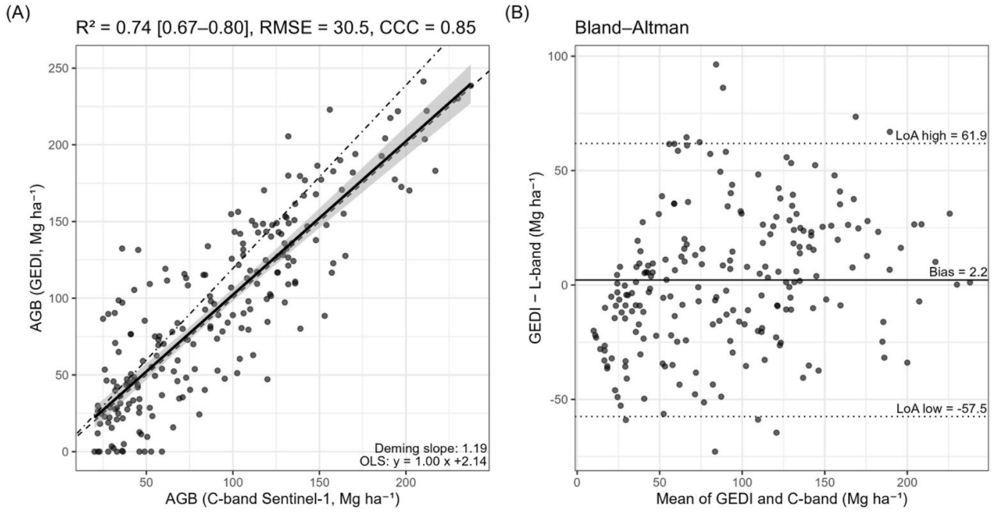

Fig. 8. Validation of C-band AGB against GEDI. (A) 1:1 comparison of C-band PALSAR AGB (x-axis) with GEDI AGB (y-axis). The dashed line denotes the 1:1 reference; the solid line is the ordinary least-squares fit with its 95% confidence band; the dot–dash line (if present) shows the Deming slope to account for error on both axes. Headline statistics (R^2, RMSE, CCC) are reported in the panel title. (B) Bland–Altman plot of GEDI − C-band (Mg ha^{-1}) versus the pairwise mean, with the solid line indicating the mean difference (bias) and dotted lines the 95% limits of agreement. Point pairs were matched within 25 m to ensure spatial coincidence.

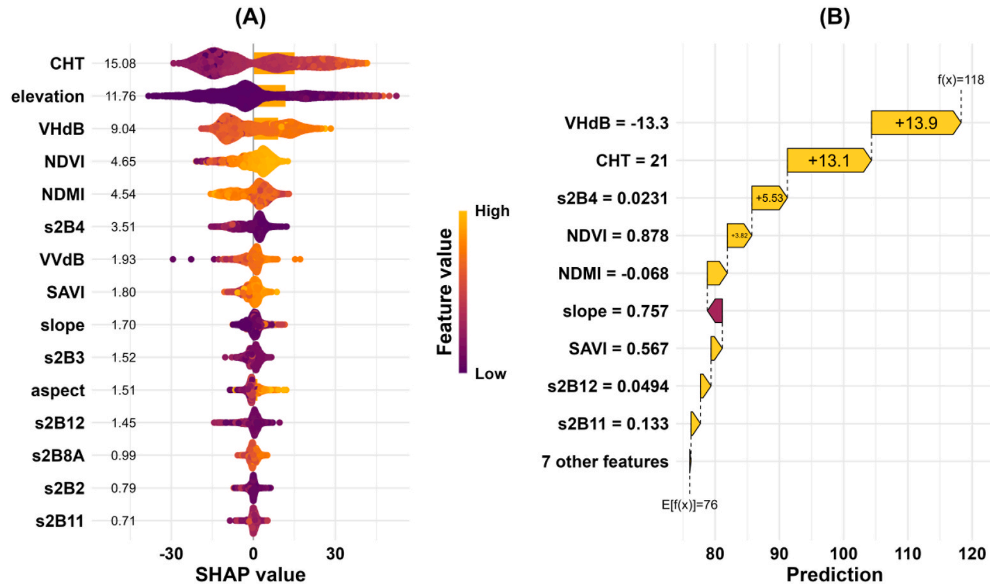

Fig. 9. (A) SHAP summary plot showing the influence of individual predictors on AGB predictions. Feature importance is ranked by mean absolute SHAP values, with warmer colors representing higher feature values. (B) SHAP waterfall plot for representative observation, highlighting the cumulative impact of each variable on the model prediction relative to the mean baseline.

backscatter and Sentinel-2 red, SWIR, red-edge, and vegetation/moisture indices. In operational monitoring, these variables should be treated as core predictors because they capture the main structural, hydrological, and canopy-condition controls on mangrove AGB.

3.5. Model stability and uncertainty assessment via Monte Carlo simulations

Model stability and generalization consistency were examined through 100 Monte Carlo simulations. Each simulation involved bootstrapping the training dataset, retraining the XGBoost model, and evaluating its performance on an independent test subset (Table S7). This resampling-based approach allowed for the quantification of prediction uncertainty and variability in model performance across diverse training configurations. The distribution of RMSE values across the simulations ranged from 23.25 to 29.12

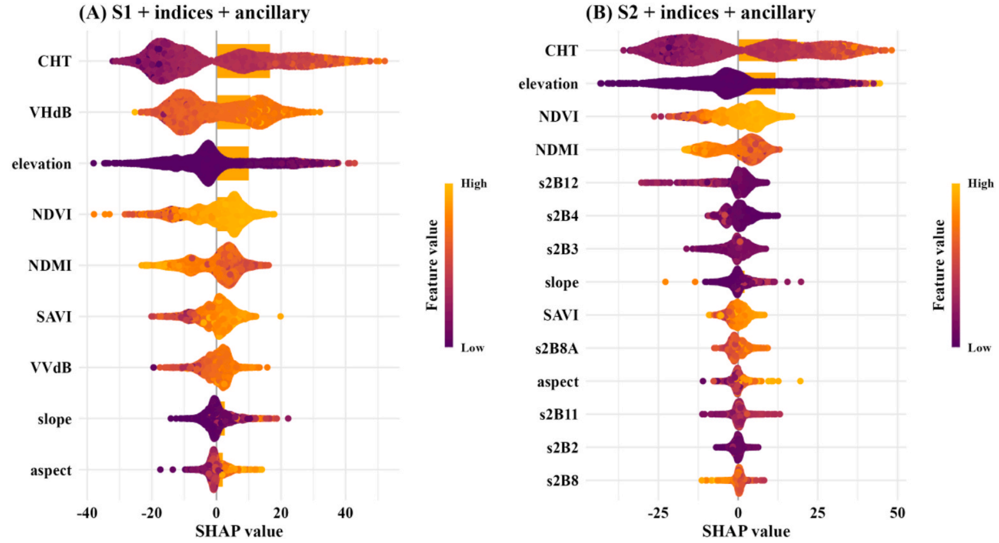

Fig. 10. SHAP summary plots for (A) Sentinel-1 + indices + ancillary and (B) Sentinel-2 + indices + ancillary AGB models. Structural (CHT) and topographic (elevation) variables remain the leading predictors in both cases, while sensor-specific variables (VH for S1; red/SWIR and NDVI for S2) shift in importance, supporting the use of the fused predictor stack.

