Agroforest Syst (2023) 97:1097–1111
https://doi.org/10.1007/s10457-023-00850-2

Remote sensing and machine learning applications for aboveground biomass estimation in agroforestry systems: a review

Bhuwan Thapa · Sarah Lovell · Jeffrey Wilson

Received: 12 November 2021 / Accepted: 25 April 2023 / Published online: 10 May 2023

Abstract The estimation of aboveground biomass (AGB) in agroforestry systems using remote sensing has proliferated in the last decades. Similarly, machine learning is also being used in AGB assessments. This study reviews the applications of remote sensing and machine learning for AGB estimation in agroforestry systems (AFS). A detailed review was conducted using 33 recent papers by extracting and comparing information on agroforestry type, data sources, methodology, and model accuracy. Statistical tests were performed to evaluate the differences in performances. High- and very-high-resolution imageries (less than 2 m) are widely used for AGB assessment because they helped to delineate heterogeneous features of AFS. Object-based image analysis yielded classification accuracy of up to 90 percent in some cases. Random Forest, Stochastic Gradient Boosting, and Support Vector Regression are the most common algorithms used for AGB estimation. However, there are no statistically significant differences in the performance between machine learning and other models. Similarly, scholars incorporated spectral indices with spectral bands, texture, and biophysical variables as covariate categories into AGB estimation models. The study finds no significant differences in results (R-squared) by adding more covariate categories. The accuracy of AGB estimates depends upon multiple factors, such as the spectral and spatial resolution, number and types of covariates, methods for AFS delineation and AGB estimation, and types and sizes of AFS. Despite some of the methodological challenges around measuring understory vegetation, advancements in cloud computing like Google Earth Engine and the availability of high-resolution datasets present opportunities for wider use of remote sensing for biomass estimation of AFS. Remote sensing and machine learning have the potential to estimate aboveground biomass over a large area with high accuracy and contribute to carbon monitoring.

Keywords Spectral indices · LiDAR · Carbon · Object-based classification · Random forest

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10457-023-00850-2.

B. Thapa (✉) · S. Lovell
Center for Agroforestry, University of Missouri, Columbia, USA
e-mail: thapa108@gmail.com; thapab@appstate.edu

B. Thapa
Department of Geography and Planning, Appalachian State University, Boone, USA

J. Wilson
Department of Geography, Indiana University–Purdue University Indianapolis, Indianapolis, USA

Introduction

Agroforestry systems (AFS) consist of landuse practices that intentionally integrate trees and shrubs into

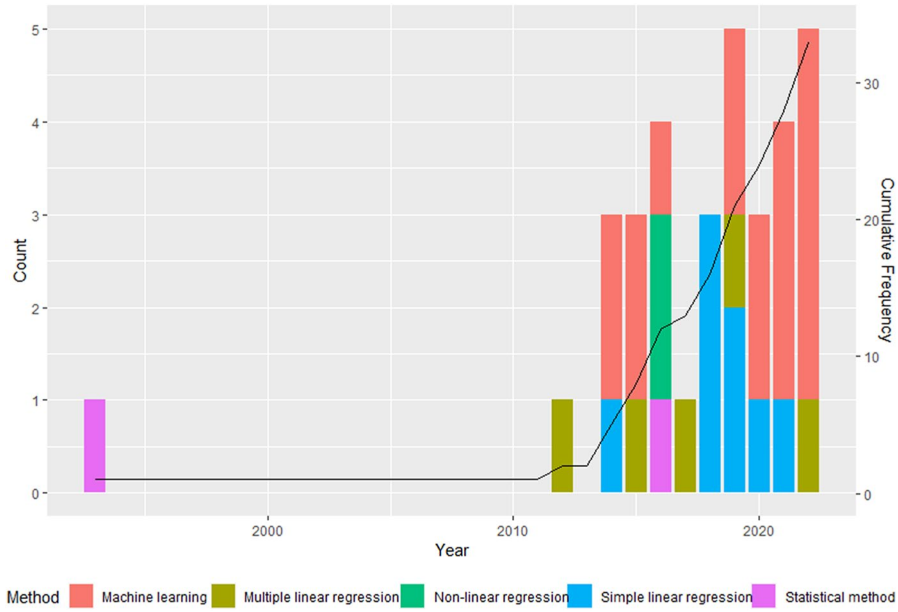

Fig. 1 Cumulative and yearly frequency of the reviewed papers on AGB estimation by different methods

crop and/or animal farming systems. Five broad categories of AFS sequester more carbon than traditional agriculture (Wilson and Lovell 2016). *Silvopasture* integrates trees in pastureland with livestock and can store 45 percent more aboveground biomass (AGB) than perennial pasture (Udawatta and Jose 2012). While rows of trees and/or shrubs are integrated with agronomic or horticultural crops in *alley cropping*, *windbreaks and shelterbelts* are planted adjacent to crops or grazing areas in a linear fashion for wind protection. *Riparian buffers* allow perennial vegetation adjacent to water bodies. The most complex form of AFS is *forest farming*, which includes a diverse mix of perennial species modeled to mimic natural woodlands (Gene Garrett and Buck 1997). Multilayer understory crops, woody biomass, and soil carbon in AFS can sequester around 0.29 to 15.21 Mg/ha/yr of AGB (Nair et al. 2010), making it one of the promising natural climate solutions (Griscom et al. 2017).

The promotion of AFS as a natural climate solution requires cost-effective monitoring of carbon biomass. There are two approaches to measuring AGB. While the *in-situ* measurement through field observations is more accurate than the *ex-situ* method that relies on remote instruments, it is resource intensive and have limited spatial coverage. Remote sensing is an *ex-situ* technique that assesses the physical characteristics of landcover by measuring the reflected and emitted radiation from a distance (Wang et al. 2019). It has been widely used in AGB estimation of forestry (Timothy et al. 2016; Pádua et al. 2017; Ahmad et al. 2021). AFS are more complicated than forestry because of understory crops, diverse vegetation, and heterogeneous structure (Udawatta and Jose 2012). Remote sensing has been applied in AFS since 1990s (Unruh et al. 1993; Houghton et al. 1993). However, with the greater access to high-resolution data and advances in AGB estimation methodologies, their use has increased significantly in the last decade (Fig. 1) (Czerepowicz et al. 2012; Chen et al. 2015; Schneider et al. 2018). Optical imageries, Light Detection and Ranging (LiDAR), Synthetic Aperture Radar (SAR), and other remotely derived imageries along with field observations provide information on canopy cover, tree height, and other characteristics need to accurately estimate AGB.

Machine learning (ML) can model complex spatial patterns using a variety of large input data with higher accuracy (Wu et al. 2016). It has been increasingly used in AGB estimation of forests (Maxwell et al. 2018; Zhang et al. 2020) and gradually applied in AFS as well (Filippi et al. 2014; Suchenwirth et al. 2014; Güneralp et al. 2014). With the increasing use of remote sensing and ML in AFS, there is a need to synthesize the emerging literature. The main goal of the study is to review remote sensing and machine learning techniques for AGB estimation of AFS. The paper is outlined as follows. The methods used for the literature review are briefly discussed. The main findings related to various factors affecting AGB performances, including data sources, methodological approaches, and covariates use are discussed next. The paper concludes with a discussion on the opportunities and challenges of using remote sensing in AFS. The synthesis can aid researchers, practitioners, and growers in better understanding the possibilities and limitations of remote sensing and machine learning applications for carbon monitoring.

