Ecological Informatics 82 (2024) 102732

Contents lists available at ScienceDirect

Ecological Informatics

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

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

Shoaib Ahmad Anees [a,b,**,1], Kaleem Mehmood [c,d,e,1], Waseem Razzaq Khan [f,g,h], Muhammad Sajjad [i], Tahani Awad Alahmadi [j], Sulaiman Ali Alharbi [k], Mi Luo [a,*]

[a] Key Laboratory of Environment Change and Resources Use in Beibu Gulf, Ministry of Education, Nanning Normal University, Nanning 530001, PR China
[b] Department of Forestry, The University of Agriculture, Dera Ismail Khan 29050, Pakistan
[c] College of Forestry, Beijing Forestry University, Beijing 100083, PR China
[d] Institute of Forest Science, University of Swat, Main Campus Charbagh, 19120 Swat, Pakistan
[e] Key Laboratory for Silviculture and Conservation of Ministry of Education, Beijing Forestry University, Beijing 100083, PR China
[f] Department of Forestry Science and Biodiversity, Faculty of Forestry and Environment, Universiti Putra Malaysia UPM, Serdang 43400, Selangor, Malaysia
[g] Advanced Master in Sustainable Blue Economy, National Institute of Oceanography and Applied Geophysics - OGS, University of Trieste, Trieste 34127, Italy
[h] Institut Ekosains Borneo (IEB), Universiti Putra Malaysia, Bintulu Campus, Sarawak 97008, Malaysia
[i] Department of Geography, Hong Kong Baptist University, Hong Kong Special Administrative Region
[j] Department of Pediatrics, College of Medicine and King Khalid University Hospital, King Saud University, Medical City, PO Box-2925, Riyadh 11461, Saudi Arabia
[k] Department of Botany and Microbiology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

ARTICLE INFO

Keywords:
Himalayas
Machine learning
Field data
Remote sensing
Aboveground biomass
Random forests

ABSTRACT

Accurately estimating aboveground biomass (AGB) in forest ecosystems facilitates efficient resource management, carbon accounting, and conservation efforts. This study examines the relationship between predictors from Landsat-9 remote sensing data and several topographical features. While Landsat-9 provides reliable data crucial for long-term monitoring, it is part of a broader suite of available remote sensing technologies. We employ machine learning algorithms such as Extreme Gradient Boosting (XGBoost), Support Vector Regression (SVR), and Random Forest (RF), alongside linear regression techniques like Multiple Linear Regression (MLR). The primary objectives of this study encompass two key aspects. Firstly, the research methodically selects optimal predictor combinations from four distinct variable groups: Landsat-9 (L1) data, a fusion of Landsat-9 data and Vegetation-based indices (L2), and the integration of Landsat-9 data with the Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM) (L3) and the combination of best predictors (L4) derived from L1, L2, and L3. Secondly, the research systematically assesses the effectiveness of different algorithms to identify the most precise method for establishing any potential relationship between field-measured AGB and predictor variables. Our study revealed that the Random Forest (RF) model was the most efficient method utilizing Landsat-9 OLI and SRTM DEM (L3) predictors, achieving remarkable accuracy. This conclusion was reached by assessing its outstanding performance when compared to an independent validation dataset. The RF model exhibited remarkable accuracy, presenting relative mean absolute error (RMAE), relative root mean square error (RRMSE), and R^2 values of 14.33%, 22.23%, and 0.81, respectively. The XGBoost model is the subsequent choice with RMAE, RRMSE, and R^2 values of 15.54%, 23.85%, and 0.77, respectively. The study further highlights the significance of specific spectral bands, notably B4 and B5 from Landsat 9 OLI data, in capturing spatial AGB distribution patterns. Integration of vegetation-based indices, including TNDVI, NDVI, RVI, and GNDVI, further refines AGB mapping precision. Elevation, slope, and the Topographic Wetness Index (TWI) are crucial proxies for representing biophysical and biological mechanisms impacting AGB. Through the utilization of openly accessible fine-resolution data and employing the RF algorithm, the research demonstrated promising outcomes

* Corresponding author.
** Corresponding author at: Key Laboratory of Environment Change and Resources Use in Beibu Gulf, Ministry of Education, Nanning Normal University, Nanning 530001, PR China.
E-mail addresses: anees.shoaib@gmail.com (S.A. Anees), luomi0910@163.com (M. Luo).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ecoinf.2024.102732
Received 21 April 2024; Received in revised form 17 July 2024; Accepted 18 July 2024
Available online 22 July 2024
1574-9541/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

in the identification of optimal predictor-algorithm combinations for forest AGB mapping. This comprehensive approach offers a valuable avenue for informed decision-making in forest management, carbon assessment, and ecological monitoring initiatives.

1. Introduction

Forests are crucial as they provide essential habitats and support a wide range of biodiversity. They play a vital role in maintaining ecological balance and supporting biodiversity (Anees et al., 2022b; S. Chen et al., 2022; Mi et al., 2021; Xie et al., 2020). An important factor in determining how well a forest ecosystem can store carbon and keep a positive carbon balance is the estimation of the forest's biomass (Baul et al., 2021; Gogoi et al., 2022; Haider et al., 2017; Jallat et al., 2021). The accurate estimate of forest biomass is particularly crucial when investigating the terrestrial carbon cycle on a large scale, given its essentiality to the understanding of several issues, including climate change, vegetation health, service regularization, etc. (Andreevich et al., 2020; Anees et al., 2022a; Luo et al., 2024; Mehmood et al., 2024c; Usoltsev et al., 2020; Usoltsev et al., 2022; Xiao et al., 2019; Zhao et al., 2022).

Traditional field measurements or remote sensing methods commonly assess AGB in forests (Luo et al., 2024; Xiao et al., 2019; Yu et al., 2022). Direct field measurements are more accurate for smaller forest areas (Khan et al., 2020; Trautenmüller et al., 2023) but are impracticable for regional-scale evaluations due to their higher costs, labor-intensive nature, and time constraints (Yu et al., 2022; Zaher et al., 2020). Utilizing satellite imaging presents a distinct advantage over traditional forest inventories and surveys conducted with Light Detection and Ranging (LiDAR) equipment because of its capability to affordably and effectively cover larger geographical areas (Ecke et al., 2022; Puletti et al., 2020). Incorporating reference values with satellite data is a standard method for increasing the accuracy of airborne LiDAR-derived AGB estimates or precise forest inventories (Do et al., 2022; Ecke et al., 2022). Using spatial prediction algorithms to produce accurate spatial distributions of AGB comes next (Jiang et al., 2022; Wai et al., 2022a). Substantial advancements in mapping forest AGB have been achieved through the integration of modeling techniques and improved predictor utilization derived from satellite data (Behera et al., 2023; L. Chen et al., 2023a; Luo et al., 2024). These developments are the result of the integration of numerous remote sensing techniques. Previous research, for example, has indicated that remote sensing is proficient in accurately measuring and monitoring forest biomass on a regional scale (Behera et al., 2023; Galidaki et al., 2017; Lu et al., 2016a; Timothy et al., 2016; Zhang and Shao, 2021; Y. Zhu et al., 2020a; Luo et al., 2024). As a result, existing research has shown that various remote sensors, including passive and active sensor technologies, can estimate AGB in an area of interest (Luo et al., 2024).

Optical remote sensing images, which present spatial range, spectral, and temporal resolutions, are a frequently utilized tool for estimating AGB across various scales (Doughty et al., 2021; Naik et al., 2021; Tsitsi, 2016). Estimating AGB on a large scale, such as at a global, continental, or country level, commonly involves leveraging moderate- and coarse-resolution data, like that available through the Moderate Resolution Imaging Spectroradiometer (MODIS) (Rodríguez-Veiga et al., 2017; Yuan et al., 2016). Conversely, estimating AGB at local scales typically involves the utilization of medium-resolution data, like that obtained from Sentinel-2 and Landsat satellites (Yingchang Li et al., 2020a; Purohit et al., 2021a; Wang et al., 2019). To estimate AGB at the level of forest stands, utilizing higher-resolution commercial satellite data, such as the imagery provided by IKONOS, QuickBird, and WorldView-2, proves to be well-matched for this purpose (Mohd Zaki and Abd Latif, 2017; Suratman et al., 2023). Moreover, estimating AGB on regional scales with intermediate spatial resolutions involves leveraging microwave radar remote sensing data, such as synthetic aperture radar (SAR), interferometric SAR (InSAR), and polarimetric InSAR (PolInSAR) data (Kaasalainen et al., 2015; Quegan et al., 2019; Sinha et al., 2015).

The first and critical step in developing precise AGB estimation models is choosing the optimal algorithm (Fan et al., 2022; Gao et al., 2018a; Y. Li et al., 2019a; Lu et al., 2016b; Sinha et al., 2015). While the traditional statistical regression method is simple and easy to calculate, it has been commonly employed for AGB estimation in previous studies (Calders et al., 2015; C. Li et al., 2019b; Zhu and Liu, 2015). This technique fits a regression model using test data and remote sensing characteristics (Meng et al., 2016). However, it fails to adequately capture the intricate nonlinear interaction between forest AGB and remote sensing data (Gao et al., 2018b; Rana et al., 2016). Additionally, non-parametric models, interpolation methods, and geostatistical methods like kriging are frequently used in predicting and mapping AGB. The variation of AGB data has been investigated using geostatistical methods, and the best sampling designs for satellite images and on-site forest inventories have been developed (Li et al., 2020b; Li et al., 2020; Su et al., 2020a).

In large hilly terrains, accurately mapping continuous forest characteristics poses challenges. Nevertheless, key geographical elements like soil type, texture, nutrient availability, solar radiation, soil moisture, and water retention capacity strongly impact fundamental tree attributes within distinct stand types, including diameter, height, and volume. Forest tree parameters and AGB assessed during the inventory establishment demonstrate spatial autocorrelation within constrained areas of stand types (Bruenig, 2016; Oliveras and Malhi, 2016). However, this spatial autocorrelation varies in diverse topographical situations, places where communities live, and locations where commercial logging is practiced (Gibson, 2018). In the estimation of AGB, numerous studies have emphasized the amalgamation of remote sensing techniques with geostatistical and machine learning approaches (Li et al., 2020; X. Zhu and Liu, 2015; Luo et al., 2024). This combination proves particularly useful for predicting broad areas with varied bioclimatic conditions and uneven topography (Luo et al., 2024; Su et al., 2020b; Wai et al., 2022b).

Remote sensing-based AGB estimation makes use of machine learning techniques, including decision trees, RF, and SVR. These methods enhance the model's capability to accurately estimate biomass, particularly in scenarios where nonlinearity is a significant factor. Existing literature shows that decision tree-based algorithms with remarkable performance in biomass estimation modeling include RF and Gradient boosting (GB) (Cameron et al., 2022; T. Chen et al., 2023b; Liu et al., 2020). Furthermore, machine learning algorithms include many modifiable and hyperparameters that substantially impact the models. The process of adjusting these parameters has occasionally been disregarded. Earlier research has indicated that the responsiveness of parameters differs in stochastic gradient boosting and RF, and the models' performance is significantly affected by the tuning process (Freeman et al., 2016a; Yingchang Li et al., 2020a; Prakash et al., 2022).

Building upon the preceding discussion, the specific aims of this study are to address the following elements: We aim to identify the optimal combination of predictors from four different sets of variables: Landsat-9 spectral bands (L1), the combined Landsat-9 and Vegetation indices (L2), the joint use of Landsat-9 series data and SRTM DEM (L1 + SRTM DEM) (L3), and incorporating the sets of all best variables among L1, L2, and L3 (L4). This selection process ensures that we leverage the most informative variables for accurate AGB estimation. We also seek to establish the most accurate algorithm for correlating field-acquired AGB measurements with the predictor groups. This involves evaluating machine learning algorithms such as XGBoost, SVR, and RF alongside linear

regression methods (MLR) to determine which model provides the highest predictive accuracy. Our objective is to find the most effective combination of predictors and algorithms suitable for mapping AGB in forest environments. This will help in creating reliable and precise spatial distributions of AGB, essential for forest management and conservation efforts.

In short, this work focuses on improving AGB prediction by identifying the most reliable predictors, contrasting various modeling approaches, examining data source synergy, and offering insights for AGB mapping. By doing so, the present study delivers insights into machine learning-based AGB estimation and provides a unique set of conditioning factors influencing the distribution of AGB in temperate forest regions. This research goes beyond the geographical boundaries of the study region (i.e., the Himalayan area) by identifying variables crucial for AGB sustainability. Although the study primarily examines temperate forests within a specific ecological area, the methodologies and models developed possess global applicability, particularly in other temperate regions. The efficacy of these methodologies enhances our understanding of AGB dynamics, aiding in the development of sustainable policies, carbon accounting methodologies, and well-informed decision-making procedures in forestry and environmental management on a broader scale. This research provides essential tools for sustainable resource management and conservation efforts worldwide by demonstrating the effectiveness of remote sensing and machine learning techniques in temperate forests.

2. Materials and methods

2.1. Study area

The study area, which is located in Pakistan's geographically elevated Hindu Kush Himalayan (HKH) region, was carefully considered. The study focuses on District Swat in the Himalayan Region, well-known for its abundant biodiversity and distinctive ecological characteristics (Bahadur et al., 2023; Rahman et al., 2021). The Kalam Forest Division in this district spans from 34°0.77′N, 72°. 56′ E to 35°0.54′N, 73°. 52′ E (Fig. 1). The total area of the study site is approximately 6606.25 ha. The diverse land use (Hussain et al., 2024a) and vegetation types (Mehmood et al., 2024c) create an intriguing framework for examining approaches to estimate AGB (Abbas et al., 2020a). In addition, the temperate coniferous forests in the Kalam Forest Division have significant cultural and economic importance for indigenous communities. This study underscores the crucial requirement for reliable estimation of AGB to inform sustainable resource management strategies and provide advantages to the residents. The climatic intricacy of the region, which is marked by arid temperate conditions, seasonal patterns of precipitation, and snowfall, is of noteworthy significance (Anwar et al., 2019). These variables substantially impact the distribution of AGB and, therefore, need to be considered in modeling endeavors. The significance of comprehending the variations in AGB is further underscored by the dynamics of seasonal migration, which largely occur as a response to unfavorable winter weather conditions (Abbas et al., 2020b; Anees et al., 2024; Pan et al., 2023; Shobairi et al., 2022). The constant environmental conditions in the region, which are minimally affected by monsoonal influences, provide an ideal setting for conducting rigorous scientific research on estimating AGB. To summarize, the choice of the study area is based on its ecological diversity (Akram et al., 2022; Aslam et al., 2022), cultural importance, climatic intricacies, and potential to provide valuable knowledge on AGB estimation. This research endeavor aims to contribute to sustainable forest management and the overall welfare of indigenous communities.

2.2. Forest inventory and biomass estimation

The field research for forest inventory was carried out between June–September 2022. Sites for sampling were carefully selected while avoiding non-forest areas, using a random distribution method. In total, 371 sampling plots were set up in dry temperate, Kalam Forest,

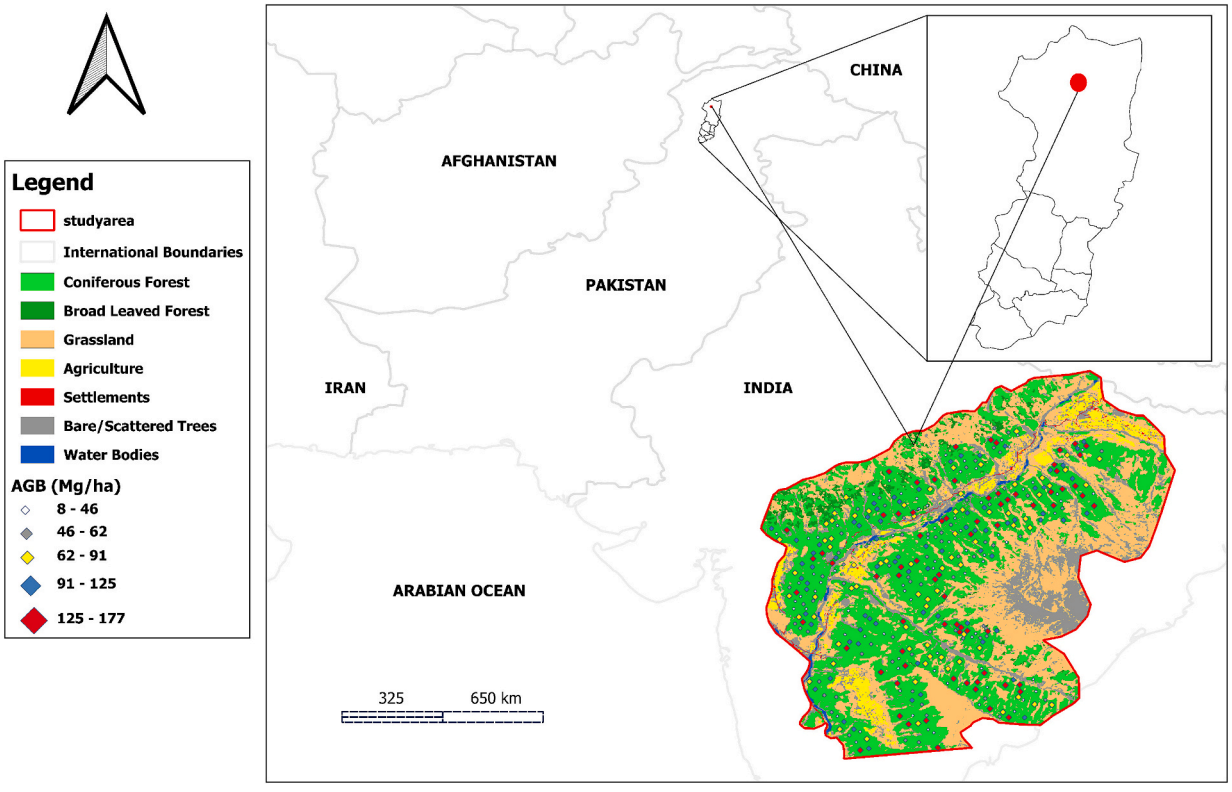

Fig. 1. The study area's location, LULC classification, and the distribution of observed AGB values within the field plots.