Mg ha^{-1}, with a mean of 26.25 ± 1.33 Mg ha^{-1} (Fig. 11A). The MAE averaged 15.56 ± 0.80 Mg ha^{-1}, while the R^2 ranged from 0.73 to 0.84, with a mean value of 0.783 ± 0.023 (Fig. 11B). These narrow standard deviations across all performance metrics suggest minimal sensitivity of the model to data variation and strong repeatability in biomass prediction outcomes. The RMSE distribution was unimodal and symmetric, centered near the average model error observed in the single-model evaluation. Likewise, R^2 values were consistently concentrated around the 0.78–0.80 interval, indicating a persistent high level of explained variance, even under varying training subsets. This confirms that the XGBoost model-maintained performance consistency across multiple runs and was not overly reliant on any specific training configuration.

Together, these results reinforce the credibility of the fused S1 and S2 predictor set in modeling mangrove above-ground biomass. The simulation framework adds confidence to the prediction outputs and supports their application for spatial biomass estimation at broader scales.

4. Discussion

The fusion of Sentinel-1 and Sentinel-2 data demonstrated strong performance in mapping mangrove extent in Sarawak, achieving

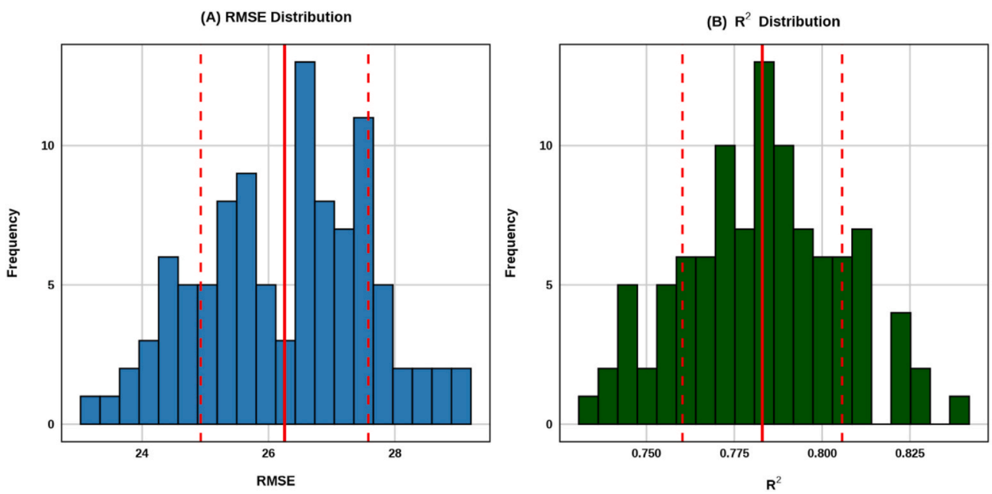

Fig. 11. (A) Histogram of RMSE values across 100 Monte Carlo simulations. The solid red line denotes the mean RMSE (26.25), while dashed lines represent one standard deviation above and below the mean. (B) Distribution of R^2 values across simulations, with analogous markings for the mean (0.783) and standard deviation. The stability of the distributions reflects strong generalization capacity and resilience to sampling variation.

an overall accuracy of 88.8% and a Cohen's kappa of 0.78, indicating substantial agreement beyond chance. These values are comparable to multi-sensor applications in other regions, such as Iran's Hara Reserve (>93% accuracy) and the Global Mangrove Watch, which reported ~94–95% accuracy using L-band radar and optical data (Bunting et al., 2022; Ghorbanian et al., 2021). Comparison of the 2023 Sarawak map with the global 10-m mangrove layer of Jia et al. (2023) shows that the locally trained product reports the same order of mangrove extent (1078.7 km^2 in this study) but adds only about 78.7 km^2 of mangrove that is not present in the global layer. This small surplus is concentrated along narrow riverine and estuarine fringes that global products tend to generalize, as indicated by the difference map (red class, ~78.7 km^2) (Fig. 12). The estimated mangrove extent (~1000 km^2) remains within the range of national inventories and represents about one-quarter of Malaysia's ~0.64 million ha of mangroves, supporting classification reliability (Kanniah et al., 2015; Shukor and Hamid, n.d.). Alignment with national figures, together with a modest and spatially explainable increase over the global layer, indicates that the Sarawak-specific product is refining, rather than inflating, existing baselines for management use. Slight departures from the highest-accuracy studies are likely related to Sarawak's complex land-cover gradients, especially transitions to peat swamps and Nypa stands. The high sensitivity (92.8%) and specificity (84.7%) further confirm the capacity of the classification to separate mangrove from non-mangrove classes. SAR–optical fusion therefore remains advantageous over single-sensor or coarser-resolution approaches for mangrove mapping in heterogeneous tropical coasts (Bunting et al., 2018, 2023).

The AGB modelling demonstrated strong predictive performance when evaluated using held-out predictions to avoid optimistic bias. Out-of-fold cross-validated predictions across all plots (n = 245) yielded $R^2 = 0.72$, Pearson's r = 0.85, RMSE = 28.59 Mg ha^{-1}, and MAE = 20.80 Mg ha^{-1}, indicating that the model explains a substantial fraction of plot-level AGB variability under independent (held-out) prediction. These results are comparable to prior studies in Southeast Asia. For instance, Pham et al. (2020) reported $R^2 \approx 0.80$ and RMSE ~28 Mg ha^{-1} in the Can Gio mangroves of Vietnam, while an XGBoost model applied in the Red River Delta yielded $R^2 \approx 0.68$ and RMSE ~25 Mg ha^{-1} under a multi-sensor approach (Liu et al., 2021). Traditional optical-only methods have demonstrated lower performance, with R^2 values often in the 0.5–0.7 range; for example, QuickBird imagery achieved an R^2 of ~0.65 for mangrove AGB in Southern Thailand (Pesiu et al., 2022). Our fusion-based approach remained within the upper performance range reported in the literature while explicitly avoiding in-sample validation. The relative error magnitude (~33% of mean AGB) aligns with norms for biomass modelling in structurally heterogeneous mangrove systems. Observed versus predicted AGB values were tightly clustered around the 1:1 line, with only minor underestimation in high-biomass plots, underscoring the model's reliability. The mapped values were contextualised by comparing the Sarawak AGB layer with the global mangrove structural/biomass product of Simard et al. (2019) and with Malaysian blue-carbon syntheses advocating higher-resolution, state-level baselines (Lee et al., 2024). At the province scale, our total mapped AGB was modestly higher than the Simard et al. (2019) surface over the same mangrove footprint, which is expected because our map was generated at 10 m, used locally calibrated plots, and retained narrow, estuarine and regenerating stands that are often generalised in global 25–100 m products. A further source of divergence lies in the allometry: Equations applied which