Materials and methods

Keyword searches of AFS-related terms were conducted using major databases, including Google Scholar, Web of Science, Scopus, and related journals. The following query was used: [Types of AFS] AND [aboveground biomass] AND [Types of remote sensing]. A total of 125 manuscripts written between 1991 and 2022 were shortlisted for review of which 33 articles were selected for detailed review based on the inclusion/exclusion criteria (see supplementary materials for more details).

Results and discussion

The performance of AGB estimation depends upon multiple factors, such as the spectral and spatial resolution of data sources, types of covariates, methods of AFS delineation and AGB estimation, and types and sizes of AFS. Some of the key factors potentially affecting the accuracy of AGB, measured by the coefficient of determination (R-squared) are discussed as follows. Statistical tests like the t-test and Kruskal–Wallis test performed to evaluate the statistical significance of the findings are also discussed.

Effect of data sources

There are three types of remote sensing data used in AGB assessment of AFS: optical imageries, Light Detection and Ranging (LiDAR) and Synthetic Aparture Radar (SAR), and multi-source. Optical imageries are the most frequently used (31 out of 33 studies) data source. The visible, near-infrared, and short-wave infrared reflectance from objects generate vegetation indices, texture information, and other parameters needed for AGB assessment. While satellites like Landsat provide global coverage, their coarse spatial resolution reduces the accuracy (Timothy et al. 2016). Therefore, high (0.5–2.0 m) and very-high resolution (lower than 0.5 m) imageries are commonly used in AFS studies (Macedo et al. 2018). Optical imageries produced an average R-squared value of 0.696.

LiDAR and SAR are also used for AGB assessments. Nine out of thirty-three studies used LiDAR whereas only two studies used SAR. LiDAR transmits and receives pulses of laser energy to measure the height, canopy closure, volume, and biomass of forest stands (Gatziolis and Andersen 2008). When combined with other data sources, some LiDAR studies have measured AGB with high accuracy. For example, Kanmegne Tamga et al. (2022) estimated the AGB of food forest by including tree height and tree density information from multispectral and LiDAR data (R-squared = 0.91, RMSE = 3.78 Mg/ha). Accuracy is also affected by the level of aggregation. LiDAR information can be aggregated at the plot- or plant-level (Chen et al. 2015). While plot-level assessment aggregates the AGB of individual trees for a given area, tree-level assessment produces higher AGB as it provides disaggregated information of individual trees (Graves et al. 2018). Unlike LiDAR which is airborne, SAR is a space-borne sensor that emits and receives long wavelength energy produced by its sensors. Since the low frequency (L- and P-band) bands can penetrate through trees, it performs better than optical imageries. For example, Naidoo et al. (2016) improved the AGB accuracy (R-squared between 0.83 and 0.88) by integrating L-band HH and HV backscatter of SAR with Landsat image reflectance.

While optical and LiDAR data sources are useful, combining these data products can help integrate information on different aspects of AFS. A total of 18 out of 33 papers used more than one data sources (Table 1). Multispectral and hyperspectral imageries along with LiDAR and field observations provide nuanced information about the biophysical characteristics of vegetation and thereby improve the accuracy. Wang et al. (2016) estimated the AGB of individual trees in shelterbelts with accuracy of 97 percent and

Table 1 Summary of key variables of the reviewed literature

Sensors	Resolution (meter)	Agroforestry system	Study area	AFS delineation method	AGB estimation method	Covariate types	Precision & accuracy	References
Optical imageries								
Sentinel-2	60	Alley cropping	Columbia	Supervised classification	Machine learning	Spectral indices, texture, biophysical variables	CA = 94%, KC = 0.82; AGB R^2 = 0.92, RMSE = 31.04	(Bolívar-Santamaría and Reu 2021)
Sentinel-2	20	Alley cropping	India	Manual delineation using GIS and imagery	Machine learning	Spectral band, spectral indices	AGB R^2 = 0.86, RMSE = 7.50 Mg/ha	(Kalita et al. 2016)
Landsat 7 ETM +	30	Forest farming	Thailand	None	Non-linear regression	Biophysical variables	AGB R^2 = 0.95	(Laosuwan and Uttaruk 2016)
QuickBird	0.7	Forest farming	Portugal	OBIA	Simple linear regression	Spectral bands, spectral indices	CA = 0.89, KC = 0.75; AGB R^2 = 0.75	(Macedo et al. 2018)
Landsat ETM + 7	30	Forest farming	Indonesia	None	Multiple linear regression	Spectral bands, spectral indices	AGB R^2 = 0.44	(Prasondita et al. 2019)
NOAA	15,000	Forest farming	Sub-Saharan Africa	Unsupervised classification	Statistical method	Spectral indices	CA = 85–95%	(Unruh et al. 1993)
World-View 2	1.84	Forest farming*	Thailand	None	Machine learning	Spectral bands	AGB R^2 = 0.66, RMSE = 11.97	(Yasen and Koedsin 2015)
Sentinel-1, Sentinel-2	10	Forest farming*	Burkina Faso, Ghana, Togo and Benin	None	Machine learning	Spectral bands, spectral indices, biophysical variables	AGB R^2 = 0.90, RMSE = 54.5 Mg/ha	(Forkuor et al. 2020)
UAV	0.05	Forest farming*	China	None	Machine learning	Spectral indices, texture	AGB R^2 = 0.75, RMSE = 28.72 Mg/ha	(Liang et al. 2022)
UAV	0.025	Riparian forest	Italy	OBIA	Simple linear regression	Biophysical variables	AGB R^2 = 0.63	(Matese et al. 2021)
Earth Observation-1 Hyperion	30	Riparian forest	India	None	Simple linear regression	Spectral indices	AGB R^2 = 0.78	(Pandey et al. 2019)
QuickBird II	2.4,0.6	Shelterbelt	New Zealand	Manual delineation	Multiple linear regression	Spectral bands, spectral indices, biophysical variables	CA = 0.92, AGB R^2 = 0.80	(Czerepowicz et al. 2012)
Resoursesat-2/LISS-II	23.5	Silvopasture	India	Supervised landcover classification	Statistical method	None	CA = 0.82, KC = 0.80	(Rizvi et al. 2016)

Table 1 (continued)