Himalayan Temperate Zone, to calculate AGB. The circular plots, which followed UNFCCC requirements, had a radius of 17.84 m and constituted a 0.1-ha area (Ellis and Moarif, 2015; Khan et al., 2021). Stratified random sampling was used to choose the plots, and time, resources, accessibility, and terrain were all considered during the field work. The stratified random sampling was employed to guarantee adequate representation of the diverse forest types and ecological variations in the Himalayan Temperate Zone. To retrieve remote sensing parameters from field plots, we used the plot center coordinates and created an 18-m buffer around each plot center to minimize the impact of spatial inconsistencies. The mean pixel value within the buffer zone encircling each plot center was then calculated to represent the plot's characteristic values accurately. This approach ensures that the buffer adequately represents the plot in the GIS environment, addressing any potential spatial misalignment between the plot boundaries and the pixel grid in the remote sensing data (Maharjan, 2012). Geodetic coordinates and elevation were initially obtained by GPS technology. Consequently, a digital elevation model was employed to derive aspects based on these coordinates. For accurate results, it is underlined that field measurement errors should be fixed before data processing (Erasmi et al., 2014; Khan et al., 2024).

According to the standard field measurement standard, the Diameter at Breast Height (DBH) measurement was made at a height of 4.5 ft, or 1.3 m, above the ground for trees with a diameter of less than 5 cm (Magarik, 2021; Yimam and Kifle, 2020). The Haga Altimeter was used to calculate the height of trees using trigonometric calculations based on the top and base of the trees and a certain distance. The Sunnto Compass was utilized to ensure accurate directional alignment across plots, while a standard measuring tape was employed to measure distances both between plots and within individual plots (Ojoatre et al., 2019; Usman et al., 2022).

We used site-specific allometric equations to estimate the biomass of the tree species that were inventoried. Table 1 delineates the allometric equations employed to facilitate AGB assessment based on the provided field data. The allometric equations pertinent to the estimation of tree height (expressed in meters) and AGB (quantified in kilograms) for the botanical species, namely *Pinus roxburghii*, *Pinus wallichiana*, *Picea smithiana*, *Abies pindrow*, and *Cedrus deodara* were formulated through meticulous research conducted within the domain of forestry (Ismail et al., 2018; Khan et al., 2021). It is essential to highlight that these equations are only valid in this context since they were expressly designed to account for the distinctive qualities of each species and the research location (Ismail et al., 2018). Additionally, given that our plots were one-tenth of a hectare in size, we selected each plot's AGB by a factor of 10 to convert it from kilograms per plot to kilograms per hectare (0.1 ha). The analysis carefully omitted plots on non-forestry terrains, such as agriculture, water regions, and construction zones. The overall number of plots in the final dataset used in this study was 371. In a broader sense, it was observed that the computed average AGB for the study region was 49.44 megagrams per hectare (Mg/ha), with a standard deviation (SD) of 41.75 Mg/ha and falling within a range of 8 to 176.67 Mg/ha. This thorough field research provides insightful information on biomass distribution throughout the assessed forest areas (Fig. 2). The locations of the plots are shown in Fig. 1.

Table 1
Allometric Equations (Ismail et al., 2018).

Specie Names (Scientific Name)	Specie Names (Common Name)	Biomass Equation	Height Equation
Pinus roxburghii	Chir pine or longleaf Indian pine	$Y=0.0224 (D^2\times H)^{0.9767}$	H = −0.0044* D^2 + 0.6863*D-0.7196
Pinus wallichiana	Himalayan pine or Himalayan white pine	$Y=0.0631* (D^2\times H)^{0.8798}$	H = −28.244 + 14.456lnD
Picea smithiana	Morinda spruce (also Western Himalayan spruce	$Y=0.0843* (D^2\times H)^{0.8472}$	H = −23.491 + 12.555lnD
Abies pindrow	Pindrow fir or west Himalayan fir,	$Y=0.0954* (D^2\times H)^{0.8114}$	H = −11.394 + 9.727lnD
Cedrus deodara	Himalayan cedar, or deodar	$Y=0.1779* (D^2\times H)^{0.8103}$	H = −34.394 + 15.355 lnD

Where, Y is the dry biomass in Kgs; D is DBH in cm; H is tree height in meters; Ln is the natural log.

2.3. Satellite data pre-processing and derivation of variables

For this study, we accessed Landsat 9 Operational Land Imager (OLI) satellite data through the USGS Earth Explorer data portal at https://earthexplorer.usgs.gov/. These data are crucial to our analysis and were taken from June 2023 Landsat 9 collections. Notably, the obtained images showed satisfactorily minimal cloud covering (i.e., below the threshold of 10%). The Landsat 9 dataset, known for its reliability in providing consistent data, underwent rigorous processing including orthorectification, georectification, and registration. This dataset is part of a broader array of remote sensing technologies used in environmental studies (Frazier and Hemingway, 2021). Environment for Visualizing Images (ENVI) was used for the per-processing of this dataset. Each band in the Multispectral Image (MSI) dataset, which has an ensemble of 7 spectral bands and a spatial resolution of 30 m, was selected (Masek et al., 2020; Wulder et al., 2022). The USGS repository (https://earthexplorer.usgs.gov/), which also provides SRTM DEM data rendered at a resolution of 30 m, was used to supplement our dataset. Fig. 3 clarifies the entire data processing. Flowchart. (See Fig. 4.)

It is crucial to overcome errors in the estimation of AGB that result from modeling variables collected from remote sensing (Battude et al., 2016; Chen et al., 2015). Previous studies have demonstrated that the higher spatial resolution inherent to the Landsat series significantly reduces uncertainties and improves the precision of AGB mapping, mainly when focusing on finer spatial scales (Montesano et al., 2015; Vaglio Laurin et al., 2016). Using the Landsat dataset series has shown encouraging results, demonstrating the ability to improve biomass estimation by utilizing derived vegetation indicators (Mehmood et al., 2024c; Mourad et al., 2020; Zhu et al., 2020b). It is crucial to recognize the complex interactions between vegetation indices (Anees et al., 2022a; Mehmood et al., 2024c), biophysical factors, and reflectance in the context of spectral sensitivity since they can significantly affect the uncertainties associated with AGB estimates (David et al., 2022; Luo et al., 2024; Sa and Fan, 2023). In this regard, existing literature highlights the exceptional sensitivity of the Landsat red edge band to phenological changes within vegetative ecosystems, positing it as an excellent tool for reducing uncertainty. To enable a thorough evaluation, we carefully chose and extracted thirteen (13) variables that were deemed essential in reducing the inherent errors in forest AGB estimate, paving for insightful comparative studies (Chasmer et al., 2020; Furniss et al., 2020; Nelson et al., 2022; Sun et al., 2023) (Table 2).

2.4. Extraction of remote sensor parameters from field plots

Several methodological procedures were involved in retrieving remote sensing parameters from field plots (Gandhi et al., 2015; He et al., 2014; Malakar et al., 2018). The southwest vertex coordinates of the plot were used to get the spatial coordinates of each plot's center. Despite efforts to resample the remote sensing images to match the field plot dimensions, there were differences between the plot boundary and the associated pixel grid. A circle with a radius of ~18 m was made around each plot center to commence the analysis. The goal of this buffer was to lessen the impact of spatial inconsistencies brought about by the misalignments mentioned above. This buffer radius was chosen to balance minimizing location-based distortions and limiting incursions into regions with clearly different tree canopy cover. The mean pixel value within the buffer zone encircling each plot center was then

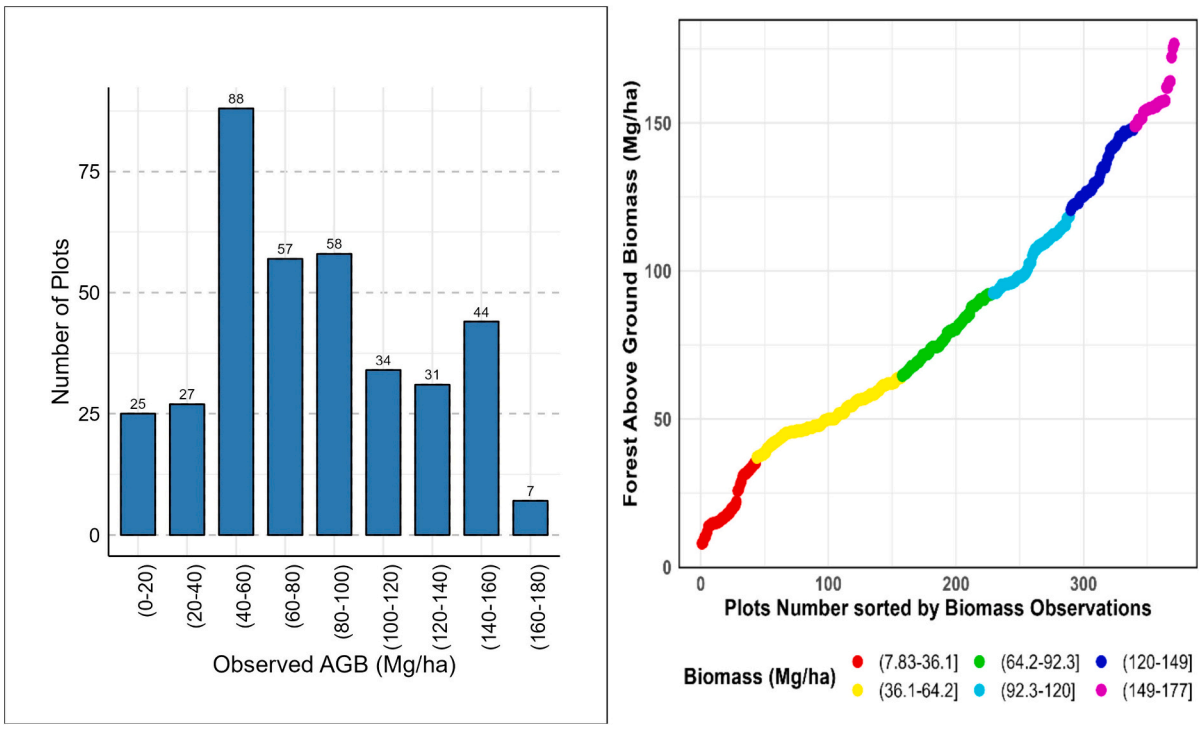

Fig. 2. Presenting the recorded values of AGB involved illustrating field plot profiles depicting AGB variations across the study site, encompassing Plot 1 through Plot 371.

calculated. The representative measurement of the plot's characteristic values was used to retrieve the mean pixel value (Malakar et al., 2018; Turner et al., 2015).

A wide range of image data was included in collecting remote sensing predictor variables. Primary images were generated from Landsat bands and SRTM DEM. Additionally, seven vegetation indexes and six SRTM DEM-based derived predictors were included. Table 3 contains more information about the makeup of the predictor variables. Notably, different predictor variables were used for the model development, including the Landsat, SRTM DEM, and Vegetation indices. Subsequently, a fusion of these datasets was adopted to enhance the predictive capacity of the developed models.

2.5. Modeling techniques and assessment

The machine learning algorithms were chosen for their ability to address the complex nature of AGB prediction in forest ecosystems, characterized by nonlinear variable relationships, high-dimensional data, and multifaceted influencing factors. Their ensemble, regularization, and outlier robustness make them well-suited choices for this task. A paired study using Pearson's product-moment correlation was first conducted to determine the relationship between the indices from multisensory sources and the field-measured AGB (Chen et al., 2018). Concurrently, multi-collinearity was assessed to avoid any bias in the input dataset (Supplementary Table 1). In this analysis, we calculated the variance inflation factor (VIF) to identify and subsequently eliminate any redundancy present within the variables, if detected (Thompson et al., 2017; Vahedi, 2014). In the context of regression analysis, variables that showed strong correlation (with a coefficient magnitude exceeding or equal to 0.8) and those that displayed high VIF values (VIF exceeding or equal to 10) were systematically removed from the pool of predictors (Kristensen et al., 2015; Mehmood et al., 2024a; Pérez-Girón et al., 2020). These analytical processes were carried out using R Statistical Computation software. Four separate groupings (i.e., L1, L2, L3, and L4) were produced by combining the Landsat and SRTM DEM data (Table 2). The first collection of variables, designated as L1, included data from Landsat-9 spectral bands (Band-2, Band3, Band4 and Band 5). The second group of variables, L2, comprised derived vegetative indices and the multispectral bands of Landsat-9 (L1). The topographic indices extrapolated from the SRTM DEM dataset included the third assemblage, L3, with the composite representation of the L1 group. Finally, the L4 set of variables combined the best performance predictors from the L1, L2, and L3 sets of predictors. The comprehensive details regarding the combination of predictors for L4 are presented in Table 3.

2.6. Enumeration of tested algorithms

We used machine learning methods like XGBoost, SVR, and RF because they can effectively capture complex and nonlinear associations within environmental systems such as forest ecosystems (Ding et al., 2021). These algorithms possess the capacity to automatically identify relevant predictor variables and effectively handle datasets with a high number of dimensions. Their robustness, ensemble capabilities, and advanced techniques align with our goal of precise AGB estimation (Bouras et al., 2021; Taghizadeh-Mehrjardi et al., 2020). The rigorous assessment of various algorithms enhances scientific rigor, guaranteeing the optimal selection of a methodology for estimating AGB in temperate forests.