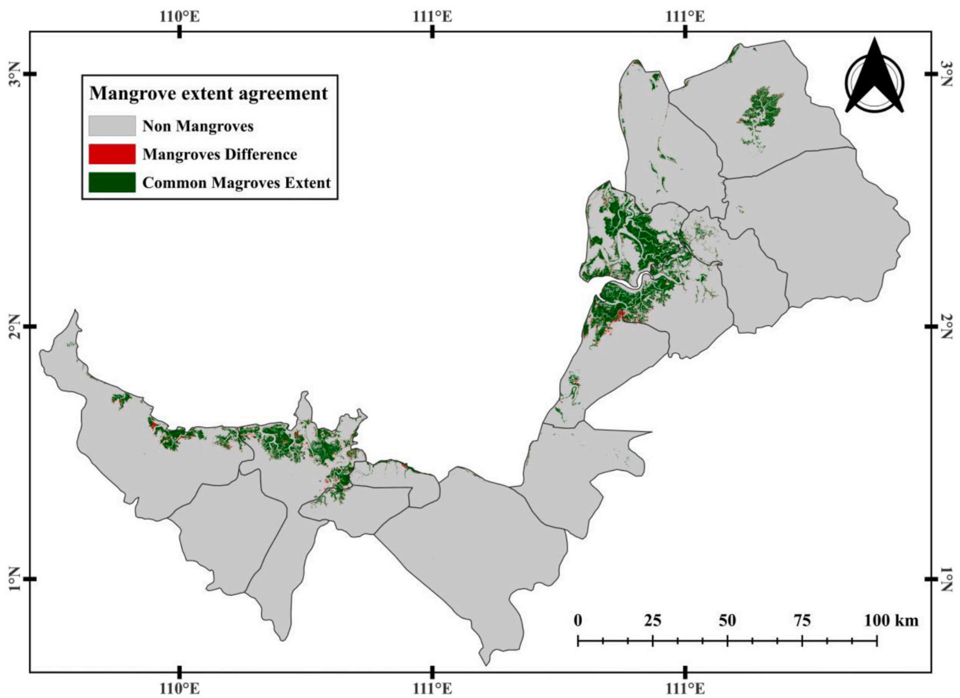

Fig. 12. Mangrove extent agreement map for Sarawak showing areas where the Sentinel-1/2 fusion classification and the global 10-m mangrove dataset (Jia et al., 2023) coincide and where they differ. Green polygons represent mangroves mapped by both datasets (common mangrove extent), red patches indicate mangrove detected only by the present study (mangrove difference, ~500 ha, mainly along estuarine and riverine fringes), and grey areas are non-mangrove zones.

are developed for Malaysian and regional mangroves, whereas global layers rely on pan-tropical canopy–biomass relationships that smooth species-level wood-density differences and can therefore understate tall Rhizophora–Sonneratia assemblages typical of Sarawak's less disturbed sites. Mean AGB values in Sarawak's mature stands (90–100 Mg ha^{-1}) are consistent with regional estimates and slightly below the global mean of ~115 Mg ha^{-1} reported by Bunting et al. (2022), which is expected given the inclusion of regenerating stands. The prominence of elevation in the SHAP ranking should be interpreted in the context of mangrove ecohydrology rather than as purely topographic control. Across Sarawak's estuaries and deltaic plains, modest elevation gradients correspond to systematic differences in tidal flooding regime, salinity stress, sediment dynamics, and freshwater inputs. These hydrogeomorphic conditions shape stand development by influencing oxygen availability in the root zone, nutrient delivery, and the stability of substrates that support tall canopies (Liu et al., 2021). In this sense, elevation functions as a spatial proxy for the environmental setting that enables (or constrains) biomass accumulation, and it complements canopy height and radar/optical indicators by helping the model distinguish consistently high-biomass zones from frequently inundated or physiologically stressed areas (Simard et al., 2019).

Species-level analysis of AGB revealed clear structural variation among dominant mangrove taxa. *Rhizophora mucronata* exhibited the highest mean biomass **(**~102 Mg ha^{-1}**),** followed closely by *Sonneratia alba* (~100 Mg ha^{-1}), while *Bruguiera parviflora* recorded the lowest **(**~75 Mg ha^{-1}**).** These differences are attributed to inherent species traits such as wood density and growth form. Similar trends have been reported in Peninsular Malaysia, where stands dominated by dwarf or early successional species (e.g., young *Avicennia*) held only ~50% of the biomass of mature stands (Hamdan et al., 2014; Muhd-Ekhzarizal et al., 2018). In Sarawak, the presence of tall, undisturbed *Rhizophora* and *Sonneratia* likely contributes to elevated local AGB, whereas disturbed zones host lower-biomass species. Despite interspecific differences, all mature stands in our dataset, including *Bruguiera*, exceeded 70 Mg ha^{-1}**,** affirming their carbon significance. The full range of observed biomass (~20 to >230 Mg ha^{-1}) highlights the structural heterogeneity of Sarawak's mangroves and supports the use of species-informed strategies for conservation and carbon management (Phua et al., 2023).

The methodological strength of this study lies in the integration of multi-source satellite data with the XGBoost learning algorithm. The fusion of Sentinel-2 optical imagery with Sentinel-1 SAR significantly enhanced both classification and AGB prediction. Optical data captured canopy biochemical properties, while SAR provided structural information unaffected by cloud cover and individual sensor limitations, such as saturation in dense canopies and low backscatter sensitivity in complex forests (Sarker, 2010; Wu et al., 2022). Sensor fusion improved classification accuracy by over 30% compared to SAR alone, as similarly reported by Gegiuc et al. (2018). Our own preliminary trials confirmed that single-sensor models omitted key mangrove features, whereas fusion mitigated these errors. XGBoost further improved model performance by capturing non-linear interactions among variables such as species composition, canopy height, and texture features. This aligns with Ou et al. (2023), who found that gradient boosting outperformed other algorithms in estimating mangrove biomass. Moreover, SHAP-based interpretability revealed that VH backscatter, red-edge, and SWIR bands were among the most influential predictors, consistent with known spectral responses to vegetation biomass. Overall, the fusion XGBoost approach provides a scalable and interpretable framework for high-resolution mangrove carbon assessments across diverse tropical coastlines.

The predictor-set comparison also has practical relevance for teams with different levels of technical capacity. The full Sentinel-1/Sentinel-2 fusion remains the preferred configuration where radar and optical processing are both feasible, because it combines structural, moisture, and spectral information. However, a minimal viable configuration using Sentinel-2 vegetation and moisture indices with basic topographic variables could still support first-order AGB screening where SAR processing expertise, computing resources, or radar-preprocessing workflows are limited. Such a simplified configuration would be most suitable for preliminary biomass stratification, restoration prioritisation, and selection of field-verification sites rather than final carbon accounting. In persistently cloudy areas, or where canopy structure is the main source of biomass variation, Sentinel-1 backscatter combined with elevation and slope offers a stronger low-cost alternative. These results indicate that the workflow can be implemented progressively, from optical–topographic screening to full radar–optical biomass modelling, depending on available data, expertise, and monitoring objectives.