Sensors	Resolution (meter)	Agroforestry system	Study area	AFS delineation method	AGB estimation method	Covariate types	Precision & accuracy	References
LiDAR								
LiDAR	0.81	Forest farming	Brazil	Manual delineation using GIS and imagery	Multiple linear regression	Tree height	AGB $R^2=0.47$, RMSE = 60.4 Mg/ha	(Chen 2015)
LiDAR	0.025	Forest farming*	Trinidad		Machine learning	Tree height	AGB $R^2=0.83$, RMSE = 0.062 Mg/ha	(Sankar et al. 2022)
Data fusion								
Sentinel-2, Pleiades	20,0.5	Alley cropping	Burkina Faso	Manual delineation	Simple linear regression	Spectral indices	AGB $R^2=0.80$, RMSE = 0.31 ton/ha	(Karlson et al. 2020)
Landsat-8	15,30	Forest farming	India	Hybrid classification	Multiple linear regression	Spectral indices	CA = 0.89, KC = 86.16%, AGB $R^2=0.79$, RMSE = 51.04 t/ha	(Bordoloi et al. 2022)
QuickBird, WorldView-2, ASTER	0.7,0.5,30	Forest farming	Portugal	OBIA	Simple linear regression	Biophysical variables	AGB $R^2=0.92$	(Gonçalves et al. 2019)
Sentinel-1, Sentinel-2, ALOS, LiDAR	10	Forest farming	Burkina Faso, Ivory Coast	OBIA	Machine learning	Spectral indices, texture	AGB $R^2=0.91$, RMSE = 3.78 Mg/ha	(Kanmegne Tamga et al. 2022)
Landsat 8, WorldView-2	30,15,0.5	Forest farming	Burkina Faso	OBIA	Machine learning	Spectral indices, texture	AGB $R^2=0.57$	(Karlson et al. 2015)
QuickBird, ASTER	0.65,30	Forest farming	El Salvador	None	Multiple linear regression	Reflectance band, texture	AGB $R^2=0.52$	(Kearney et al. 2017)
QuickBird, WorldView-2	0.7,0.5	Forest farming	Portugal	OBIA	Machine learning	Spectral bands, spectral indices, texture	AGB $R^2=0.82$	(Lourenço et al. 2021)
NAIP, (Naidoo et al. 2016) LiDAR	1	Forest farming*	USA		Machine learning	Spectral bands, spectral indices	AGB $R^2=0.37$, RMSE = 1.08 Mg/ha	(Ku and Popescu 2019)
ALOS PALSAR, Landsat-5, LiDAR	30	Forest farming*	South Africa		Machine learning	Spectral bands, spectral indices, texture	AGB $R^2=0.88$, RMSE = 8.51%	(Karlson 2016)

Table 1 (continued)

Sensors	Resolution (meter)	Agroforestry system	Study area	AFS delineation method	AGB estimation method	Covariate types	Precision & accuracy	References
LiDAR, AISAFenix	1	Riparian forest	Brazil	None	Machine learning	Spectral indices, biophysical variables, geomorphometric variables	AGB $R^2=0.7$	(Almeida et al. 2019)
Hyperion, LiDAR	30,1	Riparian forest	USA	None	Machine learning	Spectral indices, biophysical variables	MARS $R^2=0.97$, RMSE = 20.73 t/ha	(Filippi et al. 2014)
Landsat 7 ETM +, SPOT 5	30,10	Riparian forest	USA	None	Machine learning	Spectral bands, spectral indices, biophysical variables, geomorphometric variables	AGB $R^2=0.96$	(Güneralp et al. 2014)
ALOS-2 PALSAR-2, Sentinel-1, Sentinel-2	6,10–20	Riparian forest	Vietnam	None	Machine learning	Spectral bands, spectral indices	AGB $R^2=0.62$, RMSE = 27.4 Mg/ha	(Pham et al. 2020)
Landsat TM, SPOT 5	30,10,20	Riparian forest	Malaysia	Hybrid classification	Simple linear regression	Spectral bands, texture	AGB $R^2=0.91$	(Singh et al. 2014)
UAV, LiDAR	0.1	Riparian forest	China	None	Machine learning	Spectral indices, texture, biophysical variables	AGB $R^2=0.78$, RMSE = 25.51 Mg/ha	(Tian et al. 2021)
CASI-1500, LiDAR	1	Shelterbelt	China	Tree delineation using LiDAR	Non-linear regression	Biophysical variables	AGB $R^2=0.61$	(Wang et al. 2016)
Carnegie Airborne Observatory, LiDAR	1.12	Silvopasture	Panama	Forest crown cover delineation using LiDAR	Simple linear regression	Biophysical variables	AGB $R^2=0.43$, RMSE = 1800 kg	(Graves et al. 2018)
World-View 2, GeoEye	0.5	Silvopasture	Ecuador	OBIA	Simple linear regression	Biophysical variables	CA = 0.79–0.87; AGB $R^2=0.27$	(Schneider et al. 2018)

Forest farming*: Mono-plantations like rubber and cocoa

CA Classification accuracy, *KC* Kappa coefficient, *OBIA* Object-based image analysis, R^2 Coefficient of determination, *RMSE* Root mean square deviation, *UAV* Unmanned aerial vehicle

R-squared of 0.61 by combining airborne LiDAR and airborne spectrographic imagery. However, fusion of data products doesn't always improve performance. Luo et al. (2017) estimated the AGB of forest biomass by integrating LiDAR and hyperspectral imagery and found that spectral characteristics were good estimators of AGB and LiDAR only marginally improved the performance (R-squared increased by 5.8 percent). This is partly attributed to complementary information contained in these data products.

Methods for AFS delineation and AGB estimation

The AGB estimation using remote sensing generally involves the following steps. AFS are delineated automatically or semi-automatically using landcover classifications, or manually using GIS, remote sensing and field observation. AGB values are estimated for sample field plots through destructive or non-destructive sampling. The final step involves AGB estimation of the overall study area by establishing a relationship between field-based AGB and remote sensing and other variables (Wang et al. 2019).

For AFS delineation, five out of thirty-three studies used automated or semi-automated landcover classification. Along with spectral characteristics, scholars also used textural information to classify trees (Lourenço et al. 2021). Seven out of thirty-three studies used Object-based Image Analysis (OBIA) for AGB delineation. Unlike pixel-based classification that separates individual pixels directly, OBIA aggregates image pixels into spectrally homogenous image objects using an image segmentation algorithm (Fig. 2) (Liu and Xia 2010). OBIA had a classification accuracy between 79 and 89 percent.

The AGB estimation consisted of regression and machine learning models (Fig. 3). Regression models were used in 15 out of 33 studies whereas 16 studies used machine learning. Two studies utilized statistical methods that upscaled field-level AGB to a larger area. Overall, machine learning had an average R-squared of 0.815 whereas that for simple linear regression was 0.686 (Fig. 2). However, when a t-test was performed between two groups (machine learning and non-machine learning), there were no statistically significant differences between them.