The details of the four evaluated models included in the collection of algorithms are presented in Table 4. For instance, we included the MLR strategy within linear regression Eq. (1). Modeling the association between a dependent variable and two or more independent variables is done statistically using MLR (Hu et al., 2019; Mehmood et al., 2024a). This method presupposes that the dependent and independent variables have a linear relationship. Inferred from this is the relationship between changes in the independent variables and proportional changes in the dependent variable (Krzywinski and Altman, 2015; Mehmood et al., 2024a; Sreehari and Srivastava, 2018). The mathematical expression for the MLR model can be written as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \quad (1)$$

Where, Y is the response variable, X_1, X_2, …, X are the predictor

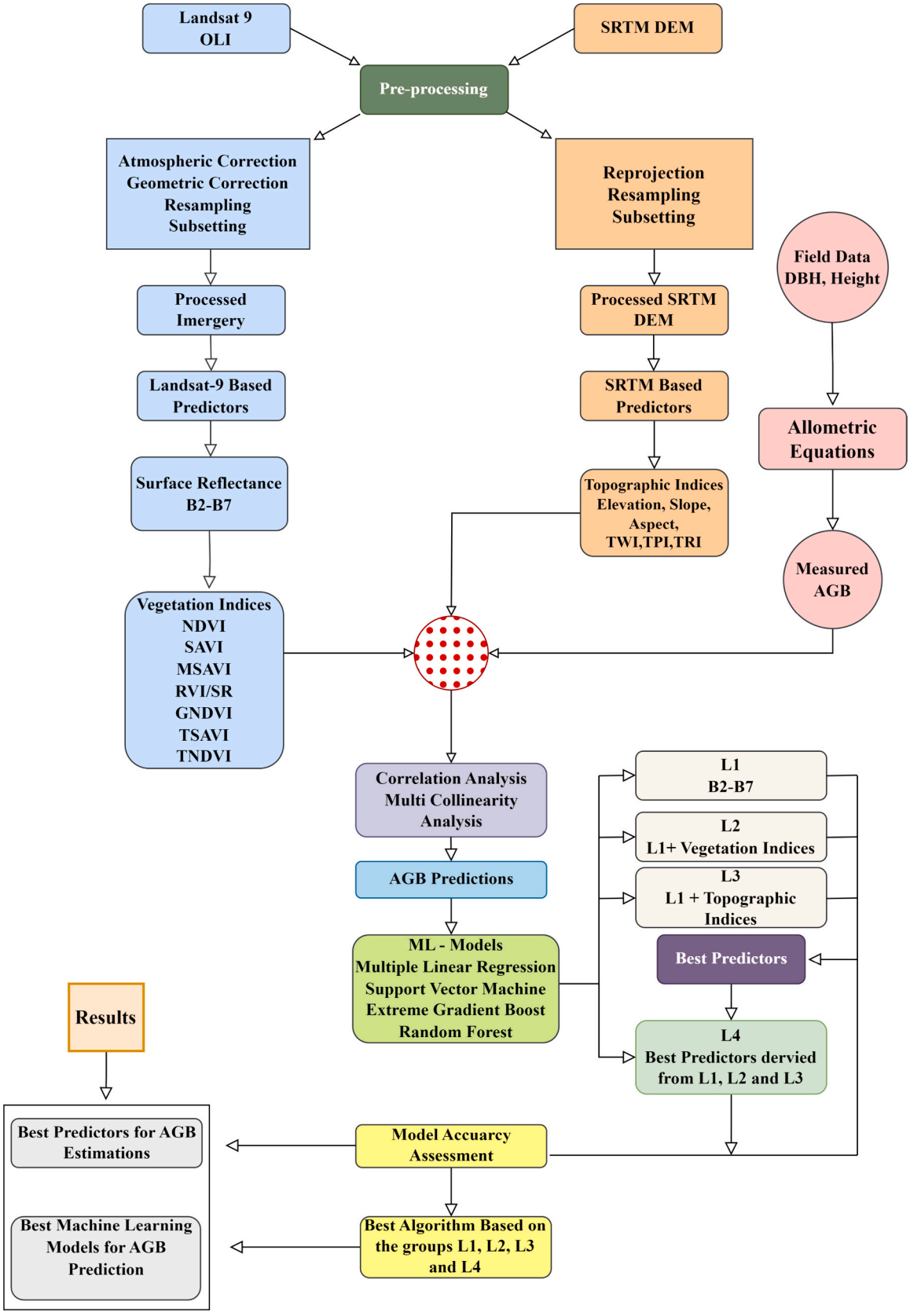

Fig. 3. Flowchart for best predictor combination and modeling algorithms estimation for forest AGB mapping using Landsat 9 and the SRTM DEM.

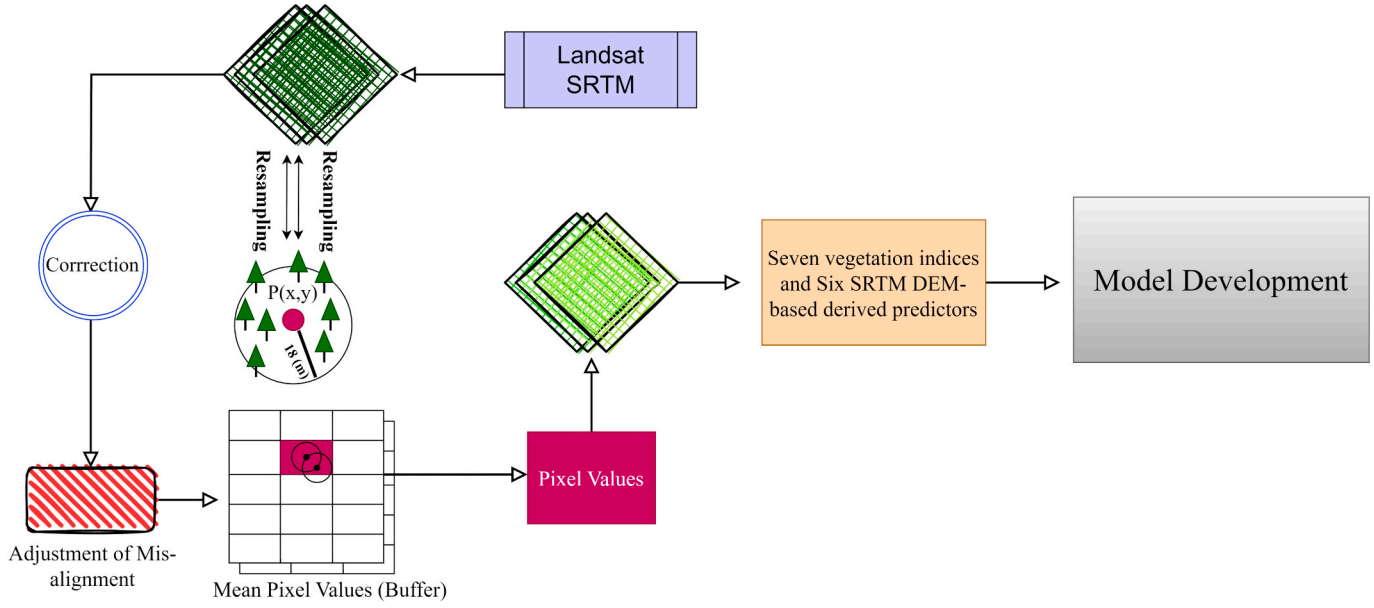

Fig. 4. Flowchart illustrating the extraction of remote sensing parameters from field plots. The process involves determining plot center coordinates, resampling images, adjusting for misalignments, creating an 18-m buffer, and calculating mean pixel values. These values are used to collect predictor variables from Landsat bands, SRTM DEM, and vegetation indexes for model development.

variables, β_0 is the intercept, β_1, β_2, …, β are the coefficients associated with each independent variable, and ε represents the error term.

2.7. Machine learning methods

RF is an ensemble learning method that combines multiple decision trees using a technique called bagging. This method builds each tree using a bootstrap sample of approximately 90% of the training data, with the remaining data used to calculate the out-of-bag (OOB) error. At each node, a random subset of explanatory variables is selected to find the optimal split. RF is widely used for its robustness in both classification and regression tasks, such as environmental modeling and habitat suitability analysis (Luo et al., 2024; Mehmood et al., 2024d; Teng et al., 2023; Zhou et al., 2016). XGBoost is a powerful ensemble learning algorithm that combines gradient boosting with advanced regularization techniques to enhance predictive performance. It sequentially builds models to correct errors from previous iterations, optimizing a user-defined loss function through second-order Taylor expansion. XGBoost's ability to handle missing data and prevent overfitting through regularization makes it highly effective for complex predictive modeling tasks (Hengl et al., 2017; Luo et al., 2021c; Luo et al., 2024; Mehmood et al., 2024a; Strandberg and Låås, 2019). SVR is a regression technique that uses kernel functions to map input data into higher-dimensional spaces where a hyperplane is fitted to minimize prediction error. This method is particularly effective for capturing nonlinear relationships in data. SVR balances model complexity and empirical risk, making it suitable for a wide range of regression applications (Lee et al., 2020; Mehmood et al., 2024a; Sharifi et al., 2016).

2.8. Optimizing model parameters

RF parameters include "ntree," which controls the number of trees, and "mtry," which determines the number of variables sampled at each split. Optimal values for these parameters were identified using grid search. Variable importance was assessed using "percent IncMSE" and "IncNodePurity" indices (Freeman et al., 2016b; Luo et al., 2024; Probst et al., 2019). Key parameters for tuning XGBoost include "nrounds" (boosting iterations), "max depth," "min child weight," "gamma," and "subsample." A grid search methodology was employed to find the best parameter combination, enhancing the model's performance. Important indices to measure variable importance in XGBoost include "Gain" and "Frequency" (Gertz et al., 2020; Li et al., 2023; Luo et al., 2024; Mehmood et al., 2024a; Ryu et al., 2020). SVR tuning involves selecting the appropriate kernel function and optimizing the "C" parameter, which balances margin width and misclassification tolerance. The versatility of SVR allows it to handle complex decision boundaries effectively (Yudong Li et al., 2020; Rodriguez-Galiano et al., 2015).

2.9. Performance evaluation of the models

The performance of the MLR, XGBoost, SVR, and RF algorithms was assessed for each variable group. These evaluations were based on measures such as the correlation coefficient *r*, RMAE, RRMSE, and mean error (ME) to produce a thorough comparison as per Eqs. (2–7) (H. Li et al., 2023; X. Li et al., 2020; Luo et al., 2024). The predictive mapping of the AGB distribution across each variable group was assigned to the method exhibiting the highest degree of accuracy. Sixteen prediction maps were produced as a result of this study using the L1, L2, L3, and L4 combinations.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}} \tag{2}$$

$$\mathrm{RRMSE} = \left(\frac{RMSE}{\bar{y}}\right) \times 100 \tag{3}$$

$$MAE = \sum_{1}^{n} \frac{|y_i - \hat{y}_i|}{n} \tag{4}$$

$$\mathrm{MAE} = \left(\frac{MAE}{\bar{y}}\right) \times 100 \tag{5}$$

$$ME = \sum_{i}^{n} \frac{(y_i - \hat{y}_i)}{n} \tag{6}$$

Table 2
To create maps of AGB, we gathered remote sensing indices from the Sentinel satellite series and utilized SRTM DEM data.

Source	Characteristics	Predictors	Descriptions
Landsat 9 – OLI 30-m	Multispectral Surface Reflectance	Band 2	Visible Blue (0.450–0.51 μm)
		Band 3	Visible Green (0.53–0.59 μm)
		Band 4	Red (0.64–0.67 μm)
		Band 5	Near-Infrared (0.85–0.88 μm)
		Band 6	SWIR 1(1.57–1.65 μm)
		Band 7	SWIR 2 (2.11–2.29 μm)
Landsat 9 – OLI	Vegetation Indices[1]	NDVI	Normalized difference vegetation index, (B5 – B4) / (B5 + B4)
		SAVI	Soil Adjusted Vegetation Index
		MSAVI	((Band 5 – Band 4) / (Band 5 + Band 4 + 0.5)) * (1.5).
		RVI/SR	Modified Soil Adjusted Vegetation Index
		GNDVI	(2 * Band 5 + 1 – sqrt ((2 * Band 5 + 1)^2–8 * (Band 5 – Band 4))) / 2
		TSAVI	Ratio Vegetation Index/ Simple Ratio
		TNDVI	(Band 5/ Band 4)
			Green Normalized Difference vegetation index, (B5 – B3) / (B5 + B3)
			Transformed soil adjusted vegetation index
			0.5 x (B5–0.5 x B4–0.5) / (0.5 x B5 + B4–0.15)
			Transformed normalized difference vegetation index, [(B5 – B4) / (B5 + B4) + 0.5] 1/2
SRTM DEM 30 m resolution	Elevation Terrain Based SRTM Indices[2]	E	Elevation
		S	Slope
		A	Aspect
		TRI	Surface roughness
		TPI	(2-pixel radius = a central point's 8 immediate neighbors)
		Twi	Topographic wetness index, Ln [Ac[3]/tan],

[1] (Ji et al., 2014; Silleos et al., 2006).
[2] (Bastin et al., 2022; Ge et al., 2022a).
[3] Ac represents the catchment area contributing to vertical water flow.

Table 3
Shows the formation of different predictor groups based on Landsat 9 and SRTM DEM.

Groups	Predictors Combination
L1	B2, B3, B4, B5, B6 and B7
L2	B1, B2, B3, B4, B5, B6 and B7 plus NDVI, SAVI, MSAVI, RVI, GNDVI, TSAVI and TNDVI
L3	B1, B2, B3, B4, B5, B6 and B7 plus elevation, slope, aspect, TWI, TPI and TRI
L4	B4, B5, elevation, aspect, TWI, NDVI, RVI, GNDVI, and TNDVI

$$r = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(\widehat{y}_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})} \times \sqrt{\sum_{i=1}^{n}(\widehat{y}_i - \overline{y})}} \quad (7)$$

In Eqs. (2) to (7), y_i denotes the observed AGB value, and n in this case = 371. $\widehat{y}_i$ represents the estimated AGB value from each model, and $\overline{y}$ is the mean of the observed AGB values. The goal is to minimize RMSE, RMAE, and ME while maximizing r to achieve a more accurate prediction.

Table 4
The study showcases the utilization of both linear regression and machine learning algorithms, executed using the R software.

Models	Description	Parameters	References
MLR	Linear regression method	Intercept (b_0), Coefficient Estimates, R-squared, adjusted R-squared, and significance levels of coefficients.	(López-Serrano et al., 2019; Luo et al., 2021a)
SVR	Support vector machine for regression	Kernel type, C regularization parameter)	(Bulut, 2023; Ge et al., 2022b; Sun et al., 2019)
XGBoost	Gradient boosting algorithm	Learning rate, Number of trees (n), Max depth	(Luo et al., 2021b; Tamiminia et al., 2022)
RF	Ensemble learning method	Number of trees (n), Max features, Max depth	(Chen et al., 2019a; Nguyen and Kappas, 2020)

3. Results

3.1. Relationship between Landsat 9 OLI and SRTM DEM based predictors and field-measured biomass

We discovered insights by examining the link between AGB observed in the field and several characteristics derived from Landsat 9 (L1) and SRTM DEM data. A notable association was identified between the L1 variables and the forest AGB (Fig. 5). The L1 subset consisted of a total of seven bands representing surface reflectance. Moreover, a positive correlation was observed among B4, B5, B6, and B7 with the measured AGB. Conversely, a negative interaction is evident between B2 and B3 to AGB. Regarding vegetation indices, B2–B3 and MSAVI reflectance showed negative connections with AGB, but the other 10 variables showed favorable associations. It's noteworthy that the reflectance values of B4, B5, NDVI, RVI, TNDVI, and GNDVI exhibit robust associations with AGB. The utilization of synthesized vegetation indices derived from the distinctive bands of Landsat 9, which are specifically tailored for monitoring vegetation chlorophyll, has proven to be significant and relevant in predicting forest AGB. Within the domain of topographical factors at the L3 level, our investigation revealed six specific variables: elevation (E), slope (S), aspect (A), topographic wetness index (TWI), topographic roughness index (TRI), and topographic position index (SPI). These variables exhibited noteworthy connections with forest AGB. The elevation, slope, and topographic wetness index (TWI) positively impacted the increase in AGB. Nevertheless, it is essential to acknowledge that a negative association between TRI, TPI, and aspect is observed (Fig. 5). An assessment was conducted among three combinations, namely L1, L2, and L3, to determine the most influential characteristics. These variables were subsequently incorporated into the L4 model, which is a hybrid model. Conclusively, the realistic depiction of forest AGB was achieved through the combined utilization of 19 variables obtained from Landsat 9 (L1), (L2), and SRTM DEM (L3).