4.1. Limitations

Several limitations should be acknowledged. Signal saturation in dense canopies remained evident, especially above 150–200 Mg ha^{-1}, where optical indices and Sentinel-1 C-band backscatter plateaued (Lucas et al., 2020). Data fusion widened the responsive range, but high-biomass underestimation persists and would likely be reduced with L-band SAR (ALOS-2 PALSAR) **or** spaceborne LiDAR coverage. The use of Sentinel-1 was driven by accessibility, revisit time and reproducibility for large-area monitoring, yet future inclusion of L-band or forthcoming P-band missions (BIOMASS) would better capture tall, closed mangroves. Model transferability is also conditioned by plot data and allometries from western/central Sarawak; as noted by Pham et al. (2020), models trained in one ecological setting may not perform equally in Avicennia-dominated or more disturbed zones. Although 245 plots represented a broad structural gradient, accessibility bias may still underrepresent remote stands; the near-zero Moran's I suggests limited spatial autocorrelation, but a denser, spatially stratified plot network would strengthen regional applicability. Classification errors were concentrated along ecological boundaries where freshwater swamp, agroforestry and Nypa-rich zones had overlapping spectral responses, reflected in mangrove precision of ~86.5%. Ancillary layers such as elevation or tidal-flooding frequency could improve separability in such areas. Sentinel 10 m resolution clearly improved on older global products, yet narrow or highly fragmented fringing mangroves may still be missed; very-high-resolution sensors (PlanetScope, WorldView-3) would reduce this gap. The use of single-date composites limited phenological discrimination; multi-seasonal stacks would help separate evergreen mangroves from deciduous or mixed intertidal vegetation. Overall, saturation, spatial transferability, edge confusion and spatial resolution define the main avenues for refinement. Despite these limitations, the present uncertainty level is sufficient for several policy-facing uses. The 10 m mangrove extent map, classification uncertainty layer, and AGB surface can support restoration prioritisation, identification of high-biomass conservation zones, screening of degraded or regenerating stands, and state-level planning for blue-carbon monitoring. The results are also suitable for improving Tier 2-style emission-factor updates where locally calibrated biomass estimates are preferable to broad default values, particularly when aggregated across districts or management units. However, plot-level carbon accounting, project-scale crediting, legal boundary decisions, and verification of high-value restoration or offset projects still require additional ground calibration, repeated field measurements, and where possible, structural data from LiDAR or longer-wavelength SAR. Thus, the framework is most appropriate for regional stratification and decision support, while site-specific carbon claims should remain supported by field-based validation.

4.2. *Future directions and implications*

Advances in mangrove monitoring will come from pairing current SAR–optical fusion with structural sensors. Integrating LiDAR (e.g. GEDI) would directly constrain canopy height and reduce biomass saturation in dense stands (Lucas et al., 2020), while upcoming L-band missions such as NISAR will improve sensitivity to mangrove structure. Hyperspectral data coupled with LiDAR can sharpen species-level mapping and support species-specific biomass estimation (Fan et al., 2024), and wider availability of missions such as PRISMA and SBGM will make such workflows more routine. Time-series use of Sentinel-1/2 at 5–12-day intervals will enable tracking of mangrove loss, storm impacts and recovery. High-resolution AGB layers developed here provide a credible basis for national carbon accounting, including Malaysia's NDCs and potential REDD + mechanisms, given the high carbon density of mangroves (Bunting et al., 2022). Mapping high-AGB zones supports conservation prioritisation, while low-AGB or regenerating areas can be targeted for Sarawak's ongoing mangrove restoration (~10,000 ha). With Asia holding ~38% of global mangroves (Rahman et al., 2024), replicating an open-data plus machine-learning workflow offers a practical pathway to a regional biomass monitoring network. Comprehensive blue-carbon accounting will ultimately require integration of belowground and soil carbon, which can store carbon over centennial scales; strategies that couple above- and below-ground pools and recognise long-term sequestration are therefore preferable (Rani et al., 2023). Beyond climate mitigation, the biomass maps provide spatial proxies for ecosystem services coastal protection, nursery habitat and biodiversity, supporting broader coastal-zone planning.

5. Conclusion

This study provides the first high-resolution (10 m) mapping of mangrove extent and AGB in Sarawak, Malaysia, using fused Sentinel-1 SAR and Sentinel-2 optical data with an XGBoost machine learning approach. The classification achieved an overall accuracy of 88.8% ($\kappa = 0.776$), while AGB estimation, evaluated using held-out out-of-fold predictions across all field plots ($n = 245$)**,** attained $R^2 = 0.72$**,** RMSE = 28.59 Mg ha^{-1}**,** and MAE = 20.80 Mg ha^{-1} **(**$r = 0.85$**).** The mapped mangrove extent of 1078.7 km^2 and spatially explicit AGB distribution capture the structural heterogeneity across species and sites, with high-biomass zones exceeding 200 Mg ha^{-1}. Model interpretability analysis (SHAP) identified structural variables (e.g., VH backscatter, canopy height proxies) and vegetation indices (NDVI, NDMI) as the most influential predictors, while uncertainty assessment through repeated resampling and cross-validation enhanced the reliability of predictions. These outputs establish a robust baseline for blue carbon accounting in Sarawak. These outputs establish a robust baseline for blue carbon accounting in Sarawak. In practical MRV terms, the mapped extent and AGB layers can be incorporated into agency workflows as annually refreshable emission-factor surfaces, allowing local biomass estimates to update default or static values at district or management-unit scale. The 10 m grid also provides a spatial change-detection layer through which mangrove loss, degradation, regeneration, or restoration gains can be tracked repeatedly using updated Sentinel-1/Sentinel-2 composites. Such outputs can help agencies move from broad inventory reporting toward spatially explicit monitoring, where changes in mangrove area and biomass are linked to conservation priorities, restoration planning, and carbon-accounting updates. The findings have direct policy relevance: high-biomass areas can be prioritised for conservation under Malaysia's Nationally Determined Contributions (NDCs) and potential REDD + projects, while low-biomass zones can be targeted for restoration.The findings have direct policy relevance: high-biomass areas can be prioritised for conservation under Malaysia's Nationally Determined Contributions (NDCs) and potential REDD + projects, while low-biomass zones can be targeted for restoration. The open-data, open-tool methodology offers a cost-effective and transferable framework for mangrove biomass monitoring in other Southeast Asian countries hosting ~38% of the world's mangroves. Future work should integrate LiDAR (GEDI), multi-seasonal imagery, and finer-resolution sensors to improve performance in dense stands and fragmented areas and extend ground sampling to under-represented regions. Including belowground carbon pools, as emphasised by Rani et al. (2023), will further enhance the completeness of blue carbon assessments.

CRediT authorship contribution statement

Abdul Qaiyum Alidin: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Azlizam Aziz:** Writing – review & editing, Visualization, Data curation. **Waseem Razzaq Khan:** Writing – review & editing, Investigation, Formal analysis. **Shazali Johari:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Ethics approval and consent to participate

Not Applicable.

Consent for publication

Not Applicable.

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

Authors acknowledge the support of the Universiti Putra Malaysia.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rsase.2026.102087.