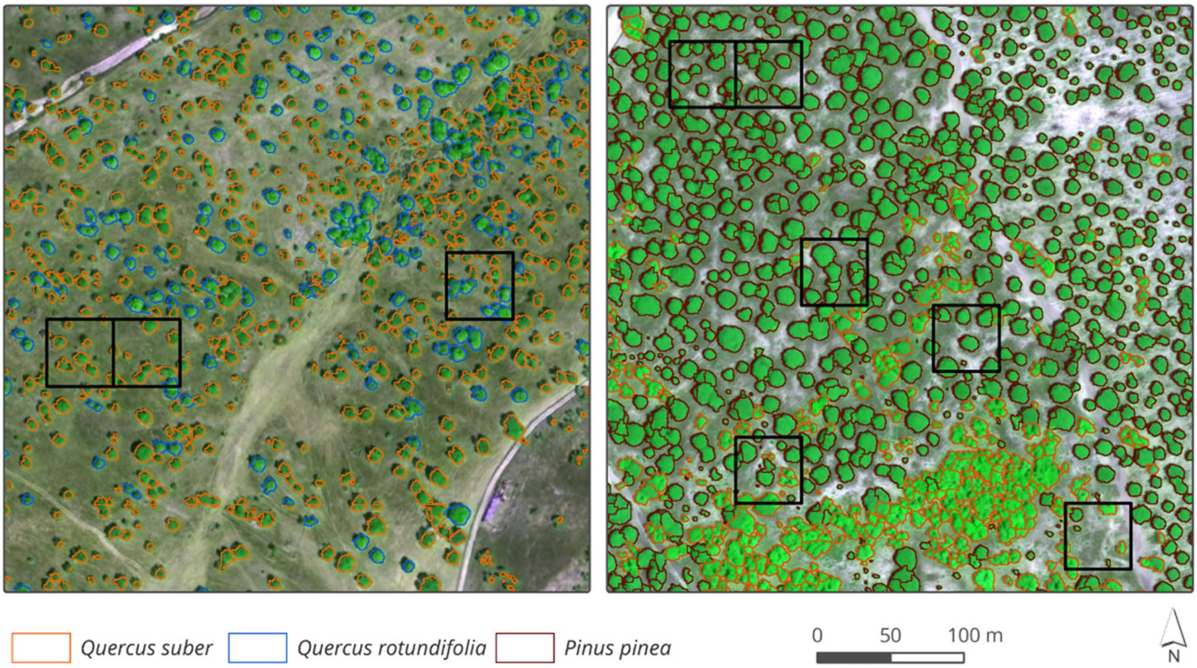

Fig. 2 Object-based classification of scattered trees (Source: Gonçalves et al. 2019)

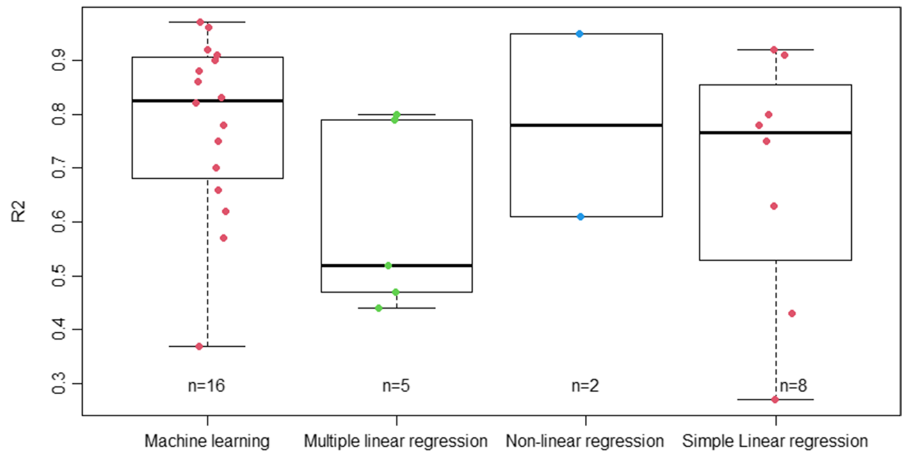

Fig. 3 R-squared by AGB estimation methods

Choice of covariates

The covariates used in AGB estimation models included spectral indices and bands, textures, biophysical variables, and geomorphometric variables (Table 2). Spectral indices are widely used in AGB assessments because of their high correlation with the vegetation index (Fig. 4) (Laosuwan and Uttaruk 2016). Of all the indices, the Normalized Difference Vegetation Index (NDVI) is the most frequently used index to measure vegetation greenness (Fig. 5). However, it can suffer from over-saturation and become insensitive to woody parts of AFS where most of the carbon is stored (Lu 2003). Indices like Atmospherically Resistant Vegetation Index (ARVI) reduces over-saturation by minimizing atmospheric aerosol brightness (Bordoloi et al. 2022).

Besides spectral indices, spectral bands, band ratios, and textures, along with biophysical and geomorphometric variables were also frequently used covariate types (Table 2). A total of 29 out of 33 papers utilized these variables. (Table 1). Among them, single band and band ratios are also good predictors of AGB. Prasondita et al. (2019) estimated the AGB of food forest using spectral bands and indices and found that single band 7 of Landsat 7 ETM+is the best predictor (R-squared = 0.44, RMSE = 52.85 tonnes/ha).

While spectral indices or spectral bands can provide some information on AGB, scholars have included more than one type of covariates (i.e., spectral, texture, biophysical, and geomorphometric). The average R-squared of studies that used more than three types of covariates was 0.87 whereas that for one type of covariate was 0.67 (Fig. 6). However, the result was not statistically significant when Kruskal–Wallis test was performed.

Machine learning

ML has been widely applied in AGB assessments of forestry (Chen et al. 2018), and they are slowly

Table 2 Highly cited covariates from the reviewed literature

Types of covariates	Covariates
Biophysical variables	Tree height, crown cover, crown diameter, Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), Leaf Area Index (LAI), Leaf Area Density (LAD)
Geomorphometric variables	Aspect, slope, elevation, Euclidean distance to water, Topographic Wetness Index (TWI), profile curvature
Spectral indices	Normalized Difference Vegetation Index (NDVI), Soil Sdjusted Vegetation index (SAVI), Enhanced Vegetation Index (EVI), Simple Ratio (SR), Normalized Difference Water Index (NDWI), Atmospherically Resistant Vegetation index (ARVI), Visible Atmospherically Resistance Index (VARI), Green Normalized Difference Vegetation Index (GNDVI)
Texture	Homogeneity, variance, mean, contrast, dissimilarity, entropy

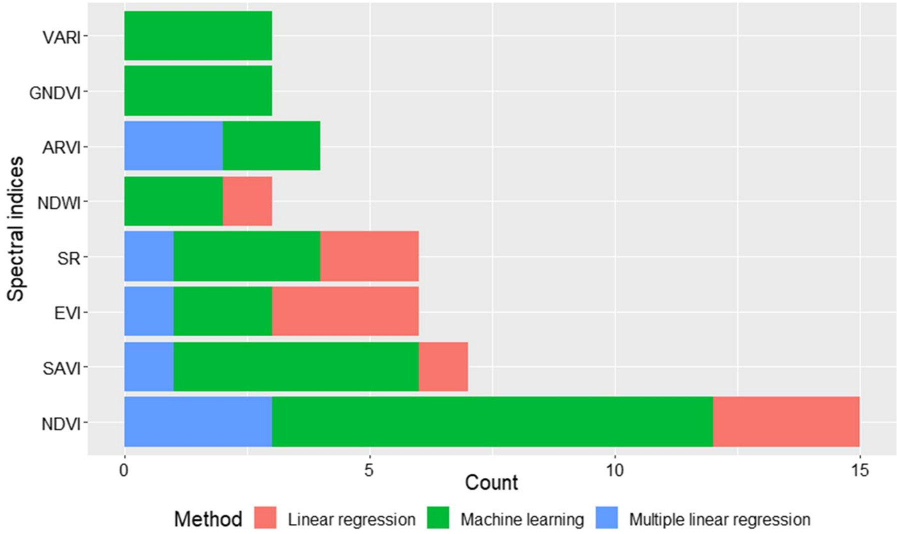

Fig. 4 Frequency of highly used spectral indices. Acronyms: *ARVI* Atmospherically Resistant Vegetation Index, *EVI* Enhanced Vegetation Index, *GNDVI* Green Normalized Difference Vegetation Index, *NDVI* Normalized Difference Vegetation Index, *NDWI* Normalized Difference Water Index, *SAVI* Soil Adjusted Vegetation Index, *SR* Simple Ratio, *VARI* Visible Atmospherically Resistance Index