The combinations of variables in Table 3 were carefully selected based on a systematic assessment of their relevance to AGB estimation. The L1 group includes the core spectral bands from Landsat-9 (B2, B3, B4, B5, B6, and B7). These bands capture the essential reflectance characteristics of vegetation, which are fundamental for estimating biomass. Bands B4 (Red) and B5 (Near-Infrared) are particularly significant as they are sensitive to chlorophyll content and vegetation structure. The L2 combination includes the spectral bands along with several vegetation indices (NDVI, SAVI, MSAVI, RVI, GNDVI, TSAVI, and TNDVI). Vegetation indices enhance the spectral information by emphasizing specific characteristics related to vegetation health, density, and stress. For instance, NDVI is widely used to assess plant health and biomass. The L3 group includes the spectral bands and topographic variables derived from the SRTM DEM (elevation, slope, aspect, TWI,

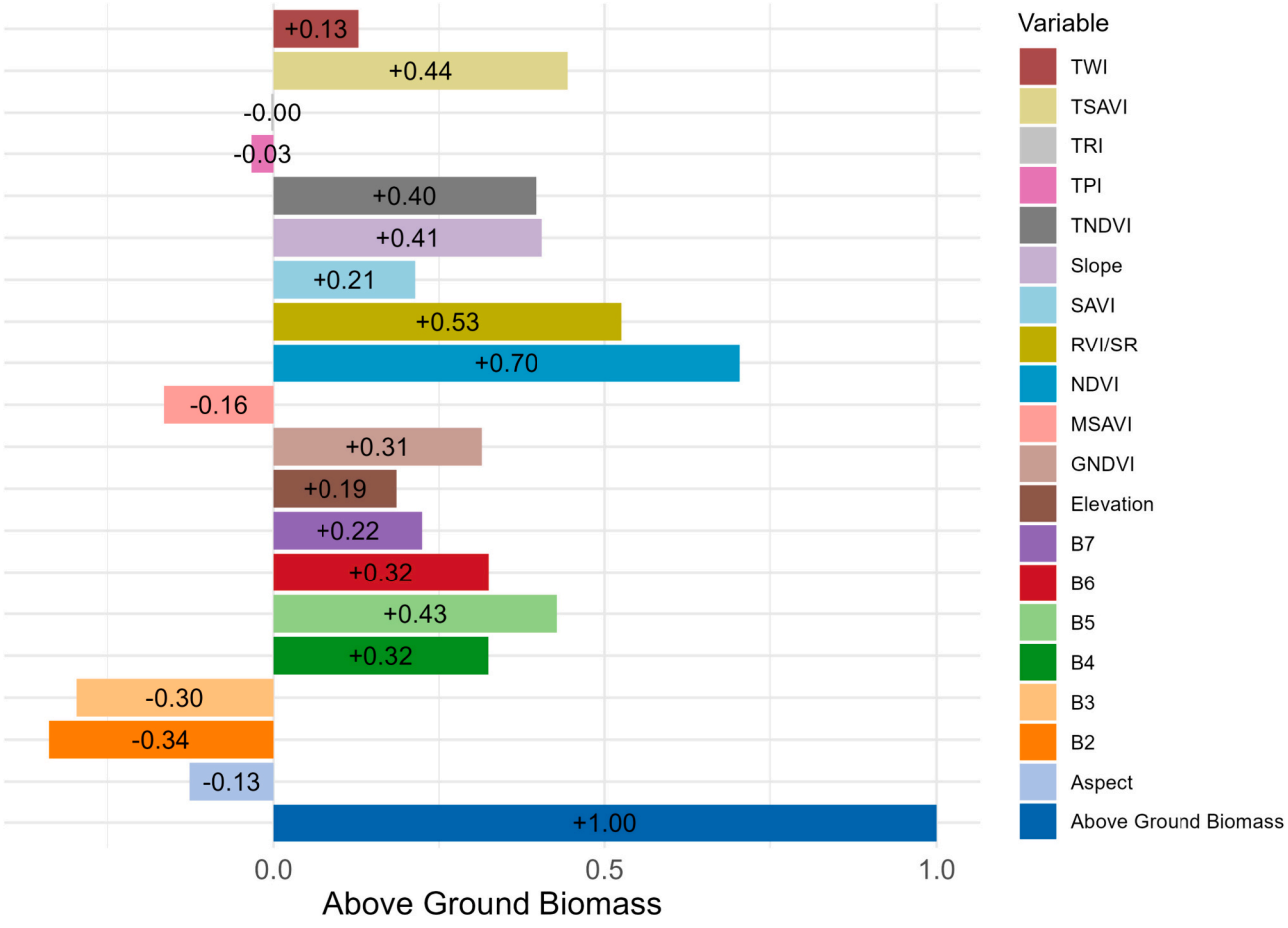

Fig. 5. Shows the relationship between field-measured biomass and various predictors.

TPI, and TRI). Topographic variables are crucial as they influence vegetation growth and biomass distribution by affecting factors like water availability, soil properties, and microclimate. The L4 combination integrates the most predictive variables from the previous groups (B4, B5, elevation, aspect, TWI, NDVI, RVI, GNDVI, and TNDVI). By selecting the top-performing variables, we aim to enhance the model's predictive capacity. This hybrid approach leverages the strengths of spectral, vegetation, and topographic data. This selection was driven by predictive power, computational efficiency, and accessibility. We prioritized variables like B4 and B5, known for their strong correlations with AGB, based on previous studies and our analyses. To avoid overfitting and manage complexity, we selected a manageable number of highly informative variables. The chosen variables are readily available from Landsat-9 and SRTM DEM datasets, ensuring our methodology can

Table 5
Model summary and accuracy assessment of four models on L1, L2, L3, and L4 predictor combination.

Groups	Models	R-squared	RMAE (%)	RRMSE (%)	MAE	RMSE	Parameters
L1	MLR	0.25	36.23	43.24	29.67	35.42	
	SVR	0.28	32.21	42.70	26.38	34.97	C[1]=1, Sigma[2] = 10
	XGBoost	0.58	23.81	32.46	19.50	26.59	max_depth[3] = 7, eta[4] = 0.3
	RF	0.66	20.94	29.77	17.15	24.38	max_depth = 25, n_estimators[5] = 1000
L2	MLR	0.21	36.95	44.51	30.26	36.45	
	SVR	0.60	20.37	31.46	16.68	25.75	C = 10, Sigma = 10
	XGBoost	0.66	20.05	29.20	16.42	23.90	max_depth = 7, eta = 0.3
	RF	0.70	18.90	27.82	15.48	22.79	max_depth = 25, n_estimators = 1000
L3	MLR	0.19	37.82	44.87	30.97	36.75	
	SVR	0.74	16.17	25.24	13.24	20.67	C = 1, Sigma = 10
	XGBoost	0.77	15.54	23.85	12.73	19.54	max_depth = 7, eta = 0.3
	RF	0.81	14.33	22.23	11.73	18.21	max_depth = 25, n_estimators = 1000
L4	MLR	0.23	36.70	43.89	30.06	35.95	
	SVR	0.62	20.02	30.78	16.40	25.21	C = 1, Sigma = 10
	XGBoost	0.72	18.16	26.53	14.87	21.72	max_depth = 7, eta = 0.3
	RF	0.75	17.56	25.61	14.38	20.97	max_depth = 25, n_estimators = 1000

[1] "C" is the regularization parameter.
[2] Bandwidths.
[3] Maximum depth of a tree.
[4] Learning rate.
[5] Number of Trees.

be easily replicated in other studies or regions.

3.2. Evaluation of different predictor combinations and machine learning algorithms

3.2.1. Variable combinations and model comparison

In this section, we will show the findings of our study about the combination of variables L1, L2, L3, and L4. The results of different comparison parameters against all the evaluated models are presented in Table 5.

All four models showed strong performance when utilizing various combinations of predictors, except the MLR model, which exhibited satisfactory performance in predicting the AGB. The MLR model produced R-squared values of 0.25, 0.21, and 0.19 in L1, L2, and L3, respectively. It is worth noting that the maximum error was seen in all combinations. The model demonstrated an RMAE value of 36.23% and an RRMSE value of 43.24%, indicating a modest degree of prediction precision. This weaker performance can be attributed to the sensitivity of MLR to its underlying assumptions, such as linearity, homoscedasticity, independence, and normality of residuals. MLR is less robust in

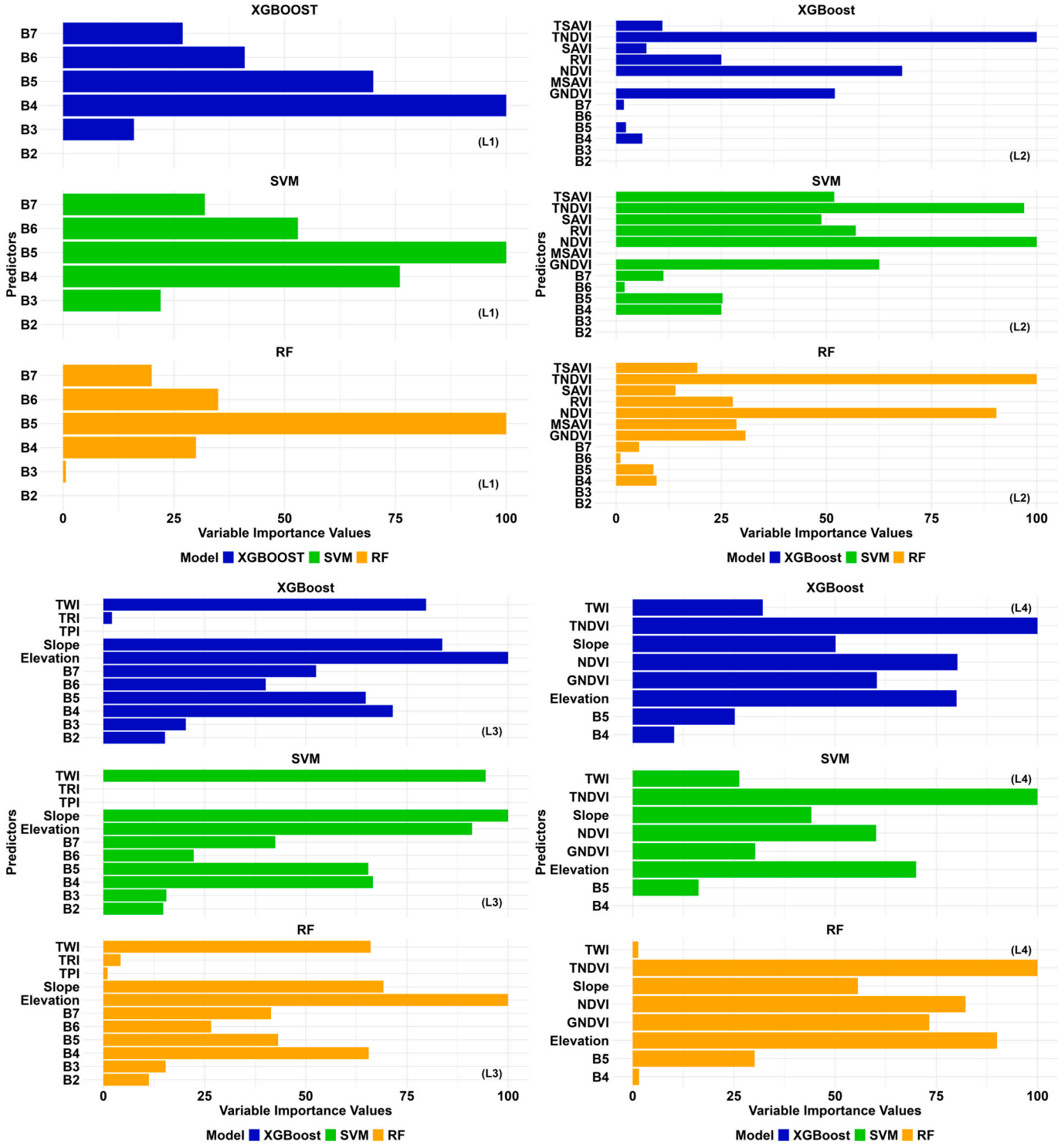

Fig. 6. Showing the variable's importance derived from MLR, SVM, XGBoost, and RF algorithms across L1, L2, L3, and L4 predictor combinations.

handling violations of these assumptions, which are often present in complex environmental data sets, thereby impacting its predictive accuracy. The SVR model exhibited R-squared values of 0.27, 0.60, and 0.74 in L1, L2, and L3, respectively. These results indicate that the SVR model outperformed the MLR model. However, it fell below both the XGBoost and RF models in all combinations. The SVR model demonstrated enhanced predictive accuracy compared to the MLR model, as evidenced by its RMAE of 32.21% and RRMSE of 42.70%. The hyperparameters C = 1 and Sigma = 10 was optimized to attain these outcomes. In contrast, XGBoost demonstrated significant enhancements in performance compared to the preceding models, with R-squared values of 0.58, 0.66, and 0.77. The obtained RMAE of 23.81% and RRMSE of 32.46% demonstrated high prediction accuracy. The higher performance can be attributed to the utilization of optimal hyperparameters, specifically a max_depth value of 7 and an eta value of 0.3. Out of the models that were assessed, it was found that the RF model exhibited the most superior predictive capability. The model demonstrated a strong fit to the data, as evidenced by R-squared values of 0.66, 0.70, and 0.81. The RF model showed exceptional error reduction performance, as seen by an RMAE of 14.33% and RRMSE of 22.23%. The hyperparameters, namely max_depth = 25, n_estimators = 1000, and mtry = 1,3,7, were optimized to attain the observed level of performance in each of the three combinations.

3.2.2. Evaluation and selection of important predictors for final model L4

A comprehensive analysis was performed to determine the significance of each predictor variable in the L1, L2, and L3 combination to the model's predictive performance. Fig. 6 summarizes the variable importance scores, which have been standardized within each model.

The variable importance scores indicate each predictor's relative impact on the models' performance. It is noted that B5 and B4 displayed considerable significance in all models in L1, indicating their consistent impact on predictive accuracy. The B5 (NIR) band is associated with its responsiveness to the photosynthetic activity and structural characteristics of vegetation (Mehmood et al., 2024a). The absorption of visible light by vegetation, namely in the red band (B4), is significant due to the presence of pigments like chlorophyll. The effectiveness of B5 (NIR) and B4 (Red) bands of Landsat imagery in predicting AGB is attributed to their sensitivity to multiple factors, including vegetation health, photosynthetic activity, canopy structure, and the biochemical as well as structural properties of vegetation (Y. Chen et al., 2023c). This sensitivity is additionally utilized through indices such as the Normalized Difference Vegetation Index (NDVI), which offers measurable indicators of vegetation vigor and biomass density (Mehmood et al., 2024b; Mehmood et al., 2024c; Shahzad et al., 2024). The indices TNDVI, NDVI, GNDVI, and RVI have shown significant importance in the context of L2. The L3 model demonstrated the most favorable combination of predictors, with a better prediction accuracy of 0.81 (R^2) and minimal error. The elevation, slope, and topographic wetness index (TWI) combination exhibited strong predictive performance in all models. Furthermore, the thermal water index (TWI) indicates regions with higher water saturation levels resulting from water accumulation. These regions have the potential to sustain distinct plant communities and increase AGB as a result of favorable water availability (Badshah et al., 2024; Bastin et al., 2022).

The L4 predictor combination was developed based on the previously described findings. The variables B4 and B5 from L1, TNDVI, NDVI, GNDVI, and RVI from L2, and elevation, slope, and TWI from L3 were chosen as components of L4, referred to as the hybrid combination of predictors. Table 5 presents the evaluation of the L4 predictor combination, which is a hybrid model consisting of variables from the remaining three models. The MLR model exhibits an R-squared value of 0.23, while the SVR model demonstrates an R-squared value of 0.62. Moreover, the XGBoost model demonstrates an R-squared value of 0.72, while the RF model outperforms the highest R-squared value of 0.75. Specifically, in L3, the RF model showcases the highest accuracy, evident in the lowest MAE and RMSE values. The RF model demonstrated superior prediction accuracy compared to the other models across all combinations of the L1, L2, L3, and L4 variables, as seen by the findings in Table 5. The enhanced performance of the models can be attributed to the combination of input variables, which possess a wide range of features, as well as the utilization of optimized hyperparameters. Moreover, it has been shown that the most reliable indicators for estimating AGB are B4 and B5 from L1, TNDVI, NDVI, GNDVI, and RVI from L2. Additionally, it has been established that elevation, slope, and TWI from L3 significantly contribute to the prediction process, as depicted in Fig. 6.

The L3 predictor combination had the highest predictive accuracy with an R^2 value of 0.81. Additionally, the RMAE and RRMSE values for this combination were determined to be the lowest compared to all other combinations. The RF model demonstrated an RMAE of 14.33% and an RRMSE of 22.23%, highlighting its superior predictive accuracy. The analysis of variable importance confirms the impact of predictors on the accuracy of each model, hence underscoring the significance of feature selection in improving predictive ability. The results mentioned above highlight the importance of carefully choosing the input variables and the suitable machine learning algorithm to attain precise predictions.

3.3. Performance evaluation and AGB mapping

To evaluate the efficacy of these models, we utilized scatterplots, which effectively illustrate the relationship between observed AGB values and the corresponding predictions (Fig. 7). Our findings show that both RF and XGBoost models consistently outperform the MLR and SVR models when applied to the same dataset. The RF model demonstrates a slightly superior performance compared to the XGBoost model. This advantageous performance is accompanied by higher R-squared values for the RF and XGBoost models in contrast to the SVR and MLR models. However, interestingly, the hybrid dataset L4, combining the information from all three predictor combinations, L1, L2, and L3, did not exhibit significantly improved model performance compared to L3.