Data availability

Data is provided within the manuscript or supplementary information files

References

Abdul Aziz, A., Dargusch, P., Phinn, S., Ward, A., 2015. Using REDD+ to balance timber production with conservation objectives in a mangrove forest in Malaysia. Ecol. Econ. 120. https://doi.org/10.1016/j.ecolecon.2015.10.014.

Agarwal, Y., Chhikara, R., Rana, S., 2023. Boruta based feature selection model for heart disease prediction. International Journal of Science and Research Archive 10 (1), 768–774. https://doi.org/10.30574/ijsra.2023.10.1.0830.

Alam, M.M., Siwar, C., Toriman, M.E., Molla, R.I., Talib, B., 2012. Climate change induced adaptation by paddy farmers in Malaysia. Mitigation and Adaptation Strategies for Global Change 17 (2), 173–186. https://doi.org/10.1007/s11027-011-9319-5.

Ali, J., Haoran, W., Mehmood, K., Hussain, W., Iftikhar, F., Shahzad, F., Hussain, K., Qun, Y., Zhongkui, J., 2025. Remote sensing and integration of machine learning algorithms for above-ground biomass estimation in Larix principis-rupprechtii Mayr plantations: a case study using Sentinel-2 and Landsat-9 data in northern China. Frontiers in Environmental Science 13, 1577298. https://doi.org/10.3389/fenvs.2025.1577298.

Allais, L., Thibodeau, B., Khan, N.S., Crowe, S.A., Cannicci, S., Not, C., 2024. Salinity, mineralogy, porosity, and hydrodynamics as drivers of carbon burial in urban mangroves from a megacity. Science of the Total Environment 912, 168955. https://doi.org/10.1016/j.scitotenv.2023.168955.

Anees, S.A., Mehmood, K., Khan, W.R., Sajjad, M., Alahmadi, T.A., Alharbi, S.A., Luo, M., 2024a. Integration of machine learning and remote sensing for above ground biomass estimation through Landsat-9 and field data in temperate forests of the Himalayan region. Ecol. Inform. 82. https://doi.org/10.1016/j.ecoinf.2024.102732.

Anees, S.A., Mehmood, K., Khan, W.R., Shahzad, F., Zhran, M., Ayub, R., Alarfaj, A.A., Alharbi, S.A., Liu, Q., 2025a. Spatiotemporal dynamics of vegetation cover: integrative machine learning analysis of multispectral imagery and environmental predictors. Earth Sci. Inf. 18. https://doi.org/10.1007/s12145-024-01673-0.

Anees, S.A., Mehmood, K., Luo, M., Abuelgasim, A., Pan, S.A., Shahzad, F., Muhammad, S., Khan, W.R., 2025b. Advancing forest fire burn severity and vegetation recovery assessments using remote sensing and machine learning approaches. Ecol. Inform., 103446 https://doi.org/10.1016/j.ecoinf.2025.103446.

Anees, S.A., Mehmood, K., Muhammad, S., Luo, M., Shahzad, F., Khan, W.R., 2025c. Spatiotemporal dynamics and environmental drivers of fractional vegetation cover in a semi-arid region using machine learning. Theor. Appl. Climatol. 156 (9), 1–23. https://doi.org/10.1007/s00704-025-05696-5.

Anees, S.A., Mehmood, K., Raza, S.I.H., Pfautsch, S., Shah, M., Jamjareegulgarn, P., Shahzad, F., Alarfaj, A.A., Alharbi, S.A., Khan, W.R., Dube, T., 2025d. Spatiotemporal analysis of surface urban heat island intensity and the role of vegetation in six major Pakistani cities. Ecol. Inform. 85. https://doi.org/10.1016/j.ecoinf.2024.102986.

Anees, S.A., Mehmood, K., Rehman, A., Rehman, N.U., Muhammad, S., Shahzad, F., Hussain, K., Luo, M., Alarfaj, A.A., Alharbi, S.A., Khan, W.R., 2024b. Unveiling fractional vegetation cover dynamics: a spatiotemporal analysis using MODIS NDVI and machine learning. Environ. Sustain. Indic., 100485 https://doi.org/10.1016/j.indic.2024.100485.

Arjasakusuma, S., Kusuma, S.S., Saringatin, S., Wicaksono, P., Mutaqin, B.W., Rafif, R., 2021. Shoreline dynamics in East Java Province, Indonesia, from 2000 to 2019 using multi-sensor remote sensing data. Land 10. https://doi.org/10.3390/land10020100.

Arthur Endsley, K., Kimball, J.S., Reichle, R.H., Watts, J.D., 2020. Satellite monitoring of global surface soil organic carbon dynamics using the SMAP level 4 carbon product. J. Geophys. Res. Biogeosci. 125. https://doi.org/10.1029/2020JG006100.

Aye, W.N., Tong, X., Tun, A.W., 2022. Species diversity, biomass and carbon stock assessment of Kanhlyashay natural mangrove forest. Forests 13. https://doi.org/10.3390/f13071013.

Barenblitt, A., Fatoyinbo, L., Thomas, N., Stovall, A., de Sousa, C., Nwobi, C., Duncanson, L., 2024. Invasion in the Niger Delta: remote sensing of mangrove conversion to invasive Nypa fruticans from 2015 to 2020. Remote Sens. Ecol. Conserv. 10. https://doi.org/10.1002/rse2.353.

Bunting, P., Hilarides, L., Rosenqvist, A., Lucas, R.M., Kuto, E., Gueye, Y., Ndiaye, L., 2023. Global mangrove watch: monthly alerts of mangrove loss for Africa. Remote Sens. 15. https://doi.org/10.3390/rs15082050.

Bunting, P., Rosenqvist, A., Hilarides, L., Lucas, R.M., Thomas, N., Tadono, T., Worthington, T.A., Spalding, M., Murray, N.J., Rebelo, L.M., 2022. Global mangrove extent change 1996–2020: Global mangrove watch version 3.0. Remote Sens. 14. https://doi.org/10.3390/rs14153657.

Bunting, P., Rosenqvist, A., Lucas, R.M., Rebelo, L.M., Hilarides, L., Thomas, N., Hardy, A., Itoh, T., Shimada, M., Finlayson, C.M., 2018. The global mangrove watch - a new 2010 global baseline of mangrove extent. Remote Sens. 10. https://doi.org/10.3390/rs10101669.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Glob. Change Biol. 20 (10), 3177–3190. https://doi.org/10.1111/gcb.12629.

Cherian, S.M., Rajitha, K., 2024. Random forest and support vector machine classifiers for coastal wetland characterization using the combination of features derived from optical data and synthetic aperture radar dataset. J. Water Clim. Change 15. https://doi.org/10.2166/wcc.2023.238.

Chien, S.C., Knoble, C., Krumins, J.A., 2024. Human population density and blue carbon stocks in mangroves soils. Environ. Res. Lett. 19. https://doi.org/10.1088/1748-9326/ad13b6.

Chuine, I., Beaubien, E.G., 2001. Phenology is a major determinant of tree species range. Ecology Letters 4 (5), 500–510. https://doi.org/10.1046/j.1461-0248.2001.00261.x.