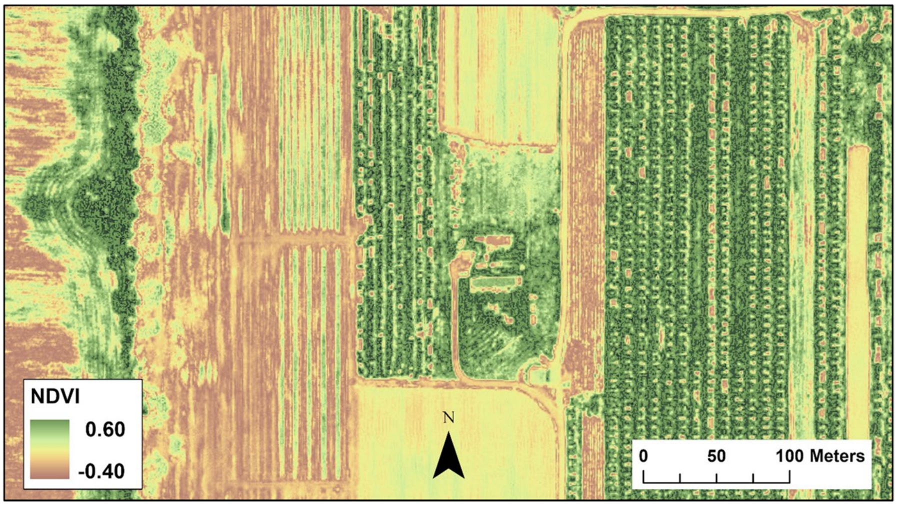

Fig. 5 NDVI image of a farm with orchards and vineyards in Central Indiana (Source: Authors, Imagery: USA National Agricultural Imagery Program, Spatial resolution: 1 m, Image Date: June 6, 2020)

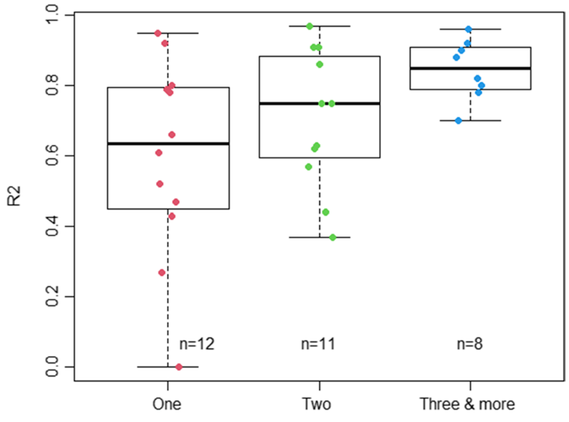

Fig. 6 R-squared by the count of covariates

emerging in AFS. The accuracy of ML depends on the input data, the number of explanatory variables used, data resolutions, and other parameters (Morais et al. 2021). Nineteen different algorithms were used in AGB estimation. Random Forest (RF), Support Vector Regression (SVR), and Stochastic Gradient Boosting (SGB) had slightly higher R-squared than other models (Fig. 7). RF is the most frequently used machine learning algorithm because it has been implemented in remote sensing literature since the early 1990s (Strahler and Jupp 1990) and it can capture the complexity and non-linearity of relationships with less sensitivity to noise in the training data (Mascaro et al. 2014; Safari et al. 2017). Other

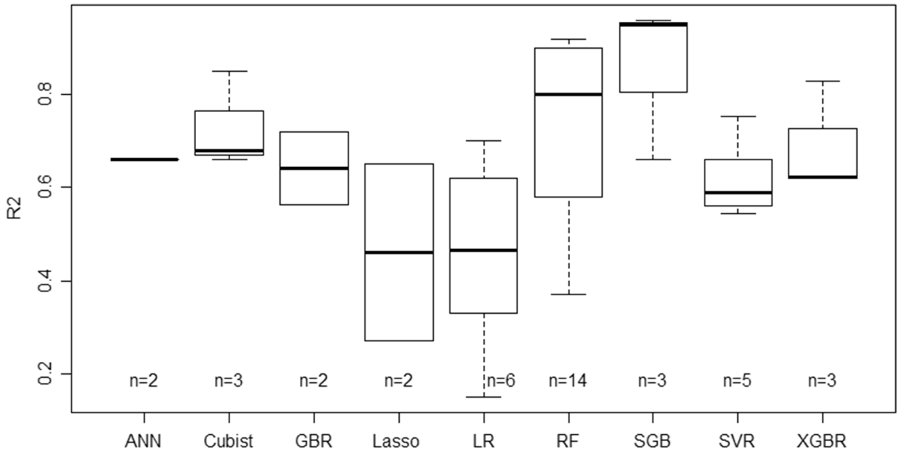

Fig. 7 The R-squared of frequently used machine learning algorithms for AGB estimation. Acronyms: *ANN* Artificial neural network, *GBR* Gradient boosting regression, *LR* Linear regression, *RF* Random forest, *SGB* Stochastic gradient boosting, *SVR* Support vector regression, *XGBR* Extreme gradient boosting regression

popular algorithms include Multivariate Adaptive Regression Splines (MARS), Stochastic Gradient Boosting (SGB), and Support Vector Regression (SVR). Filippi et al. (2014) estimated the AGB of the riparian forest using MARS, SGB, and Cubist and found SGB to be the most accurate predictor. Though there were no significant differences between MARS and SGB estimates, SGB was less sensitive to input variables than MARS and Cubist. Likewise, Safari et al. (2017) found that RF and MARS outperformed Support Vector Machine (SVM) and Boosted Regression Tree (BRT) for low biomass forests. RF also performed better than Artificial Neural Networks (ANN) that emulates human learning through interconnected processing units called nodes in a forest setting (Zhang et al. 2020).

However, some studies found other models to be better performing than RF. Pham et al. (2020) noted Extreme boosting regression (XGBR) to be performing slightly better than CatBoosting regression (CBR), Gradient boosted regression tree (GBRT), RF, and SVR in a riparian system (R-squared = 0.622, RMSE = 27.36 Mg/ha).

Besides these standard models, scholars have also developed hybrid algorithms for AGB estimation in forestry. Su et al. (2020) combined RF and spatial kriging to create a random forest regression kriging model that addresses the spatial autocorrelation effect not covered by RF alone. Pham et al. (2020) used the combination of XGBR and genetic algorithm to build an XGBR-GA model that outperformed CBR, GBRT, SVR, and RF regression for mangrove forests in Vietnam.

While machine learning models had higher R-squared in many studies, they did not always improve the AGB estimation. Almeida et al. (2019) used LiDAR and hyperspectral data to estimate the AGB of the Brazilian Amazon forest. The authors compared the performance of a linear model with Ridge Regularization, SVR, RF, SGB, and Cubist and found that LiDAR and hyperspectral images had a more substantial impact on AGB estimation than ML algorithms.