3.4. Spatial distribution characteristics of AGB

The predicted forest AGB values within the geographical scope of the investigation exhibited a range spanning from 8.91 to 171.86 Mg/ha, as illustrated in Fig. 2. These values were categorized into six discrete tiers, demarcated by intervals derived from the measured AGB values, as depicted in Fig. 8, to enhance the comparative analysis among the four models. All cartographic representations consistently depicted the northern regions in the study areas characterized by elevated AGB levels, encompassing values ranging from 97.2 to 170.28 Mg/ha. Contrarily, zones denoting low AGB levels (ranging from 8.29 to 51.46 Mg/ha) were proximate to non-forested vicinities. The RF model exhibited superior performance upon analyzing the predicted versus observed AGB values (Fig. 8). Subsequently, the cartographic representation stemming from the RF model, based solely on Landsat 9 spectral bands L1 and L2 predictor combination, revealed a fragmented AGB distribution pattern characterized by spatial randomness (Figs. 8-A (L1) and 8-A (L2)).

The best distribution of predicted AGB was observed in the L3 and hybrid combination L4 (Figs. 7-A(L3) and 7-A (L4)), which showcased enhanced predictive capability in contrast to L1 and L2 variables in the context of forest AGB mapping through the employment of the RF algorithm. The integration of topographic variables from the SRTM DEM in the L3 and L4 combinations significantly improved the model's accuracy, highlighting the importance of incorporating elevation, slope, and TWI in the prediction of AGB. Overall, the RF model showed relatively better prediction performance within the desired range of measured AGB (8.23–171.82 Mg/ha) across all the predictor combinations. Nonetheless, the RF model avoided predicting overestimated forest AGB values (ranging from 175.25 to 198.28 Mg/ha) that were

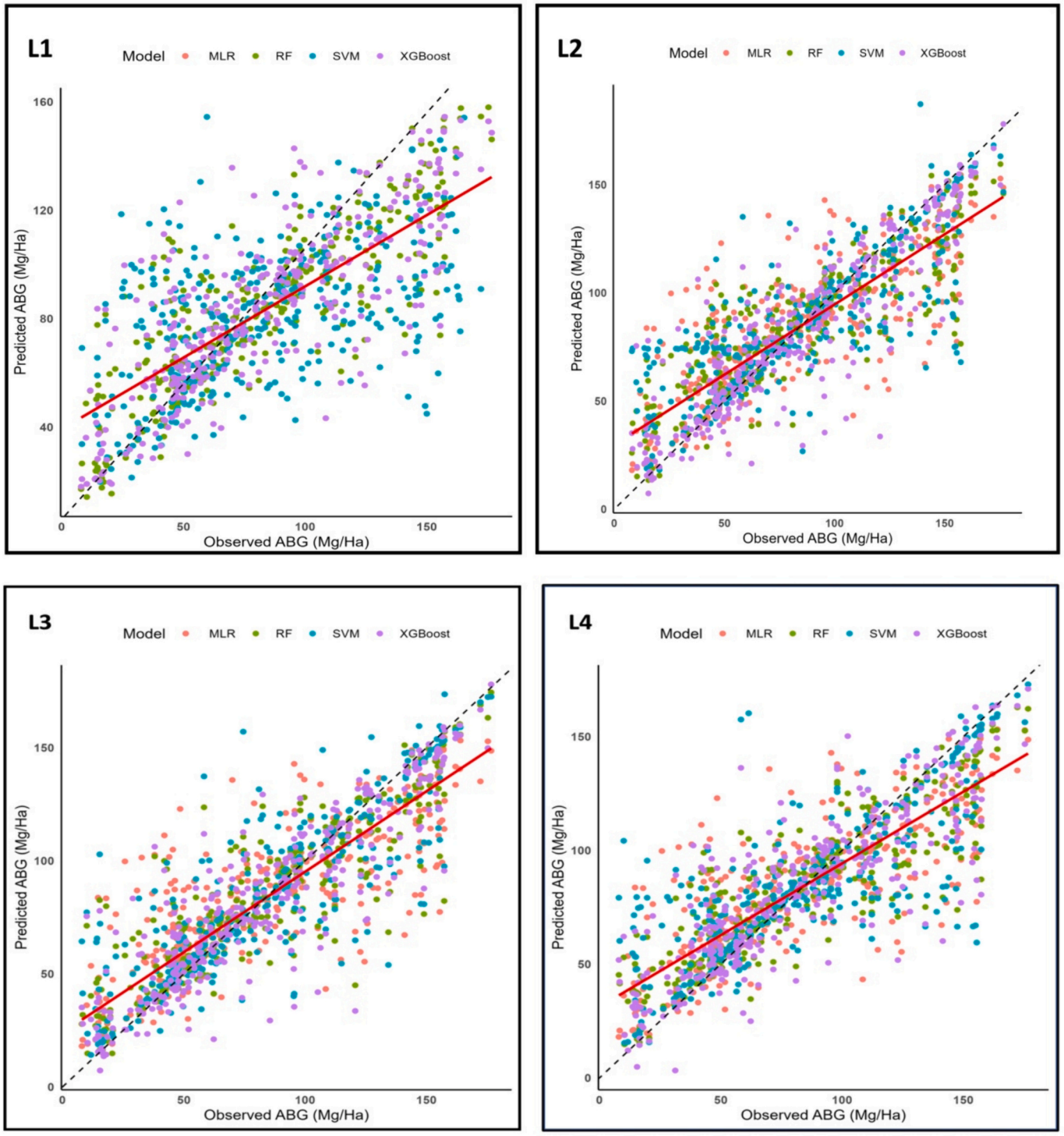

Fig. 7. Shows Predicted outcomes by the MLR, SVR, RF, and XGBoost models across L1, L2, L3 and L4 datasets.

present in the other models, such as MLR and SVR, which either over- or under-estimated AGB (Figs. 8-C and 8-D). The predicted AGB distribution generated through the RF model with the combination of L3 predictors closely paralleled the outcomes of the L2 and L4 RF models. In contrast, the RF model using L1 alone slightly underestimated AGB values, with predictions ranging from 12.9 to 166.2 Mg/ha. Incorporating topographic indices obtained from SRTM DEM alongside the spectral bands from Landsat 9 in the RF model (Fig. 8-A (L3)) notably enhanced the accuracy of forest AGB mapping, providing a more detailed and reliable spatial distribution of AGB.

4. Discussion

The techniques and methodologies used in this study are well-established within the field of remote sensing and machine learning. Our approach effectively combined Landsat-9 data and machine learning algorithms, focusing on spectral bands, vegetation indices, and topographical features. This combination allowed us to capture essential characteristics of forest biomass with high accuracy. The RF model demonstrated superior prediction performance, achieving an R^2 value of 0.81, which is considered excellent in AGB estimation studies.

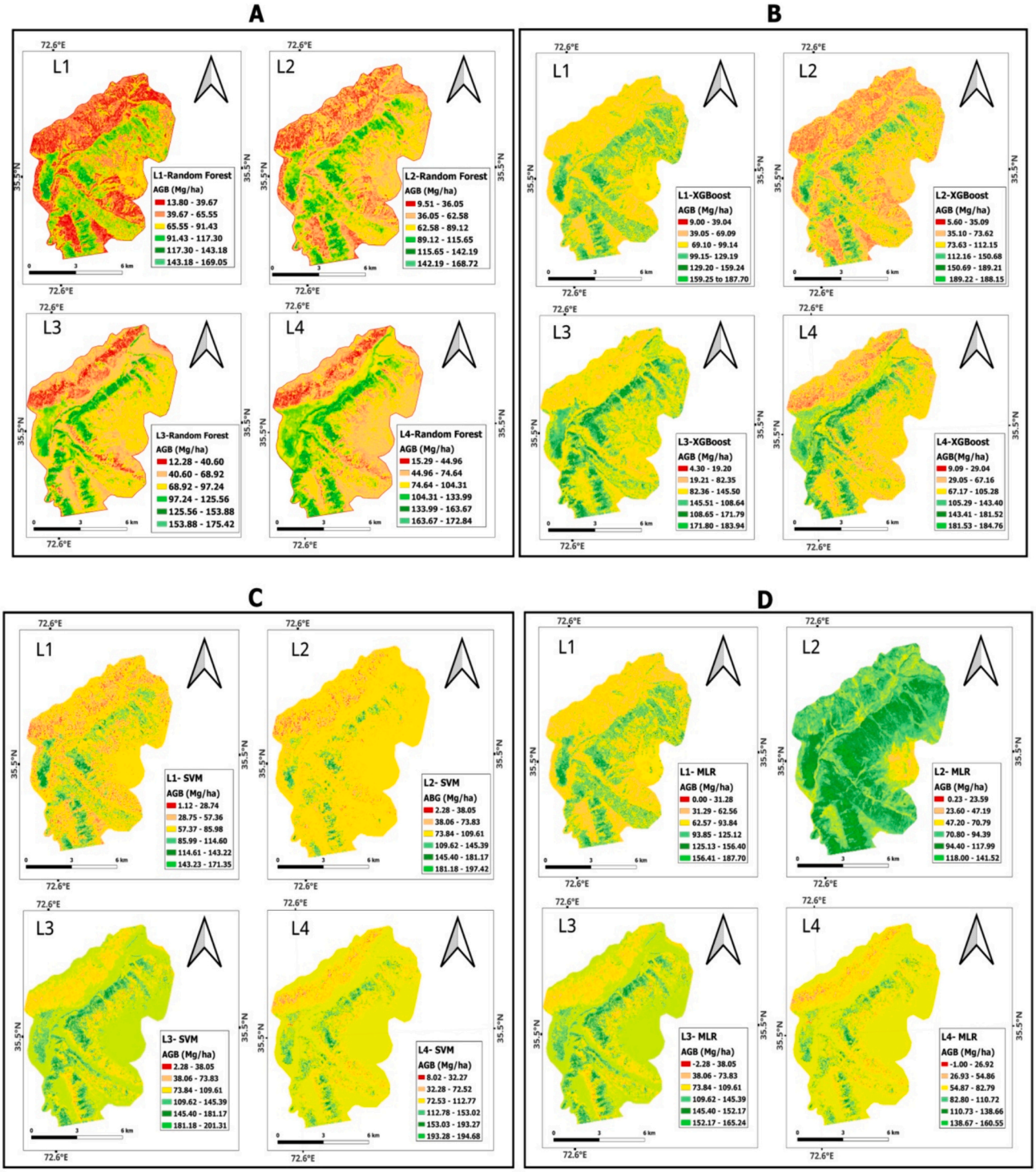

Fig. 8. The study generated maps predicting AGB by employing the optimal combination of predictors and models, utilizing data from the Landsat 9 OLI series and the SRTM DEM: (A) L1, L2, L3 and L4 across RF; (B) L1, L2, L3 and L4 across XGBoost; (C) L1, L2, L3 and L4 across SVM; (D) L1, L2, L3 and L4 across MLR.

Tamiminia et al. (2024) found that integrating multiple remote sensing data sources, such as LiDAR, optical, and SAR, with tree-based machine learning models, achieved an R^2 value of 0.81 for AGB estimation, similar to our results. However, the use of object-based image analysis (OBIA) provided better accuracy compared to pixel-based approaches. Lu et al. (2023) reported that gradient boosting decision tree (GBDT) models combined with spectral indices achieved an R^2 of 0.99 in mixed-species forest types, indicating slightly higher accuracy when using multi-source data and advanced machine learning techniques. Predictor variables were carefully selected and optimized, which included specific Landsat-9 bands known for their sensitivity to vegetation properties (such as B4 and B5), vegetation indices (like NDVI, SAVI, and TNDVI) for their effectiveness in vegetation health assessment and biomass estimation, and topographical features (such as elevation, slope, and the Topographic Wetness Index) derived from the SRTM DEM. By combining these carefully selected variables, the model's predictive power was enhanced without the need for an extensive array of remote sensing variables. This approach demonstrates that a well-thought-out

selection of relevant predictors can yield highly accurate results, even with a limited number of variables.

This study explained the multifaceted associations between the measured AGB and various predictors derived from Landsat 9 OLI -based data and topographical parameters. These relationships were determined by applying algorithms tailored to the following parameters derived from the MLR, along with attribute importance evaluated from machine learning as XGBoost and RF models. Among the variables derived from Landsat 9 OLI (L1) data, B4 and B5 emerged as the most pivotal factors in explicating forest AGB's spatial distribution patterns. This significance was prominently evident in both XGBoost and SVR and was consistently validated by the RF models in all variable combinations. In the pursuit of forest AGB mapping, vegetation-based indices (L2) were characteristically computed. Notably, its utilization heightened potential in forest AGB mapping, as substantiated by correlation analyses (Fig. 5). RF analysis systematically assessed the significance of various vegetation indices, namely TNDVI, NDVI, RVI, and GNDVI, as crucial predictors across a spectrum of variable combinations. This prominence was consistently observed within the RF model and corroborated by alternative machine learning techniques such as SVR and XGBoost algorithms (Chen et al., 2020). Concurrently, the vegetation indices derived from L1 data and their subsequent computation proved to be valuable and shared predictors in the modeling, as validated by prior research (Guo et al., 2023; Lemenkova, 2020; Mehmood et al., 2024a; Yang et al., 2022). Vegetation indices like TNDVI, NDVI, and GNDVI indicate these factors, as they are based on the differential reflectance of specific wavelength bands that respond to chlorophyll content, leaf area, and photosynthetic activity. Models incorporating these indices can indirectly capture vegetation health and density variations, which strongly influence AGB (Khunrattanasiri, 2023; Luo et al., 2024). Vegetation indices respond to changes in canopy structure (Anees et al., 2022b; Hussain et al., 2024b,a; Jackson and Huete, 1991; Mehmood et al., 2024c). For instance, the GNDVI is sensitive to the early stages of vegetation growth and changes in green biomass. Models incorporating GNDVI can capture changes in canopy density and growth dynamics that are relevant for predicting AGB (Tamiminia et al., 2024). In AGB prediction, accurate estimation of vegetation cover and characteristics is essential. By reducing the impact of non-vegetation factors, TNDVI can help produce more accurate forecasts of AGB by isolating the true vegetation signal (Fujiki, 2017). These indices enhance the sensitivity of remote sensing data to vegetation-related factors that influence AGB. NDVI, for instance, emphasizes the difference in reflectance between NIR and red bands, which are influenced by chlorophyll absorption and leaf structure (Hussain et al., 2024b,a). By capturing these variations, NDVI provides insight into the overall health and productivity of vegetation which are closely linked to AGB (Anees et al., 2022a; Mehmood et al., 2024c; Mngadi et al., 2022). Moreover, it emphasizes the calibration and validation of AGB prediction models. Field measurements of AGB can be correlated with remote sensing-derived indices to establish relationships and calibrate the models (Wu et al., 2016). Moreover, the accuracy of AGB predictions can be validated using these indices against ground-based measurements, enhancing the reliability of the modeling approach (Bui et al., 2024).

Elevation, slope, and topographic wetness index (TWI) are proxy indicators in AGB prediction models representing various biophysical and biological mechanisms impacting vegetation growth and spatial distribution (Hojo et al., 2023). Researchers can account for the intricate relationships between topography, climate, and vegetation (Mehmood et al., 2024c) by incorporating these topographic variables into the models. This leads to improved accuracy and ecologically significant projections of AGB in diverse environments (Salinas-Melgoza et al., 2018). The elevation variable demonstrated the highest level of correlation, as supported by the empirical evidence presented in Fig. 5. This conclusion is consistent with previous research that highlights the significant impact of topography on variables such as water distribution and solar irradiation, as documented in previous studies (Fararoda et al., 2021; Liu et al., 2021). The Topographic Wetness Index (TWI), considered an important indication of soil-water retention which is essential for vegetation, has been previously recognized as a key factor in estimating AGB in forest ecosystems using nonlinear regression models (Martinuzzi et al., 2022). Nevertheless, the present investigation showed that TWI exhibited a higher degree of relevance to the measured AGB compared to the topographic position index and topographic roughness index, which displayed a greater relevance to the measured AGB. The RF model revealed that elevation, slope, and topographic wetness index (TWI) were the most influential predictors, with B3 and B5 following closely behind. The incorporation of the same has been notably indicated by its presence in the XGBoost and SVR models.