Dagne, S.S., Hirpha, H.H., Tekoye, A.T., Dessie, Y.B., Endeshaw, A.A., 2023. Fusion of sentinel-1 SAR and sentinel-2 MSI data for accurate urban land use-land cover classification in Gondar City, Ethiopia. Environ. Syst. Res. 12 (1), 40. https://doi.org/10.1186/s40068-023-00324-5.

Elbakidze, M., Ražauskaitė, R., Manton, M., Angelstam, P., Mozgeris, G., Brūmelis, G., Brazaitis, G., Vogt, P., 2016. The role of forest certification for biodiversity conservation: Lithuania as a case study. European Journal of Forest Research 135 (2), 361–376. https://doi.org/10.1007/s10342-016-0940-4.

Fan, G., Zhang, B., Zhou, J., Wang, R., Xu, Q., Zeng, X., Lu, F., Luo, W., Cai, H., Wang, Y., Dong, Z., Gao, C., 2024. Satellite image fusion airborne LiDAR point-clouds-driven machine learning modeling to predict the carbon stock of typical subtropical plantation in China. Forests 15. https://doi.org/10.3390/f15050751.

Ferreira, A.C., Lacerda, L.D., 2016. Degradation and conservation of Brazilian mangroves, status and perspectives. Ocean & Coastal Management 125, 38–46. https://doi.org/10.1016/j.ocecoaman.2016.03.011.

Gegiuc, A., Similä, M., Karvonen, J., Lensu, M., Mäkynen, M., Vainio, J., 2018. Estimation of degree of sea ice ridging based on dual-polarized C-band SAR data. Cryosphere 12. https://doi.org/10.5194/tc-12-343-2018.

Ghorbanian, A., Zaghian, S., Asiyabi, R.M., Amani, M., Mohammadzadeh, A., Jamali, S., 2021. Mangrove ecosystem mapping using sentinel-1 and sentinel-2 satellite images and random forest algorithm in google earth engine. Remote Sens. 13. https://doi.org/10.3390/rs13132565.

Ghosh, S.M., Behera, M.D., 2021. Aboveground biomass estimates of tropical mangrove forest using Sentinel-1 SAR coherence data - the superiority of deep learning over a semi-empirical model. Comput. Geosci. 150. https://doi.org/10.1016/j.cageo.2021.104737.

Hamdan, O., Khali Aziz, H., Mohd Hasmadi, I., 2014. L-band ALOS PALSAR for biomass estimation of Matang Mangroves, Malaysia. Remote Sens. Environ. 155. https://doi.org/10.1016/j.rse.2014.04.029.

Hemati, M., Mahdianpari, M., Shiri, H., Mohammadimanesh, F., Salehi, B., Wu, J., 2025. Advancing wetland aboveground biomass mapping: insights from high-resolution PlanetScope data in the Mississippi River Delta. Remote Sens. Appl.: Soc. Environ. 38, 101604. https://doi.org/10.1016/j.rsase.2025.101604.

Hu, L., Xu, N., Liang, J., Li, Z., Chen, L., Zhao, F., 2020. Advancing the mapping of mangrove forests at national-scale using Sentinel-1 and Sentinel-2 time-series data with google Earth Engine: a case study in China. Remote Sens. 12. https://doi.org/10.3390/RS12193120.

Huang, K., Yang, G., Yuan, Y., Sun, W., Meng, X., Ge, Y., 2022. Optical and SAR images combined mangrove index based on multi-feature fusion. Sci. Remote Sens. 5. https://doi.org/10.1016/j.srs.2022.100040.

Hussain, K., Mehmood, K., Yujun, S., Badshah, T., Anees, S.A., Shahzad, F., Nooruddin, Ali, J., Bilal, M., 2025. Analysing LULC transformations using remote sensing data: insights from a multilayer perceptron neural network approach. Annals of GIS 31 (3), 473–500. https://doi.org/10.1080/19475683.2024.2343399.

Ihinegbu, C., Mönnich, S., Akukwe, T., 2023. Scientific evidence for the effectiveness of mangrove forests in reducing floods and associated hazards in coastal areas. Climate. https://doi.org/10.3390/cli11040079.

Islam, M.M., Alam, M.J., Maniruzzaman, M., Ahmed, N.A.M.F., Ali, M.S., Rahman, M.J., Roy, D.C., 2023. Predicting the risk of hypertension using machine learning algorithms: A cross sectional study in Ethiopia. PLOS ONE 18 (8), e0289613. https://doi.org/10.1371/journal.pone.0289613.

Jia, M., Wang, Z., Mao, D., Ren, C., Song, K., Zhao, C., Wang, Y., 2023. Mapping global distribution of mangrove forests at 10-m resolution. Sci. Bull. (Taipei) 68 (12), 1306–1316. https://doi.org/10.1016/j.scib.2023.05.004.

Kanniah, K.D., Sheikhi, A., Cracknell, A.P., Goh, H.C., Tan, K.P., Ho, C.S., Rasli, F.N., 2015. Satellite images for monitoring mangrove cover changes in a fast growing economic region in southern Peninsular Malaysia. Remote Sens. 7. https://doi.org/10.3390/rs71114360.

Karvonen, J., Gegiuc, A., Niskanen, T., Montonen, A., Buus-Hinkler, J., Rinne, E., 2022. Iceberg detection in dual-polarized C-Band SAR imagery by segmentation and nonparametric CFAR (SnP-CFAR). IEEE Trans. Geosci. Rem. Sens. 60. https://doi.org/10.1109/TGRS.2021.3070312.

Khan, W.R., Giani, M., Bevilacqua, S., Anees, S.A., Mehmood, K., Nazre, M., Haddy, A.A.B.A., Median, A.N.B.A., Bujang, J.B., Mohamad-Ismail, F.N., Mohamed, J., 2025. Derivation of allometric equations and carbon content estimation in mangrove forests of Malaysia. Environ. Sustain. Indic., 100618

Khan, W.R., Nazre, M., Akram, S., Anees, S.A., Mehmood, K., Ibrahim, F.H., et al., 2024. Assessing the productivity of the Matang mangrove forest reserve: review of one of the best-managed mangrove forests. Forests 15 (5), 747. https://doi.org/10.3390/f15050747.

Komiyama, A., Poungparn, S., Kato, S., 2005. Common allometric equations for estimating the tree weight of mangroves. J. Trop. Ecol. 21 (4), 471–477. https://doi.org/10.1017/S0266467405002476.

Lee, S.L., Chee, S.Y., Huxham, M., Jamilah, M., Choo, J., Kaur, C.R., Amir, A.A., Ooi, J.L.S., Rozaimi, M., Omar, H., Sharma, S., Moritz, M., Then, A.Y.-H., 2024. Blue carbon ecosystems in Malaysia – status, threats, and the way forward for research and policy. J. Environ. Dev. 34 (1), 225–265. https://doi.org/10.1177/10704965241284366 (Original work published 2025).

Lestari, S.D., Leon, F.M., Riyadi, S., Qodariah, Q., Putra, A.H.P.K., 2024. Comparison and implementation of environmental law policies in handling climate change in ASEAN countries: A comparative study of Indonesia, Malaysia, and Thailand. International Journal of Energy Economics and Policy 14 (2), 687–700. https://doi.org/10.32479/ijeep.14998.

Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., Chen, H., 2020. High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data. Int. J. Appl. Earth Obs. Geoinf. 92. https://doi.org/10.1016/j.jag.2020.102163.

Liu, J., Failler, P., Ramrattan, D., 2024. Blue carbon accounting to monitor coastal blue carbon ecosystems. J. Environ. Manag. 352. https://doi.org/10.1016/j.jenvman.2023.120008.

Liu, X., Fatoyinbo, T.E., Thomas, N.M., Guan, W.W., Zhan, Y., Mondal, P., Lagomasino, D., Simard, M., Trettin, C.C., Deo, R., Barenblitt, A., 2021. Large-scale high-resolution coastal mangrove forests mapping across West Africa with machine learning ensemble and satellite big data. Front. Earth Sci. 8. https://doi.org/10.3389/feart.2020.560933.

Lucas, R., Van De Kerchove, R., Otero, V., Lagomasino, D., Fatoyinbo, L., Omar, H., Satyanarayana, B., Dahdouh-Guebas, F., 2020. Structural characterisation of mangrove forests achieved through combining multiple sources of remote sensing data. Remote Sens. Environ. 237. https://doi.org/10.1016/j.rse.2019.111543.

Malik, A., Sideng, U., Jaelani, 2022. Biomass carbon stock assessment of mangrove ecosystem in Pannikiang island South Sulawesi Indonesia. Indones. J. Geogr. 54. https://doi.org/10.22146/ijg.46989.

Martinez, M., Osland, M.J., Grace, J.B., Enwright, N.M., Stagg, C.L., Kaalstad, S., Anderson, G.H., Armitage, A.R., Cebrian, J., et al., 2024. Integrating remote sensing with ground-based observations to quantify the effects of an extreme freeze event on black mangroves (Avicennia germinans) at the landscape scale. Ecosystems 27 (1), 45–60. https://doi.org/10.1007/s10021-023-00871-z.

Mehmood, K., Anees, S.A., Luo, M., Akram, M., Zubair, M., Khan, K.A., Khan, W.R., 2024a. Assessing Chilgoza Pine (Pinus gerardiana) forest fire severity: remote sensing analysis, correlations, and predictive modeling for enhanced management strategies. Trees For. People 16. https://doi.org/10.1016/j.tfp.2024.100521.

Mehmood, K., Anees, S.A., Muhammad, S., Albasher, G., Shahzad, F., Hussain, K., Liu, Q., Ayub, R., Khan, W.R., 2025a. Spatial and temporal vegetation dynamics from 2000 to 2023 in the Western Himalayan regions. Stoch. Environ. Res. Risk Assess. 1–22. https://doi.org/10.1007/s00477-025-02971-9.

Mehmood, K., Anees, S.A., Muhammad, S., Hussain, K., Shahzad, F., Liu, Q., Ansari, M.J., Alharbi, S.A., Khan, W.R., 2024b. Analyzing vegetation health dynamics across seasons and regions through NDVI and climatic variables. Sci. Rep. 14 (1), 11775. https://doi.org/10.1038/s41598-024-62464-7.

Mehmood, K., Anees, S.A., Muhammad, S., Shahzad, F., Liu, Q., Khan, W.R., Shrahili, M., Ansari, M.J., Dube, T., 2025b. Machine learning and spatio temporal analysis for assessing ecological impacts of the billion tree afforestation project. Ecol. Evol. 15 (2), e70736. https://doi.org/10.1002/ece3.70736.

Mehmood, K., Anees, S.A., Rehman, A., Rehman, N.U., Muhammad, S., Shahzad, F., Liu, Q., Alharbi, S.A., Alfarraj, S., Ansari, M.J., 2024c. Assessment of climatic influences on net primary productivity along elevation gradients in temperate ecoregions. Trees For. People 18, 100657.

Mehmood, K., Anees, S.A., Shahzad, F., Muhammad, S., Liu, Q., Khan, W.R., Shah, M., Jamjareegulgarn, P., 2025c. Exploring vegetation health in Southern Thailand under climate stress from temperature and water impacts between 2000 and 2023. Sci. Rep. 15 (1), 30491.

Mo, L., Crowther, T.W., Maynard, D.S., Van den Hoogen, J., Ma, H., Bialic-Murphy, L., Liang, J., De-Miguel, S., Nabuurs, G.J., Reich, P.B., Phillips, O.L., 2024. The global distribution and drivers of wood density and their impact on forest carbon stocks. Nat. Ecol. Evol. 8 (12), 2195–2212. https://doi.org/10.1038/s41559-024-02564-9.

Mohammadpour, P., Viegas, D.X., Viegas, C., 2022. Vegetation mapping with Random Forest using Sentinel-2 and GLCM texture feature—A case study for Lousã region, Portugal. Remote Sensing 14 (18), 4585. https://doi.org/10.3390/rs14184585.

Monga, E., Mangora, M.M., Trettin, C.C., 2022. Impact of mangrove planting on forest biomass carbon and other structural attributes in the Rufiji Delta, Tanzania. Glob. Ecol. Conserv. 35. https://doi.org/10.1016/j.gecco.2022.e02100.

Muhammad, S., Nizami, S.M., Mehmood, K., Hussain, M., Anees, S.A., Shahzad, F., Khan, W.R., 2025. Machine learning-based forest fire vulnerability assessment in subtropical chir pine forests of Pakistan. Fire Ecol. 21 (1), 1–17.

Muhd-Ekhzarizal, M.E., Mohd-Hasmadi, I., Hamdan, O., Mohamad-Roslan, M.K., Noor-Shaila, S., 2018. Estimation of aboveground biomass in mangrove forests using vegetation indices from SPOT-5 image. J. Trop. For. Sci. 30. https://doi.org/10.26525/jtfs2018.30.2.224233.

Nguyen, D.B., Gruber, A., Wagner, W., 2016. Mapping rice extent and cropping scheme in the Mekong Delta using Sentinel-1A data. Remote Sensing Letters 7 (12), 1209–1218. https://doi.org/10.1080/2150704X.2016.1225172.

Oliveira, J.M.D., Oliveira, J.P.M.D., Cardoso, L.S., Ataíde, D.H.S., Curto, R.A., Araújo, E.J.G., 2021. Evaluation of smartphone applications for measuring the total height of trees. Scientific Electronic Archives 14 (2), 7–14. https://doi.org/10.36560/14220211221.

Ou, J., Tian, Y., Zhang, Q., Xie, X., Zhang, Y., Tao, J., Lin, J., 2023. Coupling UAV hyperspectral and LiDAR data for mangrove classification using XGBoost in China's Pinglu Canal Estuary. Forests 14. https://doi.org/10.3390/f14091838.

Pandey, B., Koirala, S., Aung, H., Li, R., Timilsina, A., Htun, S.T.T., Kontny, J., Naing, H.T., 2024. Understanding carbon storage dynamics in Ayeyarwady delta's mangrove ecosystem in Myanmar: insights for restoration efforts. Environ. Res. Commun. 6. https://doi.org/10.1088/2515-7620/ad23f1.