Opportunities and challenges for AGB estimation of agroforestry systems

Some of the unique opportunities and challenges for estimating the AGB of AFS are as follows:

Opportunities

1. *Advances in AGB estimation of forest and agriculture systems* There have been significant advances in AGB estimation of forests and agriculture (Lu et al. 2016). As a combination of agriculture and forestry, AFS can benefit from methodological advances in these sectors (Bégué et al. 2018; Ahmad et al. 2021). For example, riparian buffers and food forests are very similar to forestry systems, and therefore forest-related methodologies could be applied to these systems with some modifications. Similarly, the AGB estimation of alley cropping could benefit from multitemporal remote sensing applications in agriculture settings (Karlson et al. 2020).

2. *Advances in cloud computing* One of the main barriers to the wider adoption of remote sensing has been data and computation needs. Over the last decade, there have been significant advances in cloud computing platforms such as Google Earth Engine (GEE), Microsoft Azure, and Amazon Web Service. In addition, some of these services are tailored toward remote sensing analysis. For example, GEE, a platform released by Google, allows users to have access to multiple geospatial datasets and achieve parallel programming using GEE's in-built library (Kumar and Mutunga 2018; Amani et al. 2020). These platforms are increasingly used for biomass estimation (Yang et al. 2019; Xie et al. 2022), and they are likely to expand in the future.
3. *High-resolution data sources* There has been rapid advancement of spatial data acquisition methods that produce high spatial and temporal resolution datasets. High-resolution sensors like GeoEye, WorldView, IKONOS, EROS-B, Pleiades, and PlanetLab are already being used in AFS research. In addition, Unmanned Areal Systems (UAS) has also expanded data collection capacity at a very-high resolution (Pádua et al. 2017). The availability of these data products and widely accessible cloud computing platforms is expected to expand biomass research of AFS.

Challenges

1. *Identification of agroforestry features* Accurate identification of agroforestry over large areas would require very high-resolution imagery and specialized methods due to the small size and/or narrow width of AFS (Czerepowicz et al. 2012). These systems have been mapped with high accuracy for Trees Outside Forests, windbreaks, and riprain buffer (Liknes et al. 2017), however, narrow features and presence of shadows hinder accurate AFS delineation.
2. *Measurement of understory vegetation* The areal remote sensing is generally less accurate in estimating the understory canopy biomass. The insufficient LiDAR returns from heterogeneous vegetation cover create uncertainty in assessing the understory canopy AGB (Li et al. 2015). Advanced models like the radiative transfer model can help detect the contribution of understory crops towards measuring the vegetation cover, but they are applicable to a small areal extent (Hornero et al. 2021).
3. *Allometric equations* The trees grown in the open space of AFS accumulate more branch biomass than those grown in forest areas. Many allometric equations used for agroforestry AGB estimations are derived from the forest environment, potentially underestimating the biomass calculation (Zhou et al. 2011). In addition, since AGB differs by geography, agroforestry type, tree composition, tree age, and site quality, existing equations inadequately capture these variations (Chave et al. 2005).
4. *Biomass estimation error* The diverse tree density of AFS creates additional errors for the AGB estimation. Algorithms underestimate AGB in a forest with dense canopy cover due to saturation of pixels but over-estimate it in the area with thin canopy cover due to sub-pixel heterogeneity. Background soil reflectance can be problematic to estimate AGB in forested areas with low tree density. To some extent, the issue could be minimized by using high-resolution data and multiple parameters for feature extraction. In addition, the accuracy of remote sensing relies on the *in-situ* AGB measurements of sample plots for model calibration and validation (Wang et al. 2019).
5. *Access to geospatial data and technology* The cost and access to satellite imagery and technology are still major barriers for its wider application (Smith and Doldirina 2008). While many medium- and coarse-resolution imageries are freely available, high-resolution imageries are costly. Similarly, accessibility of satellite imageries is another hurdle (Turner et al. 2015). While data providers have developed specific websites and tools to facilitate easy access (e.g., https://sentinel.esa.int/), the limited information and technical barriers restrict their use, especially in developing countries where AGB is widely practiced.

Conclusion

This study reviewed methods of AGB estimation of AFS using remote sensing. A total of 33 papers were reviewed in detail covering all types of AFS across the world. Since remote sensing can assess a large area of land in a quick and efficient manner, it has been widely implemented in AGB research. High-resolution imageries are increasingly used for detecting heterogenous AFS through pixel-and object-based classification methods. For AGB estimation, scholars used regression, machine learning and statistical models. However, the study finds no statistically significant differences between the models in terms of performances (R-squared). Scholars have incorporated different types of covariates, including spectral bands, spectral indices, texture, biophysical and geomorphometric variables. Similarly, there's no statistically significant differences in model performance with the addition of covariate types. While the performance of ML varies by the input parameters, spectral and spatial resolutions, and type of sensors, non-linear algorithms such as Random Forest and Support Vector Machine are some of the most widely used algorithms. The advancements in cloud computing like Google Earth Engine and the availability of high-resolution dataset presents opportunities for wider use of remote sensing in AFS biomass estimation.

Remote sensing and machine learning technologies will become increasingly valuable in the coming years, as working lands are tapped as spaces to store carbon. Compared with other types of agricultural landscapes, AFS is uniquely suited as habitats for applying for carbon credits, because much of the sequestered carbon is detectable in the aboveground biomass. Unlike carbon stored the in soil, the aboveground pool can be verified, estimated, and monitored without extensive on-the-ground sampling. The ability to apply carbon credits to site-scale plantings, with direct financial benefits for individual landowners, would be an important breakthrough for distributing carbon storage broadly across the landscape. The technologies explored in this paper offer a foundation for moving that dialog forward.

Acknowledgements This material is based upon work supported by the University of Missouri Center for Agroforestry and the U.S. Department of Agriculture, Agricultural Research Service, under agreement No. 58-6020-0-007. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture.

Funding The authors have no financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

References

Ahmad A, Gilani H, Ahmad SR (2021) Forest aboveground biomass estimation and mapping through high-resolution optical satellite imagery—a literature review. Forests 12:914. https://doi.org/10.3390/f12070914

Amani M, Ghorbanian A, Ahmadi SA et al (2020) Google earth engine cloud computing platform for remote sensing big data applications: a comprehensive review. IEEE J Sel Top Appl Earth Observations Remote Sens 13:5326–5350. https://doi.org/10.1109/JSTARS.2020.3021052

Bégué A, Arvor D, Bellon B et al (2018) Remote Sensing and cropping practices: a review. Remote Sens 10:99. https://doi.org/10.3390/rs10010099

Bolívar-Santamaría S, Reu B (2021) Detection and characterization of agroforestry systems in the Colombian Andes using sentinel-2 imagery. Agrofor Syst 95:499–514. https://doi.org/10.1007/s10457-021-00597-8

Bordoloi R, Das B, Tripathi OP et al (2022) Satellite-based integrated approaches to modeling spatial carbon stock and carbon sequestration potential of different land uses of Northeast India. Environ Sustain Indic 13:100166. https://doi.org/10.1016/j.indic.2021.100166

Chave J, Andalo C, Brown S et al (2005) Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145:87–99. https://doi.org/10.1007/s00442-005-0100-x

Chen Q, Lu D, Keller M et al (2015) Modeling and mapping agroforestry aboveground biomass in the Brazilian amazon using airborne lidar data. Remote Sens 8:21. https://doi.org/10.3390/rs8010021

Chen L, Ren C, Zhang B et al (2018) Estimation of forest above-ground biomass by geographically weighted regression and machine learning with sentinel imagery. Forests 9:582. https://doi.org/10.3390/f9100582