We thoroughly studied the process of selecting predictors to map AGB in forested areas accurately. This study successfully identified the primary predictors for the mapping in question. The predictors included in this study were derived from four distinct variable groups, namely L1, L2, L3, and L4. The RF method demonstrated exceptional proficiency as a modeling technique for all variable groups, indicating its strong potential for predicting forest AGB. The preference for the RF algorithm aligns with previous research findings reported in the scholarly literature (Chen et al., 2019b; Luo et al., 2021d; Nandy et al., 2021; Purohit et al., 2021b). Conversely, the XGBoost algorithm demonstrated the second-highest potential for modeling across all variable groups, showcasing a considerable capacity for accurate prediction of AGB. SVR had suboptimal performance concerning L1 variables, whereas the MLR algorithm demonstrated inferior performance concerning all modeling scenarios. The study showed the suitability of the SVR algorithm for making predictions with restricted datasets, a finding that aligns with previous research investigations (Kaveh et al., 2023; Mehmood et al., 2024a; Zhang et al., 2020; Zhang et al., 2022). An important conclusion is derived from Fig. 6. It highlights that MLR heavily depends on predictor variables, demonstrating a more pronounced reliance on these elements than machine learning methods. Furthermore, it is worth noting that relying solely on spectral bands to predict AGB is unlikely to produce favorable outcomes. Integrating these spectral bands with other variable sets, including those produced from the Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM) and vegetation-derived indices, will lead to more favorable outcomes.

Within the scope of this study, the optimal combination that demonstrated the highest predictive power for AGB estimation was denoted as "L4." This configuration distinctly outperformed alternative predictor combinations, substantiating its superior efficacy. Notwithstanding the varied predictor combinations explored and acknowledging the limitations of the saturation phenomenon witnessed in the Shuttle SRTM DEM and optical multispectral data regarding their ability to discern the sensitivity of forest AGB, it is noteworthy that the Landsat satellite series, along with the SRTM DEM, delivered essential insights for forest AGB estimation. This contrasted with the Landsat series used with its derived vegetation indices. This choice of data sources, characterized by their fine spatial resolution and comprehensive geographical coverage, emerged as indispensable resources for accurate forest AGB estimation.

This study has made notable advancements in estimating AGB in temperate Himalayan forests. These advancements were made possible by combining Landsat-9 remote sensing data with robust machine learning methods, including Random Forest, XGBoost, and Support Vector Regression. This methodology has not only improved the precision of AGB predictions, which are crucial for the effective management of forests and carbon accounting but has also underscored the need to integrate distinct spectral bands, vegetation indices, and topographical characteristics into ecological models. One significant contribution of this study is the systematic assessment of different combinations of predictors and algorithms, which sheds light on the complex capacities of machine learning models in effectively analyzing intricate and nonlinear data from forest ecosystems. The results of our study establish a standard in the realm of forest biomass prediction, making a

substantial contribution to the disciplines of ecological monitoring and conservation (Khan et al., 2024). Nevertheless, it is important to recognize several constraints in our study, including the potential for data saturation in dense forest canopies and the applicability of our models to other ecological contexts. Subsequent investigations must prioritize resolving these obstacles, delving into the enhancement of algorithms, incorporating nascent technologies such as LiDAR, and evaluating the enduring ramifications of climatic dynamics on AGB.

5. Conclusion

This study demonstrates significant advancements in AGB estimation in temperate Himalayan forests by leveraging Landsat-9 remote sensing data combined with robust machine learning algorithms. Through a systematic assessment of different predictor combinations and algorithms, we identified the optimal variables and methods for accurate AGB prediction. Among the evaluated models, the RF algorithm consistently outperformed others, demonstrating superior prediction accuracy. The RF model achieved an R^2 value of 0.81, which is notably high for AGB estimation studies. This performance can be attributed to the model's ability to handle complex interactions between variables and its robustness in dealing with high-dimensional data. The optimal predictor combination, denoted as "L4," integrated variables from Landsat-9 spectral bands (L1), vegetation indices (L2), and topographical features (L3) derived from the SRTM DEM. The hybrid combination showcased the highest predictive accuracy, with the lowest RMAE and RRMSE values. Specifically, the L3 predictor combination had an R^2 value of 0.81, making it the most effective for AGB estimation. Key variables that emerged as reliable indicators for AGB included B4 and B5 from L1, TNDVI, NDVI, GNDVI, and RVI from L2, and elevation, slope, and Topographic Wetness Index (TWI) from L3. These variables significantly contributed to the prediction process, as corroborated by their importance scores in the RF model. In summary, this study underscores the efficacy of the RF algorithm as the premier modeling approach for predicting AGB, with XGBoost as the subsequent choice. Although SVR showed potential for small datasets, the MLR algorithm generally exhibited subpar results. The integration of diverse data sources, including Landsat-9 and SRTM DEM, highlighted the importance of high-resolution spatial data in accurately modeling AGB. These findings contribute significantly to understanding the intricate relationships between AGB and various predictors, offering promising prospects for employing advanced machine learning techniques in ecological modeling. The results are relevant for implementing efficient forest management strategies, accurately evaluating carbon sequestration, and monitoring ecological conditions in the temperate Himalayan region of Pakistan and other similar ecosystems worldwide.

Funding

Ecosystem Soil and Water Conservation Function Assessment Project in Beibu Gulf, Guangxi Province (Grant No.84-Y50-G29–9001-22/23), and the MNR-CN Key Laboratory of China-ASEAN Satellite Remote Sensing Applications (Grant No. ZDMY202310), Science and Technology Base and Talent Project of Guangxi (grant no. Guike-AD23026073), Guangxi Young and Middle-aged University Teachers' Scientific Research Ability Enhancement Project (Grant No.2023KY0399).

CRediT authorship contribution statement

Shoaib Ahmad Anees: Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Data curation, Validation, Supervision, Writing – original draft, Writing – review & editing. **Kaleem Mehmood:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Data curation, Validation, Writing – original draft, Writing – review & editing. **Waseem Razzaq Khan:** Formal analysis, Investigation, Writing – review & editing. **Muhammad Sajjad:** Writing – review & editing. **Tahani Awad Alahmadi:** Writing – review & editing. **Sulaiman Ali Alharbi:** Writing – review & editing. **Mi Luo:** Formal analysis, Writing – review & editing.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

The authors confirm that the data links supporting the findings of this study are available within the article.

Acknowledgements

We are grateful to the Department of Forestry, The University of Agriculture, Dera Ismail Khan, 29050, Pakistan, for providing assistance and platforms for this research. We are also grateful to the Key Laboratory for Silviculture and Conservation of Ministry of Education, Beijing Forestry University, Beijing, (100083), P. R. China, for providing assistance and platforms for this research. This project was supported by Researchers Supporting Project number (RSP2025R230) King Saud University, Riyadh, Saudi Arabia. Authors also 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.ecoinf.2024.102732.

References

Abbas, S., Shirazi, S.A., Hussain, M.S., Yaseen, M., Shakarullah, K., Wahla, S.S., Khurshid, M., 2020a. Impact of climate change on forest cover: implications for carbon stock assessment and sustainable development in HKH region-Pakistan. Pakistan Vision 21, 66.

Abbas, S., Shirazi, S.A., Hussain, M.S., Yaseen, M., Shakarullah, K., Wahla, S.S., Khurshid, M., 2020b. Impact of climate change on forest cover: implications for carbon stock assessment and sustainable development in HKH region-Pakistan. Pakistan Vision 21, 66.

Akram, M., Hayat, U., Shi, J., Anees, S.A., 2022. Association of the female flight ability of Asian spongy moths (*Lymantria dispar* asiatica) with locality, age and mating: a case study from China. Forests 13 (8), 1158.

Andreevich, U.V., Reza, S.S.O., Stepanovich, T.I., Amirhossein, A., Meng, Z., Anees, S.A., Petrovich, C.V., 2020. Are there differences in the response of natural stand and plantation biomass to changes in temperature and precipitation? A case for two-needled pines in Eurasia. J. Resources Ecol. 11 (4), 331.

Anees, S.A., Zhang, X., Khan, K.A., Abbas, M., Ghramh, H.A., Ahmad, Z., 2022a. Estimation of fractional vegetation cover dynamics and its drivers based on multi-sensor data in Dera Ismail Khan, Pakistan. J. King Saud. Univ. Sci. 34 https://doi.org/10.1016/j.jksus.2022.102217.

Anees, S.A., Zhang, X., Shakeel, M., Al-Kahtani, M.A., Khan, K.A., Akram, M., Ghramh, H. A., 2022b. Estimation of fractional vegetation cover dynamics based on satellite remote sensing in Pakistan: a comprehensive study on the FVC and its drivers. J. King Saud Univ. Sci. 34 (3), 101848.

Anees, S.A., Yang, X., Mehmood, K., 2024. The stoichiometric characteristics and the relationship with hydraulic and morphological traits of the Faxon fir in the subalpine coniferous forest of Southwest China. Ecol. Indic. 159, 111636.

Anwar, S., Khan, S.M., Ahmad, Z., Ullah, Z., Iqbal, M., 2019. Floristic composition and ecological gradient analyses of the Liakot forests in the Kalam region of district swat, Pakistan. J. For. Res. (Harbin) 30, 1407–1416.

Aslam, M.S., Huanxue, P., Sohail, S., Majeed, M.T., Rahman, S.U., Anees, S.A., 2022. Assessment of major food crops production-based environmental efficiency in China, India, and Pakistan. Environ. Sci. Pollut. Res. 1–10.

Badshah, M.T., Hussain, K., Rehman, A.U., Mehmood, K., Muhammad, B., Wiarta, R., Silamon, R.F., Khan, M.A., Meng, J., 2024. The role of random forest and Markov chain models in understanding metropolitan urban growth trajectory. Front. For. Glob. Change 7. https://doi.org/10.3389/ffgc.2024.1345047.

Bahadur, S., Ahmad, M., Zafar, M., Begum, N., Ali, M., Kumar, T., 2023. Ethnomedicinal relevance of selected monocot taxa from different geographical regions of Pakistan. Ethnobot. Res. Appl. 26, 1–17.

Bastin, J.-F., Depoortere, P., Meersmans, J., Fayolle, A., Charles, C., 2022. Influence of Topography on Tropical Forest Structure and Composition in Democratic Republic of Congo.

Battude, M., Al Bitar, A., Morin, D., Cros, J., Huc, M., Sicre, C.M., Le Dantec, V., Demarez, V., 2016. Estimating maize biomass and yield over large areas using high