Park, Y., Guldmann, J.-M., 2020. Measuring continuous landscape patterns with Gray-Level Co-Occurrence Matrix (GLCM) indices: An alternative to patch metrics? Ecological Indicators 109, 105802. https://doi.org/10.1016/j.ecolind.2019.105802.

Pesiu, E., Lee, G.E., Salam, M.R., Salim, J.M., Lau, K.H., Yong, J.W.H., Abdullah, M.T., 2022. Species composition, diversity, and biomass estimation in coastal and marine protected areas of Terengganu, peninsular Malaysia. Agronomy 12. https://doi.org/10.3390/agronomy12102380.

Pham, T.D., Le, N.N., Ha, N.T., Nguyen, L.V., Xia, J., Yokoya, N., To, T.T., Trinh, H.X., Kieu, L.Q., Takeuchi, W., 2020. Estimating mangrove above-ground biomass using extreme gradient boosting decision trees algorithm with fused sentinel-2 and ALOS-2 PALSAR-2 data in can Gio biosphere reserve, Vietnam. Remote Sens. 12. https://doi.org/10.3390/rs12050777.

Phua, M.-H., Wong, C.J., James, D., Besar, A., N, 2023. Estimating mangrove above-ground biomass (AGB) in sabah, Malaysia using field measurements, shuttle radar topography mission and landsat data. Borneo Sci. J. Sci. Technol. 41. https://doi.org/10.51200/bsj.v41i2.4447.

Potin, P., Rosich, B., Grimont, P., Miranda, N., Shurmer, I., O'Connell, A., Krassenburg, M., 2016. Sentinel-1 mission status. In: Proceedings of EUSAR 2016: 11Th European Conference on Synthetic Aperture Radar. VDE, pp. 1–6.

Rahman, Ceanturi, A., Tuahatu, J.W., Lokollo, F.F., Supusepa, J., Hulopi, M., Permatahati, Y.I., Lewerissa, Y.A., Wardiatno, Y., 2024. Mangrove ecosystems in Southeast Asia region: mangrove extent, blue carbon potential and CO2 emissions in 1996–2020. Sci. Total Environ. 915. https://doi.org/10.1016/j.scitotenv.2024.170052.

Rani, V., Schwing, P.T., Jayachandran, P.R., Preethy, C.M., Sreelekshmi, S., Joseph, P., Nandan, S.B., 2023. Carbon stocks and sequestration rate in mangroves and its major influencing factors from highly urbanised port city, southern India. J. Environ. Manag. 335, 117542. https://doi.org/10.1016/j.jenvman.2023.117542.

Richards, D.R., Thompson, B.S., Wijedasa, L., 2020. Quantifying net loss of global mangrove carbon stocks from 20 years of land cover change. Nat. Commun. 11. https://doi.org/10.1038/s41467-020-18118-z.

Rodigheri, G., Sanches, I.D.A., Richetti, J., Tsukahara, R.Y., Lawes, R., Bendini, H.N., Adami, M., 2023. Estimating crop sowing and harvesting dates using satellite vegetation index: A comparative analysis. Remote Sensing 15 (22), 5366. https://doi.org/10.3390/rs15225366.

Sahputra, E., Harahap, R.H., Wahyuningsih, H., Utomo, B., 2022. Assessing the sustainability status of mangrove forest ecosystem management by coastal community in Jaring Halus Village, North Sumatra, Indonesia. Biodiversitas 23. https://doi.org/10.13057/biodiv/d230101.

Sainuddin, F.V., Malek, G., Rajwadi, A., Nagar, P.S., Asok, S.V., Reddy, C.S., 2024. Estimating above-ground biomass of the regional forest landscape of northern Western Ghats using machine learning algorithms and multi-sensor remote sensing data. J. Indian Soc. Remote Sens. 52 (4), 885–902. https://doi.org/10.1007/s12524-024-01836-y.

Sarker, MdL.R., 2010. Estimation of Forest Biomass Using Remote Sensing, 26. The Hong Kong Polytechnic University.

Shahzad, F., Mehmood, K., Anees, S.A., Adnan, M., Hussain, K., Khan, W.R., Shah, M., Jamjareegulgarn, P., Oliveira, M., Borges, J.G., 2025. Understanding green house gases emission dynamics from forest fires in Thailand using predictive models. Global Planet. Change, 105236.

Sheehy, J., Porter, J., Bell, M., Kerr, S., 2024. Redefining blue carbon with adaptive valuation for global policy. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2023.168253.

Shukor, A., Hamid, A., n.d. Title the Use of Mangroves in Malaysia.

Simard, M., Fatoyinbo, L., Smetanka, C., et al., 2019. Mangrove canopy height globally related to precipitation, temperature and cyclone frequency. Nat. Geosci. 12, 40–45. https://doi.org/10.1038/s41561-018-0279-1.

Soeprobowati, T.R., Sularto, R.B., Hadiyanto, H., Puryono, S., Rahim, A., Jumari, J., Gell, P., 2024. The carbon stock potential of the restored mangrove ecosystem of Pasarbanggi, Rembang, Central Java. Mar. Environ. Res. 193. https://doi.org/10.1016/j.marenvres.2023.106257.

van Hespen, R., Hu, Z., Borsje, B., De Dominicis, M., Friess, D.A., Jevrejeva, S., Kleinhans, M.G., Maza, M., van Bijsterveldt, C.E.J., Van der Stocken, T., van Wesenbeeck, B., Xie, D., Bouma, T.J., 2023. Mangrove forests as a nature-based solution for coastal flood protection: biophysical and ecological considerations. Water Sci. Eng. 16. https://doi.org/10.1016/j.wse.2022.10.004.

Wolswijk, G., Satyanarayana, B., Le, Q.D., Siau, Y.F., Ali, A.N. Bin, Saliu, I.S., Fisol, M.A. Bin, Gonnelli, C., Dahdouh-Guebas, F., 2020. Mercury concentration data from Matang mangrove forest reserve, Malaysia. Data Brief 29. https://doi.org/10.1016/j.dib.2020.105134.

Wu, W., Shao, Z., Huang, X., Teng, J., Guo, S., Li, D., 2022. Quantifying the sensitivity of SAR and optical images three-level fusions in land cover classification to registration errors. Int. J. Appl. Earth Obs. Geoinf. 112. https://doi.org/10.1016/j.jag.2022.102868.

Xiong, Y., Dai, Z., Long, C., Liang, X., Lou, Y., Mei, X., Nguyen, B.A., Cheng, J., 2024. Machine learning-based examination of recent mangrove forest changes in the western Irrawaddy River Delta, Southeast Asia. Catena 234. https://doi.org/10.1016/j.catena.2023.107601.

Yilmaz, O.S., Acar, U., Sanli, F.B., Gulgen, F., Ates, A.M., 2023. Mapping burn severity and monitoring CO content in Türkiye's 2021 wildfires, using Sentinel-2 and Sentinel-5P satellite data on the GEE platform. Earth Science Informatics 16 (1), 221–240. https://doi.org/10.1007/s12145-023-00933-9.

Zhang, Y., Zou, Y., Wang, Y., 2025. Remote sensing of forest above-ground biomass dynamics: a review. Forests 16 (5), 821. https://doi.org/10.3390/f16050821.