Czerepowicz L, Case BS, Doscher C (2012) Using satellite image data to estimate aboveground shelterbelt carbon stocks across an agricultural landscape. Agr Ecosyst Environ 156:142–150. https://doi.org/10.1016/j.agee.2012.05.014

de Almeida CT, Galvão LS, de Aragão LE, OC e, et al (2019) Combining LiDAR and hyperspectral data for aboveground biomass modeling in the Brazilian Amazon using different regression algorithms. Remote Sens Environ 232:111323. https://doi.org/10.1016/j.rse.2019.111323

Filippi AM, Güneralp İ, Randall J (2014) Hyperspectral remote sensing of aboveground biomass on a river meander bend using multivariate adaptive regression splines and stochastic gradient boosting. Remote Sens Lett 5:432–441. https://doi.org/10.1080/2150704X.2014.915070

Forkuor G, Benewinde Zoungrana J-B, Dimobe K et al (2020) Above-ground biomass mapping in West African dryland forest using Sentinel-1 and 2 datasets - A case study. Remote Sens Environ 236:111496. https://doi.org/10.1016/j.rse.2019.111496

Gatziolis D, Andersen H-Erik (2008) A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest. In: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR
Garrett HG, Buck L (1997) Agroforestry practice and policy in the United States of America. For Ecol Manage 91:5–15. https://doi.org/10.1016/S0378-1127(96)03884-4
Gonçalves AC, Sousa AMO, Mesquita P (2019) Functions for aboveground biomass estimation derived from satellite images data in Mediterranean agroforestry systems. Agrofor Syst 93:1485–1500. https://doi.org/10.1007/s10457-018-0252-4
Graves SJ, Caughlin TT, Asner GP, Bohlman SA (2018) A tree-based approach to biomass estimation from remote sensing data in a tropical agricultural landscape. Remote Sens Environ 218:32–43. https://doi.org/10.1016/j.rse.2018.09.009
Griscom BW, Adams J, Ellis PW et al (2017) Natural climate solutions. Proc Natl Acad Sci USA 114:11645–11650. https://doi.org/10.1073/pnas.1710465114
Güneralp İ, Filippi AM, Randall J (2014) Estimation of floodplain aboveground biomass using multispectral remote sensing and nonparametric modeling. Int J Appl Earth Obs Geoinf 33:119–126. https://doi.org/10.1016/j.jag.2014.05.004
Hornero A, North PRJ, Zarco-Tejada PJ et al (2021) Assessing the contribution of understory sun-induced chlorophyll fluorescence through 3-D radiative transfer modeling and field data. Remote Sens Environ 253:112195. https://doi.org/10.1016/j.rse.2020.112195
Houghton RA, Unruh JD, Lefebvre PA (1993) Current land cover in the tropics and its potential for sequestering carbon. Global Biogeochem Cycles 7:305–320. https://doi.org/10.1029/93GB00470
Kalita RM, Das AK, Nath AJ (2016) Carbon stock and sequestration potential in biomass of tea agroforestry system in Barak Valley, Assam, North East India. Int J Ecol Environ Sci 42:107–114
Kanmegne Tamga D, Latifi H, Ullmann T et al (2022) Estimation of aboveground biomass in agroforestry systems over three climatic regions in West Africa using sentinel-1, sentinel-2, ALOS, and GEDI data. Sensors 23:349. https://doi.org/10.3390/s23010349
Karlson M, Ostwald M, Reese H et al (2015) Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using Landsat 8 and random forest. Remote Sens 7:10017–10041. https://doi.org/10.3390/rs70810017
Karlson MO, Madelene; Reese, Heather; Bazié, HR; Tankoano, Boalidioa, (2016) Assessing the potential of multiseasonal WorldView-2 imagery for mapping West African agroforestry tree species. Int J Appl Earth Obs Geoinformation 50:80–88. https://doi.org/10.1016/j.jag.2016.03.004
Karlson M, Ostwald M, Bayala J et al (2020) The potential of sentinel-2 for crop production estimation in a smallholder agroforestry landscape burkina faso. Front Environ Sci 8:85. https://doi.org/10.3389/fenvs.2020.00085
Kearney SP, Coops NC, Chan KMA et al (2017) Predicting carbon benefits from climate-smart agriculture: Highresolution carbon mapping and uncertainty assessment in El Salvador. J Environ Manage 202:287–298. https://doi.org/10.1016/j.jenvman.2017.07.039
Ku N-W, Popescu SC (2019) A comparison of multiple methods for mapping local-scale mesquite tree aboveground biomass with remotely sensed data. Biomass Bioenergy 122:270–279. https://doi.org/10.1016/j.biombioe.2019.01.045
Kumar L, Mutunga O (2018) Google earth engine applications since inception: usage, trends, and potential. Remote sensing 10:1509. https://www.mdpi.com/2072-4292/10/10/1509
Laosuwan T, Uttaruk Y (2016) Estimating above ground carbon capture using remote sensing technology in small scale agroforestry areas. AgricultForest 62:253. https://doi.org/10.17707/AgricultForest.62.2.22
Li A, Glenn NF, Olsoy PJ et al (2015) Aboveground biomass estimates of sagebrush using terrestrial and airborne LiDAR data in a dryland ecosystem. Agric for Meteorol 213:138–147. https://doi.org/10.1016/j.agrformet.2015.06.005
Liang Y, Kou W, Lai H et al (2022) Improved estimation of aboveground biomass in rubber plantations by fusing spectral and textural information from UAV-based RGB imagery. Ecol Indic 142:109286. https://doi.org/10.1016/j.ecolind.2022.109286
Liknes GC, Meneguzzo DM, Kellerman TA (2017) Shape indexes for semi-automated detection of windbreaks in thematic tree cover maps from the central United States. Int J Appl Earth Obs Geoinf 59:167–174. https://doi.org/10.1016/j.jag.2017.03.005
Liu D, Xia F (2010) Assessing object-based classification: advantages and limitations. Remote Sens Lett 1:187–194. https://doi.org/10.1080/01431161003743173
Lourenço P, Godinho S, Sousa A, Gonçalves AC (2021) Estimating tree aboveground biomass using multispectral satellite-based data in Mediterranean agroforestry system using random forest algorithm. Remote Sens Appl: Soc Environ 23:100560. https://doi.org/10.1016/j.rsase.2021.100560
Lu H (2003) Decomposition of vegetation cover into woody and herbaceous components using AVHRR NDVI time series. Remote Sens Environ 86:1–18. https://doi.org/10.1016/S0034-4257(03)00054-3
Lu D, Chen Q, Wang G et al (2016) A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. Int J Digit Earth 9:63–105. https://doi.org/10.1080/17538947.2014.990526
Luo S, Wang C, Xi X et al (2017) Fusion of airborne LiDAR data and hyperspectral imagery for aboveground and belowground forest biomass estimation. Ecol Ind 73:378–387. https://doi.org/10.1016/j.ecolind.2016.10.001
Macedo FL, Sousa AMO, Gonçalves AC et al (2018) Aboveground biomass estimation for Quercus rotundifolia using vegetation indices derived from high spatial resolution satellite images. Eur J Remote Sens 51:932–944. https://doi.org/10.1080/22797254.2018.1521250
Mascaro J, Asner GP, Knapp DE et al (2014) A tale of two “Forests”: random forest machine learning aids tropical forest carbon mapping. PLoS ONE 9:e85993. https://doi.org/10.1371/journal.pone.0085993
Matese A, Berton A, Chiarello V et al (2021) Determination of riparian vegetation biomass from an unmanned aerial