spatial and temporal resolution Sentinel-2 like remote sensing data. Remote Sens. Environ. 184, 668–681.
Baul, T.K., Chakraborty, A., Nandi, R., Mohiuddin, M., Kilpeläinen, A., Sultana, T., 2021. Effects of tree species diversity and stand structure on carbon stocks of homestead forests in Maheshkhali Island, southern Bangladesh. Carbon Balance Manag. 16, 1–15.
Behera, D., Kumar, V.A., Rao, J.P., Padal, S.B., Ayyappan, N., Reddy, C.S., 2023. Estimating aboveground biomass of a regional forest landscape by integrating textural and spectral variables of Sentinel-2 along with ancillary data. J. Indian Soc. Remote Sens. 1–13.
Bouras, E.H., Jarlan, L., Er-Raki, S., Balaghi, R., Amazirh, A., Richard, B., Khabba, S., 2021. Cereal yield forecasting with satellite drought-based indices, weather data and regional climate indices using machine learning in Morocco. Remote Sens. 13, 3101.
Bruenig, E.F., 2016. Conservation and management of tropical rainforests: an integrated approach to sustainability. Cabi.
Bui, Q.T., Pham, Q.T., Pham, V.M., Tran, V.T., Nguyen, D.H., Nguyen, Q.H., Nguyen, H. D., Do, N.T., Vu, V.M., 2024. Hybrid machine learning models for aboveground biomass estimations. Ecol. Inform. 79 https://doi.org/10.1016/j.ecoinf.2023.102421.
Bulut, S., 2023. Machine learning prediction of above-ground biomass in pure Calabrian pine (*Pinus brutia* Ten.) stands of the Mediterranean region, Türkiye. Ecol. Inform. 74, 101951.
Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., 2015. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. Methods Ecol. Evol. 6, 198–208.
Cameron, H.A., Panda, P., Barczyk, M., Beverly, J.L., 2022. Estimating boreal forest ground cover vegetation composition from nadir photographs using deep convolutional neural networks. Ecol. Inform. 69 https://doi.org/10.1016/j.ecoinf.2022.101658.
Chasmer, L., Mahoney, C., Millard, K., Nelson, K., Peters, D., Merchant, M., Hopkinson, C., Brisco, B., Niemann, O., Montgomery, J., 2020. Remote sensing of boreal wetlands 2: methods for evaluating boreal wetland ecosystem state and drivers of change. Remote Sens. 12, 1321.
Chen, Q., Laurin, G.V., Valentini, R., 2015. Uncertainty of remotely sensed aboveground biomass over an African tropical forest: propagating errors from trees to plots to pixels. Remote Sens. Environ. 160, 134–143.
Chen, L., Ren, C., Zhang, B., Wang, Z., Xi, Y., 2018. Estimation of forest above-ground biomass by geographically weighted regression and machine learning with sentinel imagery. Forests 9, 582.
Chen, L., Wang, Y., Ren, C., Zhang, B., Wang, Z., 2019a. Assessment of multi-wavelength SAR and multispectral instrument data for forest aboveground biomass mapping using random forest kriging. For. Ecol. Manag. 447, 12–25.
Chen, L., Wang, Y., Ren, C., Zhang, B., Wang, Z., 2019b. Assessment of multi-wavelength SAR and multispectral instrument data for forest aboveground biomass mapping using random forest kriging. For. Ecol. Manag. 447, 12–25.
Chen, R.-C., Dewi, C., Huang, S.-W., Caraka, R.E., 2020. Selecting critical features for data classification based on machine learning methods. J. Big Data 7, 52.
Chen, S., Chen, J., Jiang, Chunqian, Yao, R.T., Xue, J., Bai, Y., Wang, H., Jiang, Chunwu, Wang, S., Zhong, Y., 2022. Trends in research on forest ecosystem services in the most recent 20 years: a bibliometric analysis. Forests 13, 1087.
Chen, Y., Collins, S.L., Zhao, Y., Zhang, T., Yang, X., An, H., Hu, G., Xin, C., Zhou, J., Sheng, X., He, M., Zhang, P., Guo, Z., Zhang, H., Li, L., Ma, M., 2023a. Warming reduced flowering synchrony and extended community flowering season in an alpine meadow on the Tibetan plateau. Ecology 104. https://doi.org/10.1002/ecy.3862.
Chen, T., Lv, L., Wang, D., Zhang, J., Yang, Y., Zhao, Z., Wang, C., Guo, X., Chen, H., Wang, Q., 2023b. Revolutionizing Agrifood systems with artificial intelligence: a survey. arXiv preprint arXiv:2305.01899.
Chen, L., Ren, C., Zhang, B., Wang, Z., Man, W., Liu, M., 2023c. Improved object-based mapping of aboveground biomass using geographic stratification with GEDI data and multi-sensor imagery. Remote Sens. 15, 2625.
David, R.M., Rosser, N.J., Donoghue, D.N.M., 2022. Improving above ground biomass estimates of southern Africa dryland forests by combining Sentinel-1 SAR and Sentinel-2 multispectral imagery. Remote Sens. Environ. 282, 113232.
Ding, C., Cao, X., Yu, B., Ju, Y., 2021. Non-linear associations between zonal built environment attributes and transit commuting mode choice accounting for spatial heterogeneity. Transp. Res. Part A Policy Pract 148, 22–35.
Do, A.N.T., Tran, H.D., Ashley, M., Nguyen, A.T., 2022. Monitoring landscape fragmentation and aboveground biomass estimation in can Gio mangrove biosphere reserve over the past 20 years. Ecol. Inform. 70 https://doi.org/10.1016/j.ecoinf.2022.101743.
Doughty, C.L., Ambrose, R.F., Okin, G.S., Cavanaugh, K.C., 2021. Characterizing spatial variability in coastal wetland biomass across multiple scales using UAV and satellite imagery. Remote Sens. Ecol. Conserv. 7, 411–429.
Ecke, S., Dempewolf, J., Frey, J., Schwaller, A., Endres, E., Klemmt, H.-J., Tiede, D., Seifert, T., 2022. UAV-based forest health monitoring: a systematic review. Remote Sens. 14, 3205.
Ellis, J., Moarif, S., 2015. Identifying and Addressing Gaps in the UNFCCC Reporting Framework.
Erasmi, S., Rosenbauer, R., Buchbach, R., Busche, T., Rutishauser, S., 2014. Evaluating the quality and accuracy of TanDEM-X digital elevation models at archaeological sites in the Cilician plain, Turkey. Remote Sens. 6, 9475–9493.
Fan, X., He, G., Zhang, W., Long, T., Zhang, X., Wang, G., Sun, G., Zhou, H., Shang, Z., Tian, D., 2022. Sentinel-2 images based modeling of grassland above-ground biomass using random forest algorithm: a case study on the Tibetan plateau. Remote Sens. 14, 5321.
Fararoda, R., Reddy, R.S., Rajashekar, G., Chand, T.R.K., Jha, C.S., Dadhwal, V.K., 2021. Improving forest above ground biomass estimates over Indian forests using multi source data sets with machine learning algorithm. Ecol. Inform. 65 https://doi.org/10.1016/j.ecoinf.2021.101392.
Frazier, A.E., Hemingway, B.L., 2021. A technical review of planet smallsat data: practical considerations for processing and using planetscope imagery. Remote Sens. 13, 3930.
Freeman, E.A., Moisen, G.G., Coulston, J.W., Wilson, B.T., 2016a. Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing tuning processes and model performance. Can. J. For. Res. 46, 323–339.
Freeman, E.A., Moisen, G.G., Coulston, J.W., Wilson, B.T., 2016b. Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing tuning processes and model performance. Can. J. For. Res. 46, 323–339.
Fujiki, S., 2017. Evaluation of Large-Scale Spatiotemporal Changes in the Tree-Community Composition of Bornean Rain Forests Using Remote Sensing Techniques.
Furniss, T.J., Kane, V.R., Larson, A.J., Lutz, J.A., 2020. Detecting tree mortality with Landsat-derived spectral indices: improving ecological accuracy by examining uncertainty. Remote Sens. Environ. 237, 111497.
Galidaki, G., Zianis, D., Gitas, I., Radoglou, K., Karathanassi, V., Tsakiri-Strati, M., Woodhouse, I., Mallinis, G., 2017. Vegetation biomass estimation with remote sensing: focus on forest and other wooded land over the Mediterranean ecosystem. Int. J. Remote Sens. 38, 1940–1966.
Gandhi, G.M., Parthiban, B.S., Thummalu, N., Christy, A., 2015. Ndvi: vegetation change detection using remote sensing and gis–a case study of Vellore District. Procedia Comput. Sci. 57, 1199–1210.
Gao, Y., Lu, D., Li, G., Wang, G., Chen, Q., Liu, L., Li, D., 2018a. Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. Remote Sens. 10, 627.
Gao, Y., Lu, D., Li, G., Wang, G., Chen, Q., Liu, L., Li, D., 2018b. Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. Remote Sens. 10, 627.
Ge, J., Hou, M., Liang, T., Feng, Q., Meng, X., Liu, J., Bao, X., Gao, H., 2022a. Spatiotemporal dynamics of grassland aboveground biomass and its driving factors in North China over the past 20 years. Sci. Total Environ. 826, 154226.
Ge, J., Hou, M., Liang, T., Feng, Q., Meng, X., Liu, J., Bao, X., Gao, H., 2022b. Spatiotemporal dynamics of grassland aboveground biomass and its driving factors in North China over the past 20 years. Sci. Total Environ. 826, 154226.
Gertz, M., Große-Butenuth, K., Junge, W., Maassen-Francke, B., Renner, C., Sparenberg, H., Krieter, J., 2020. Using the XGBoost algorithm to classify neck and leg activity sensor data using on-farm health recordings for locomotor-associated diseases. Comput. Electron. Agric. 173, 105404.
Gibson, J., 2018. Forest loss and economic inequality in the Solomon Islands: using small-area estimation to link environmental change to welfare outcomes. Ecol. Econ. 148, 66–76.
Gogoi, A., Ahirwal, J., Sahoo, U.K., 2022. Evaluation of ecosystem carbon storage in major forest types of eastern Himalaya: implications for carbon sink management. J. Environ. Manag. 302, 113972.
Guo, Q., Du, S., Jiang, J., Guo, W., Zhao, H., Yan, X., Zhao, Y., Xiao, W., 2023. Combining GEDI and sentinel data to estimate forest canopy mean height and aboveground biomass. Ecol. Inform. 78 https://doi.org/10.1016/j.ecoinf.2023.102348.
Haider, K., Khokhar, M.F., Chishtie, F., RazzaqKhan, W., Hakeem, K.R., 2017. Identification and future description of warming signatures over Pakistan with special emphasis on evolution of CO 2 levels and temperature during the first decade of the twenty-first century. Environ. Sci. Pollut. Res. 24, 7617–7629.
He, B., Xing, M., Bai, X., 2014. A synergistic methodology for soil moisture estimation in an alpine prairie using radar and optical satellite data. Remote Sens. 6, 10966–10985.
Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., 2017. SoilGrids250m: global gridded soil information based on machine learning. PLoS One 12, e0169748.
Hojo, A., Avtar, R., Nakaji, T., Tadono, T., Takagi, K., 2023. Modeling forest above-ground biomass using freely available satellite and multisource datasets. Ecol. Inform. 74 https://doi.org/10.1016/j.ecoinf.2023.101973.
Hu, Y.H., Yu, S.C., Qi, X., Zheng, W.J., Wang, Q.Q., Yao, H.Y., 2019. An overview of multiple linear regression model and its application. Zhonghua Yu Fang Yi Xue Za Zhi 53, 653–656.
Hussain, K., Mehmood, K., Anees, S.A., Ding, Z., Muhammad, S., Badshah, T., Shahzad, F., Haidar, I., Wahab, A., Ali, J., Ansari, M.J., 2024a. Assessing Forest Fragmentation due to Land use Changes from 1992 to 2023: A Spatio-Temporal Analysis Using Remote Sensing Data. Heliyon. https://doi.org/10.1016/j.heliyon.2024.e34710.
Hussain, K., Mehmood, K., Yujun, S., Badshah, T., Anees, S.A., Shahzad, F., Nooruddin, Ali J., Bilal, M., 2024b. Analysing LULC transformations using remote sensing data: insights from a multilayer perceptron neural network approach. Ann. GIS. https://doi.org/10.1080/19475683.2024.2343399.
Ismail, I., Sohail, M., Gilani, H., Ali, A., Hussain, Kiramat, Hussain, Kamran, Karky, B.S., Qamer, F.M., Qazi, W., Ning, W., 2018. Forest inventory and analysis in Gilgit-Baltistan: a contribution towards developing a forest inventory for all Pakistan. Int. J. Clim. Chang. Strateg. Manag. 10, 616–631.
Jackson, R.D., Huete, A.R., 1991. Interpreting vegetation indices. Prev. Vet. Med. 11, 185–200.
Jallat, H., Khokhar, M.F., Kudus, K.A., Nazre, M., Saqib, N.U., Tahir, U., Khan, W.R., 2021. Monitoring carbon stock and land-use change in 5000-year-old juniper forest stand of Ziarat, Balochistan, through a synergistic approach. Forests 12 (1), 51.

Ji, L., Zhang, L., Rover, J., Wylie, B.K., Chen, X., 2014. Geostatistical estimation of signal-to-noise ratios for spectral vegetation indices. ISPRS J. Photogramm. Remote Sens. 96, 20–27.
Jiang, F., Sun, H., Ma, K., Fu, L., Tang, J., 2022. Improving aboveground biomass estimation of natural forests on the Tibetan plateau using spaceborne LiDAR and machine learning algorithms. Ecol. Indic. 143, 109365.
Kaasalainen, S., Holopainen, M., Karjalainen, M., Vastaranta, M., Kankare, V., Karila, K., Osmanoglu, B., 2015. Combining lidar and synthetic aperture radar data to estimate forest biomass: status and prospects. Forests 6, 252–270.
Kaveh, N., Ebrahimi, A., Asadi, E., 2023. Comparative analysis of random forest, exploratory regression, and structural equation modeling for screening key environmental variables in evaluating rangeland above-ground biomass. Ecol. Inform. 77 https://doi.org/10.1016/j.ecoinf.2023.102251.
Khan, W.R., Rasheed, F., Zulkifli, S.Z., Kasim, M.R.B.M., Zimmer, M., Pazi, A.M., Kamrudin, N.A., Zafar, Z., Faridah-Hanum, I., Nazre, M., 2020. Phytoextraction potential of *Rhizophora apiculata*: a case study in Matang mangrove forest reserve, Malaysia. Trop. Conserv. Sci. 13, 1940082920947344.
Khan, I.A., Khan, W.R., Ali, A., Nazre, M., 2021. Assessment of above-ground biomass in Pakistan Forest Ecosystem's carbon Pool: a review. Forests 12, 586. REDD+ 360 1.
Khan, W.R., Nazre, M., Akram, S., Anees, S.A., Mehmood, K., Ibrahim, F.H., Al Edrus, S.S. O., Latiff, A., Fitri, Z.A., Yaseen, M., Li, P., Zhu, X., 2024. Assessing the productivity of the Matang mangrove forest reserve: review of one of the best-managed mangrove forests. Forests 15, 747. https://doi.org/10.3390/f15050747.
Khunrattanasiri, W., 2023. Application of remote sensing vegetation indices for forest cover assessments. In: Concepts and Applications of Remote Sensing in Forestry. Springer, pp. 153–166.
Kristensen, T., Næsset, E., Ohlson, M., Bolstad, P.V., Kolka, R., 2015. Mapping above-and below-ground carbon pools in boreal forests: the case for airborne lidar. PLoS One 10, e0138450.
Krzywinski, M., Altman, N., 2015. Multiple linear regression: when multiple variables are associated with a response, the interpretation of a prediction equation is seldom simple. Nat. Methods 12, 1103–1105.
Lee, H., Wang, J., Leblon, B., 2020. Using linear regression, random forests, and support vector machine with unmanned aerial vehicle multispectral images to predict canopy nitrogen weight in corn. Remote Sens. 12, 2071.
Lemenkova, P., 2020. Hyperspectral vegetation indices calculated by Qgis using Landsat tm image: a case study of northern Iceland. Adv. Res. Life Sci. 4, 70–78.
Li, C., Li, Y., Li, M., 2019a. Improving forest aboveground biomass (AGB) estimation by incorporating crown density and using landsat 8 OLI images of a subtropical forest in Western Hunan in Central China. Forests 10, 104.
Li, Y., Li, C., Li, M., Liu, Z., 2019b. Influence of variable selection and forest type on forest aboveground biomass estimation using machine learning algorithms. Forests 10, 1073.
Li, Yudong, Feng, Z., Chen, S., Zhao, Z., Wang, F., 2020. Application of the artificial neural network and support vector machines in forest fire prediction in the Guangxi autonomous region, China. Discret. Dyn. Nat. Soc. 2020, 1–14.
Li, Yingchang, Li, M., Li, C., Liu, Z., 2020a. Forest aboveground biomass estimation using Landsat 8 and sentinel-1A data with machine learning algorithms. Sci. Rep. 10, 9952.
Li, Yingchang, Li, M., Liu, Z., Li, C., 2020b. Combining kriging interpolation to improve the accuracy of forest aboveground biomass estimation using remote sensing data. IEEE Access 8, 128124–128139.
Li, Z., Jiao, Z., Wang, C., Yin, S., Guo, J., Tong, Y., Gao, G., Tan, Z., Chen, S., 2023. Seasonal effect of the vegetation clumping index on gross primary productivity estimated by a two-leaf light use efficiency model. Remote Sens. 15 https://doi.org/10.3390/rs15235537.
Liu, Y., Qian, J., Yue, H., 2020. Combined sentinel-1A with sentinel-2A to estimate soil moisture in farmland. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 1292–1310.
Liu, H., Jin, Y., Roche, L.M., O'Geen, T., Dahlgren, R.A., 2021. Understanding spatial variability of forage production in California grasslands: delineating climate, topography and soil controls. Environ. Res. Lett. 16, 014043.
López-Serrano, P.M., Cárdenas Domínguez, J.L., Corral-Rivas, J.J., Jiménez, E., López-Sánchez, C.A., Vega-Nieva, D.J., 2019. Modeling of aboveground biomass with Landsat 8 OLI and machine learning in temperate forests. Forests 11, 11.
Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., Moran, E., 2016a. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. Int. J. Digit. Earth 9, 63–105.
Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., Moran, E., 2016b. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. Int. J. Digit. Earth 9, 63–105.
Lu, Z., Chen, P., Yang, Y., Zhang, S., Zhang, C., Zhu, H., 2023. Exploring quantification and analyzing driving force for spatial and temporal differentiation characteristics of vegetation net primary productivity in Shandong Province, China. Ecol. Indic. 153 https://doi.org/10.1016/j.ecolind.2023.110471.
Luo, M., Wang, Y., Xie, Y., Zhou, L., Qiao, J., Qiu, S., Sun, Y., 2021a. Combination of feature selection and catboost for prediction: the first application to the estimation of aboveground biomass. Forests 12, 216.
Luo, M., Wang, Y., Xie, Y., Zhou, L., Qiao, J., Qiu, S., Sun, Y., 2021b. Combination of feature selection and catboost for prediction: the first application to the estimation of aboveground biomass. Forests 12, 216.
Luo, M., Wang, Y., Xie, Y., Zhou, L., Qiao, J., Qiu, S., Sun, Y., 2021c. Combination of feature selection and catboost for prediction: the first application to the estimation of aboveground biomass. Forests 12, 216.
Luo, M., Wang, Y., Xie, Y., Zhou, L., Qiao, J., Qiu, S., Sun, Y., 2021d. Combination of feature selection and catboost for prediction: the first application to the estimation of aboveground biomass. Forests 12, 216.
Luo, M., Anees, S.A., Huang, Q., Qin, X., Qin, Z., Fan, J., Han, G., Zhang, L., Shafri, H.Z. M., 2024. Improving Forest above-ground biomass estimation by integrating individual machine learning models. Forests 15 (6), 975.
Magarik, Y.A.S., 2021. "Roughly speaking": why Do US foresters measure DBH at 4.5 feet? Soc. Nat. Resour. 34, 725–744.
Maharjan, S., 2012. Estimation and Mapping above Ground Woody Carbon Stocks Using Lidar Data and Digital Camera Imagery in the Hilly Forests of Gorkha, Nepal.
Malakar, N.K., Hulley, G.C., Hook, S.J., Laraby, K., Cook, M., Schott, J.R., 2018. An operational land surface temperature product for Landsat thermal data: methodology and validation. IEEE Trans. Geosci. Remote Sens. 56, 5717–5735.
Martinuzzi, S., Cook, B.D., Helmer, E.H., Keller, M., Locke, D.H., Marcano-Vega, H., Uriarte, M., Morton, D.C., 2022. Patterns and controls on island-wide aboveground biomass accumulation in second-growth forests of Puerto Rico. Biotropica 54, 1146–1159.
Masek, J.G., Wulder, M.A., Markham, B., McCorkel, J., Crawford, C.J., Storey, J., Jenstrom, D.T., 2020. Landsat 9: empowering open science and applications through continuity. Remote Sens. Environ. 248, 111968.
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., 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 https://doi.org/10.1038/s41598-024-62464-7.
Mehmood, K., Anees, S.A., Rehman, A., Pan, S., Tariq, A., Zubair, M., Liu, Q., Rabbi, F., Khan, K.A., Luo, M., 2024c. Exploring spatiotemporal dynamics of NDVI and climate-driven responses in ecosystems: insights for sustainable management and climate resilience. Ecol. Inform. 80, 102532 https://doi.org/10.1016/J.ECOINF.2024.102532.
Mehmood, K., Anees, S.A., Rehman, A., Tariq, A., Liu, Q., Muhammad, S., Rabbi, F., Pan, S., Hatamleh, W.A., 2024d. Assessing forest cover changes and fragmentation in the Himalayan temperate region: implications for forest conservation and management. J. For. Res. (Harbin) 35, 82. https://doi.org/10.1007/s11676-024-01734-6.
Meng, S., Pang, Y., Zhang, Z., Jia, W., Li, Z., 2016. Mapping aboveground biomass using texture indices from aerial photos in a temperate forest of northeastern China. Remote Sens. 8, 230.
Mi, X., Feng, G., Hu, Y., Zhang, J., Chen, L., Corlett, R.T., Hughes, A.C., Pimm, S., Schmid, B., Shi, S., 2021. The global significance of biodiversity science in China: An overview. Natl. Sci. Rev. 8, nwab032.
Mngadi, M., Odindi, J., Mutanga, O., Sibanda, M., 2022. Estimating aboveground net primary productivity of reforested trees in an urban landscape using biophysical variables and remotely sensed data. Sci. Total Environ. 802, 149958.
Mohd Zaki, N.A., Abd Latif, Z., 2017. Carbon sinks and tropical forest biomass estimation: a review on role of remote sensing in aboveground-biomass modelling. Geocarto Int. 32, 701–716.
Montesano, P.M., Rosette, J., Sun, G., North, P., Nelson, R.F., Dubayah, R.O., Ranson, K. J., Kharuk, V., 2015. The uncertainty of biomass estimates from modeled ICESat-2 returns across a boreal forest gradient. Remote Sens. Environ. 158, 95–109.
Mourad, R., Jaafar, H., Anderson, M., Gao, F., 2020. Assessment of leaf area index models using harmonized landsat and sentinel-2 surface reflectance data over a semi-arid irrigated landscape. Remote Sens. 12, 3121.
Naik, P., Dalponte, M., Bruzzone, L., 2021. Prediction of forest aboveground biomass using multitemporal multispectral remote sensing data. Remote Sens. 13, 1282.
Nandy, S., Srinet, R., Padalia, H., 2021. Mapping forest height and aboveground biomass by integrating ICESat-2, Sentinel-1 and Sentinel-2 data using random forest algorithm in northwest Himalayan foothills of India. Geophys. Res. Lett. 48, e2021GL093799.
Nelson, P.R., Maguire, A.J., Pierrat, Z., Orcutt, E.L., Yang, D., Serbin, S., Frost, G.V., Macander, M.J., Magney, T.S., Thompson, D.R., 2022. Remote sensing of tundra ecosystems using high spectral resolution reflectance: opportunities and challenges. Eur. J. Vasc. Endovasc. Surg. 127, e2021JG006697.
Nguyen, T.D., Kappas, M., 2020. Estimating the aboveground biomass of an evergreen broadleaf forest in Xuan Lien nature reserve, Thanh Hoa, Vietnam, using SPOT-6 data and the random forest algorithm. Int. J. For. Res. 2020, 1–13.
Ojoatre, S., Zhang, C., Hussin, Y.A., Kloosterman, H.E., Ismail, M.H., 2019. Assessing the uncertainty of tree height and aboveground biomass from terrestrial laser scanner and hypsometer using airborne LiDAR data in tropical rainforests. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12, 4149–4159.
Oliveras, I., Malhi, Y., 2016. Many shades of green: the dynamic tropical forest–savannah transition zones. Philos. Trans. R. Soc. B 371, 20150308.
Pan, S.A., Anees, S.A., Li, X., Yang, X., Duan, X., Li, Z., 2023. Spatial and temporal patterns of non-structural carbohydrates in Faxon fir (Abies fargesii var. faxoniana), Subalpine Mountains of Southwest China. Forests 14 (7), 1438.
Pérez-Girón, J.C., Alvarez-Alvarez, P., Díaz-Varela, E.R., Lopes, D.M.M., 2020. Influence of climate variations on primary production indicators and on the resilience of forest ecosystems in a future scenario of climate change: application to sweet chestnut agroforestry systems in the Iberian Peninsula. Ecol. Indic. 113, 106199.
Prakash, A.J., Behera, M.D., Ghosh, S.M., Das, A., Mishra, D.R., 2022. A new synergistic approach for Sentinel-1 and PALSAR-2 in a machine learning framework to predict aboveground biomass of a dense mangrove forest. Ecol. Inform. 72 https://doi.org/10.1016/j.ecoinf.2022.101900.

Probst, P., Wright, M.N., Boulesteix, A., 2019. Hyperparameters and tuning strategies for random forest. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 9, e1301.
Puletti, N., Grotti, M., Ferrara, C., Chianucci, F., 2020. Lidar-based estimates of aboveground biomass through ground, aerial, and satellite observation: a case study in a Mediterranean forest. J. Appl. Remote. Sens. 14, 44501.
Purohit, S., Aggarwal, S.P., Patel, N.R., 2021a. Estimation of forest aboveground biomass using combination of Landsat 8 and sentinel-1A data with random forest regression algorithm in Himalayan foothills. Trop. Ecol. 62, 288–300.
Purohit, S., Aggarwal, S.P., Patel, N.R., 2021b. Estimation of forest aboveground biomass using combination of Landsat 8 and sentinel-1A data with random forest regression algorithm in Himalayan foothills. Trop. Ecol. 62, 288–300.
Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.-F., Minh, D.H.T., Lomas, M., D'alessandro, M.M., Paillou, P., Papathanassiou, K., 2019. The European Space Agency BIOMASS mission: measuring forest above-ground biomass from space. Remote Sens. Environ. 227, 44–60.
Rahman, A., Khan, N., Ali, K., Ullah, R., Khan, M.E.H., Jones, D.A., Rahman, I.U., 2021. Plant species classification and diversity of the understory vegetation in oak forests of swat, Pakistan. Appl. Sci. 11, 11372.
Rana, P., Gautam, B., Tokola, T., 2016. Optimizing the number of training areas for modeling above-ground biomass with ALS and multispectral remote sensing in subtropical Nepal. Int. J. Appl. Earth Obs. Geoinf. 49, 52–62.
Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M., 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol. Rev. 71, 804–818.
Rodríguez-Veiga, P., Wheeler, J., Louis, V., Tansey, K., Balzter, H., 2017. Quantifying forest biomass carbon stocks from space. Curr. For. Rep. 3, 1–18.
Ryu, S.-E., Shin, D.-H., Chung, K., 2020. Prediction model of dementia risk based on XGBoost using derived variable extraction and hyper parameter optimization. IEEE Access 8, 177708–177720.
Sa, R., Fan, W., 2023. Estimation of Forest parameters in boreal artificial coniferous forests using Landsat 8 and sentinel-2A. Remote Sens. 15, 3605.
Salinas-Melgoza, M.A., Skutsch, M., Lovett, J.C., 2018. Predicting aboveground forest biomass with topographic variables in human-impacted tropical dry forest landscapes. Ecosphere 9, e02063.
Shahzad, F., Mehmood, K., Hussain, K., Haidar, I., Anees, S.A., Muhammad, S., Ali, J., Adnan, M., Wang, Z., Feng, Z., 2024. Comparing machine learning algorithms to predict vegetation fire detections in Pakistan. Fire Ecol. 20 https://doi.org/10.1186/s42408-024-00289-5.
Sharifi, A., Amini, J., Tateishi, R., 2016. Estimation of forest biomass using multivariate relevance vector regression. Photogramm. Eng. Remote. Sens. 82, 41–49.
Shobairi, S.O.R., Lin, H., Usoltsev, V.A., Osmirko, A.A., Tsepordey, I.S., Ye, Z., Anees, S. A., 2022. A comparative pattern for Populus spp. and Betula spp. stand biomass in Eurasian climate gradients. Croatian journal of Forest engineering: journal for theory and application of forestry. Engineering 43 (2), 457–467.
Silleos, N.G., Alexandridis, T.K., Gitas, I.Z., Perakis, K., 2006. Vegetation indices: advances made in biomass estimation and vegetation monitoring in the last 30 years. Geocarto Int. 21, 21–28.
Sinha, S., Jeganathan, C., Sharma, L.K., Nathawat, M.S., 2015. A review of radar remote sensing for biomass estimation. Int. J. Environ. Sci. Technol. 12, 1779–1792.
Sreehari, E., Srivastava, S., 2018. Prediction of climate variable using multiple linear regression. In: 2018 4th International Conference on Computing Communication and Automation (ICCCA). IEEE, pp. 1–4.
Strandberg, R., Låås, J., 2019. A Comparison between Neural Networks, Lasso Regularized Logistic Regression, and Gradient Boosted Trees in Modeling Binary Sales.
Su, H., Shen, W., Wang, J., Ali, A., Li, M., 2020a. Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests. For. Ecosyst. 7, 1–20.
Su, H., Shen, W., Wang, J., Ali, A., Li, M., 2020b. Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests. For. Ecosyst. 7, 1–20.
Sun, X., Li, G., Wang, M., Fan, Z., 2019. Analyzing the uncertainty of estimating forest aboveground biomass using optical imagery and spaceborne LiDAR. Remote Sens. 11, 722.
Sun, C., Li, J., Liu, Y., Zhao, S., Zheng, J., Zhang, S., 2023. Tracking annual changes in the distribution and composition of saltmarsh vegetation on the Jiangsu coast of China using landsat time series–based phenological parameters. Remote Sens. Environ. 284, 113370.
Suratman, M.N., Abd Latiff, Z., Tengku Hashim, T.M.Z., Mohsin, A.F., Asari, N., Mohd Zaki, N.A., 2023. Remote sensing for forest inventory and resource assessment. In: Concepts and Applications of Remote Sensing in Forestry. Springer, pp. 3–23.
Taghizadeh-Mehrjardi, R., Schmidt, K., Amirian-Chakan, A., Rentschler, T., Zeraatpisheh, M., Sarmadian, F., Valavi, R., Davatgar, N., Behrens, T., Scholten, T., 2020. Improving the spatial prediction of soil organic carbon content in two contrasting climatic regions by stacking machine learning models and rescanning covariate space. Remote Sens. 12, 1095.
Tamiminia, H., Salehi, B., Mahdianpari, M., Beier, C.M., Johnson, L., 2022. Evaluating pixel-based and object-based approaches for Forest above-ground biomass estimation using a combination of optical, SAR, and AN extreme gradient boosting model. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 3, pp. 485–492.
Tamiminia, H., Salehi, B., Mahdianpari, M., Goulden, T., 2024. State-wide forest canopy height and aboveground biomass map for New York with 10 m resolution, integrating GEDI, Sentinel-1, and Sentinel-2 data. Ecol. Inform. 79 https://doi.org/10.1016/j.ecoinf.2023.102404.
Teng, H., Chen, S., Hu, B., Shi, Z., 2023. Future changes and driving factors of global peak vegetation growth based on CMIP6 simulations. Ecol. Inform. 75 https://doi.org/10.1016/j.ecoinf.2023.102031.
Thompson, C.G., Kim, R.S., Aloe, A.M., Becker, B.J., 2017. Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. Basic Appl. Soc. Psychol. 39, 81–90.
Timothy, D., Onisimo, M., Riyad, I., 2016. Quantifying aboveground biomass in African environments: a review of the trade-offs between sensor estimation accuracy and costs. Trop. Ecol. 57, 393–405.
Trautenmüller, J.W., Netto, S.P., Balbinot, R., David, H.C., Dalla Corte, A.P., Watzlawick, L.F., Sanquetta, C.R., Mallmann, A.A., Engel, K., Behling, A., 2023. Ratio estimators for aboveground biomass and its parts in subtropical forests of Brazil. Ecol. Indic. 154, 110530.
Tsitsi, B., 2016. Remote sensing of aboveground forest biomass: a review. Trop. Ecol. 57, 125–132.
Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A.K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., 2015. Free and open-access satellite data are key to biodiversity conservation. Biol. Conserv. 182, 173–176.
Usman, N., Hussain, M., Akram, S., Majeed, M., Shah, S., Rehman, F., Yousaf, A., Shaukat, S., Shah, S.W.A., Mishr, R.S., 2022. Yield, carbon stock, and price dynamics of agroforestry tree species in district Mardan, Khyber Pakhtunkhwa, Pakistan. Braz. J. Biol. 84.
Usoltsev, V.A., Chen, B., Shobairi, S.O.R., Tsepordey, I.S., Chasovskikh, V.P., Anees, S.A., 2020. Patterns for Populus spp. stand biomass in gradients of winter temperature and precipitation of Eurasia. Forests 11 (9), 906.
Usoltsev, V.A., Lin, H., Shobairi, S.O.R., Tsepordey, I.S., Ye, Z., Anees, S.A., 2022. The principle of space-for-time substitution in predicting Betula spp. biomass change related to climate shifts. Appl. Ecol. Environ. Res. 20 (4), 3683–3698.
Vaglio Laurin, G., Pirotti, F., Callegari, M., Chen, Q., Cuozzo, G., Lingua, E., Notarnicola, C., Papale, D., 2016. Potential of ALOS2 and NDVI to estimate forest above-ground biomass, and comparison with lidar-derived estimates. Remote Sens. 9, 18.
Vahedi, A.A., 2014. Optimal allometric biomass equations for hornbeam (*Carpinus betulus* L.) boles within the Hyrcanian forests. Iran. J. For. Poplar Res. 22.
Wai, P., Su, H., Li, M., 2022a. Estimating aboveground biomass of two different forest types in Myanmar from sentinel-2 data with machine learning and geostatistical algorithms. Remote Sens. 14, 2146.
Wai, P., Su, H., Li, M., 2022b. Estimating aboveground biomass of two different forest types in Myanmar from sentinel-2 data with machine learning and geostatistical algorithms. Remote Sens. 14, 2146.
Wang, J., Xiao, X., Bajgain, R., Starks, P., Steiner, J., Doughty, R.B., Chang, Q., 2019. Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1, Sentinel-2 and Landsat images. ISPRS J. Photogramm. Remote Sens. 154, 189–201.
Wu, C., Shen, H., Shen, A., Deng, J., Gan, M., Zhu, J., Xu, H., Wang, K., 2016. Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery. J. Appl. Remote. Sens. 10, 35010.
Wulder, M.A., Roy, D.P., Radeloff, V.C., Loveland, T.R., Anderson, M.C., Johnson, D.M., Healey, S., Zhu, Z., Scambos, T.A., Pahlevan, N., 2022. Fifty years of Landsat science and impacts. Remote Sens. Environ. 280, 113195.
Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J.A., Huete, A.R., Ichii, K., Ni, W., Pang, Y., Rahman, A.F., 2019. Remote sensing of the terrestrial carbon cycle: a review of advances over 50 years. Remote Sens. Environ. 233, 111383.
Xie, H., Zhang, Y., Zeng, X., He, Y., 2020. Sustainable land use and management research: a scientometric review. Landsc. Ecol. 35, 2381–2411.
Yang, B., Zhang, Y., Mao, X., Lv, Y., Shi, F., Li, M., 2022. Mapping spatiotemporal changes in forest type and aboveground biomass from landsat long-term time-series analysis—a case study from Yaoluoping National Nature Reserve, Anhui province of Eastern China. Remote Sens. 14, 2786.
Yimam, Y.M., Kifle, B., 2020. Tree inventory assessment in religious institution compound and its benefit for environmental management in the case of Addis Ababa, Ethiopia. Int. J. Sci. Res. Biol. Sci. 7.
Yu, Y., Pan, Y., Yang, X., Fan, W., 2022. Spatial scale effect and correction of forest aboveground biomass estimation using remote sensing. Remote Sens. 14, 2828.
Yuan, X., Li, L., Tian, X., Luo, G., Chen, X., 2016. Estimation of above-ground biomass using MODIS satellite imagery of multiple land-cover types in China. Remote Sens. Lett. 7, 1141–1149.
Zaher, H., Sabir, M., Benjelloun, H., Paul-Igor, H., 2020. Effect of forest land use change on carbohydrates, physical soil quality and carbon stocks in Moroccan cedar area. J. Environ. Manag. 254, 109544.
Zhang, Y., Shao, Z., 2021. Assessing of urban vegetation biomass in combination with LiDAR and high-resolution remote sensing images. Int. J. Remote Sens. 42, 964–985.
Zhang, Y., Ma, J., Liang, S., Li, X., Li, M., 2020. An evaluation of eight machine learning regression algorithms for forest aboveground biomass estimation from multiple satellite data products. Remote Sens. 12, 4015.
Zhang, Y., Ma, J., Liang, S., Li, X., Liu, J., 2022. A stacking ensemble algorithm for improving the biases of forest aboveground biomass estimations from multiple remotely sensed datasets. GIsci Remote Sens 59, 234–249.
Zhao, J., Liu, D., Zhu, Y., Peng, H., Xie, H., 2022. A review of forest carbon cycle models on spatiotemporal scales. J. Clean. Prod. 339, 130692.
Zhou, X., Zhu, X., Dong, Z., Guo, W., 2016. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. Crop. J. 4, 212–219.

Zhu, X., Liu, D., 2015. Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. ISPRS J. Photogramm. Remote Sens. 102, 222–231.
Zhu, Y., Feng, Z., Lu, J., Liu, J., 2020a. Estimation of forest biomass in Beijing (China) using multisource remote sensing and forest inventory data. Forests 11, 163.
Zhu, Y., Feng, Z., Lu, J., Liu, J., 2020b. Estimation of forest biomass in Beijing (China) using multisource remote sensing and forest inventory data. Forests 11, 163.