Vehicle (UAV). Forests 12:1566. https://doi.org/10.3390/f12111566

Maxwell AE, Warner TA, Fang F (2018) Implementation of machine-learning classification in remote sensing: an applied review. Int J Remote Sens 39:2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Morais TG, Teixeira RFM, Figueiredo M, Domingos T (2021) The use of machine learning methods to estimate aboveground biomass of grasslands: a review. Ecol Ind 130:108081. https://doi.org/10.1016/j.ecolind.2021.108081

Naidoo L, Mathieu R, Main R et al (2016) L-band synthetic aperture radar imagery performs better than optical datasets at retrieving woody fractional cover in deciduous, dry savannahs. Int J Appl Earth Obs Geoinf 52:54–64. https://doi.org/10.1016/j.jag.2016.05.006

Nair PKR, Nair VD, Mohan Kumar B, Showalter JM (2010) Carbon sequestration in agroforestry systems. In: Advances in agronomy. Elsevier, pp 237–307

Pádua L, Vanko J, Hruška J et al (2017) UAS, sensors, and data processing in agroforestry: a review towards practical applications. Int J Remote Sens 38:2349–2391. https://doi.org/10.1080/01431161.2017.1297548

Pandey PC, Anand A, Srivastava PK (2019) Spatial distribution of mangrove forest species and biomass assessment using field inventory and earth observation hyperspectral data. Biodivers Conserv 28:2143–2162. https://doi.org/10.1007/s10531-019-01698-8

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

Prasondita E, Nakagoshi N, Suwandana E (2019) Ecological study of aboveground biomass and plant species diversity in complex agroforestry sites, Lampung, Indonesia. The 7th Annual Meeting 27

Rizvi RH, Newaj R, Prasad R, et al (2016) Assessment of carbon storage potential and area under agroforestry systems in Gujarat Plains by Co2 fix model and remote sensing techniques. Curr Sci 110:2005. https://doi.org/10.18520/cs/v110/i10/2005-2011

Safari A, Sohrabi H, Powell S, Shataee S (2017) A comparative assessment of multi-temporal Landsat 8 and machine learning algorithms for estimating aboveground carbon stock in coppice oak forests. Int J Remote Sens 38:6407–6432. https://doi.org/10.1080/01431161.2017.1356488

Sankar S, Lewis M, Hosein P (2022) Above ground biomass estimation of a cocoa plantation using machine learning. 2022 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE, Orlando, FL, USA, pp 1–8

Schneider LC, Lerner AM, McGroddy M, Rudel T (2018) Assessing carbon sequestration of silvopastoral tropical landscapes using optical remote sensing and field measurements. J Land Use Sci 13:455–472. https://doi.org/10.1080/1747423X.2018.1542463

Singh M, Malhi Y, Bhagwat S (2014) Evaluating land use and aboveground biomass dynamics in an oil palm–dominated landscape in Borneo using optical remote sensing. J Appl Remote Sens 8:083695.https://doi.org/10.1117/1.JRS.8.083695

Smith LJ, Doldirina C (2008) Remote sensing: a case for moving space data towards the public good. Space Policy 24:22–32. https://doi.org/10.1016/j.spacepol.2007.12.002

Strahler AH, Jupp DLB (1990) Modeling bidirectional reflectance of forests and woodlands using boolean models and geometric optics. Remote Sens Environ 34:153–166. https://doi.org/10.1016/0034-4257(90)90065-T

Su H, Shen W, Wang J et al (2020) Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests. Forest Ecosystems 7:64. https://doi.org/10.1186/s40663-020-00276-7

Suchenwirth L, Stümer W, Schmidt T et al (2014) Large-scale mapping of carbon stocks in riparian forests with self-organizing maps and the k-nearest-neighbor algorithm. Forests 5:1635–1652. https://doi.org/10.3390/f5071635

Tian Y, Huang H, Zhou G et al (2021) Aboveground mangrove biomass estimation in Beibu Gulf using machine learning and UAV remote sensing. Sci Total Environ 781:146816. https://doi.org/10.1016/j.scitotenv.2021.146816

Timothy D, Onisimo M, Cletah S et al (2016) Remote sensing of aboveground forest biomass: a review. Trop Ecol 57:125–132

Turner W, Rondinini C, Pettorelli N et al (2015) Free and open-access satellite data are key to biodiversity conservation. Biol Cons 182:173–176. https://doi.org/10.1016/j.biocon.2014.11.048

Udawatta RP, Jose S (2012) Agroforestry strategies to sequester carbon in temperate North America. Agrofor Syst 86:225–242. https://doi.org/10.1605/01.301-0020846221.2012

Unruh J, Houghton R, Lefebvre P (1993) Carbon storage in agroforestry: an estimate for sub-Saharan Africa. Clim Res 3:39–52. https://doi.org/10.3354/cr003039

Wang Z, Liu L, Peng D et al (2016) Estimating woody aboveground biomass in an area of agroforestry using airborne light detection and ranging and compact airborne spectrographic imager hyperspectral data: individual tree analysis incorporating tree species information. J Appl Remote Sens 10:036007. https://doi.org/10.1117/1.JRS.10.036007

Wang Y, Pyörälä J, Liang X et al (2019) In situ biomass estimation at tree and plot levels: what did data record and what did algorithms derive from terrestrial and aerial point clouds in boreal forest. Remote Sens Environ 232:111309. https://doi.org/10.1016/j.rse.2019.111309

Wilson M, Lovell S (2016) Agroforestry—the next step in sustainable and resilient agriculture. Sustainability 8:574. https://doi.org/10.3390/su8060574

Wu C, Shen H, Shen A et al (2016) Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery. J Appl Remote Sens 10:035010. https://doi.org/10.1117/1.JRS.10.035010

Xie J, Wang C, Ma D et al (2022) Generating spatiotemporally continuous grassland aboveground biomass on the tibetan plateau through PROSAIL model inversion on google earth engine. IEEE Trans Geosci Remote Sens 60:1–10. https://doi.org/10.1109/TGRS.2022.3227565

Yang Z, Li W, Chen Q et al (2019) A scalable cyberinfrastructure and cloud computing platform for forest aboveground biomass estimation based on the google earth engine. Null

12:995–1012. https://doi.org/10.1080/17538947.2018.1494761

Yasen K, Koedsin W (2015) Estimating aboveground biomass of rubber tree using remote sensing in Phuket Province, Thailand. J Med Bioeng 4:451–456

Zhang Y, Ma J, Liang S et al (2020) An evaluation of eight machine learning regression algorithms for forest aboveground biomass estimation from multiple satellite data products. Remote Sensing 12:4015. https://doi.org/10.3390/rs12244015

Zhou X, Brandle JR, Awada TN et al (2011) The use of forest-derived specific gravity for the conversion of volume to biomass for open-grown trees on agricultural land. Biomass Bioenerg 35:1721–1731. https://doi.org/10.1016/j.biombioe.2011.01.019

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.