Revista ÁRVORE
ISSN 1806 - 9088

Scientific Article

ASSESSING MULTIPLE REMOTELY SENSED DATA AND MODEL-ASSISTED INFERENCE FOR BIOMASS ESTIMATION OF AN ATLANTIC FOREST FRAGMENT

Talles Bhering de Matos[2*], Amanda Camargos de Moura[2], Mikaely Vasconcelos Paulo[2], Vivian Silva Santos[2], Luiana Rolim de Azevedo[2], Alexandre Simões Lorenzon[3], Carlos Moreira Miquelino Eleto Torres[3], Elpidio Inácio Fernandes Filho[4], Ernani Lopes Possato[3], José Marinaldo Gleriani[3] and Diogo Nepomuceno Cosenza[3]

1 Received on 12.07.2025 accepted for publication on 16.09.2025. Editors: Bruno Leão Said Schettini.
2 Universidade Federal de Viçosa, Programa de Pós-Graduação em Ciência Florestal, Viçosa, Minas Gerais, Brasil. E-mail: <talles.matos@ufv.br>, <amanda.moura@ufv.br>, <mikaely.paulo@ufv.br>, <vivian.s.santos@ufv.br> and <luiana.azevedo@ufv.br>.
3 Universidade Federal de Viçosa, Departamento de Engenharia Florestal, Viçosa, Minas Gerais, Brasil. E-mail: <alexandre.lorenzon@ufv.br>, <carlos.eleto@ufv.br>, <ernani.possato@ufv.br>, <gleriani@ufv.br> and <diogo.cosenza@ufv.br>.
4 Universidade Federal de Viçosa, Departamento de Solos, Viçosa, Minas Gerais, Brasil. E-mail: <elpidio@ufv.br>.
*Corresponding author.

ABSTRACT

Forest biomass quantification (Mg ha^{-1}) is essential for ecosystem monitoring, especially in areas under anthropogenic pressure, such as Atlantic Forest fragments. This study aimed to compare remote sensors in biomass mapping and population stock estimation of an Atlantic Forest fragment. Ten 0.1 ha plots were randomly distributed within a 17 ha fragment. Data from Sentinel-1 (S1), Sentinel-2 (S2), digital aerial photogrammetry (DAP), and their fusion were evaluated for the construction of predictive models. Linear models with two predictors were fitted: one for each sensor and another using data fusion, selecting the best predictors among all. The models were applied to estimate stand biomass using a regression estimator. The data fusion model showed the best predictive performance (RMSE = 41%), while the DAP-based model had the highest error (RMSE = 64%). However, the most accurate population estimate was obtained with the S2-based model (SE = 21 Mg ha^{-1}), with a relative efficiency 7% higher compared to the traditional inventory (SE = 22 Mg ha^{-1}). Estimates based on DAP, S1, and fusion were less accurate than those from the field inventory. The selected metrics, such as vegetation indices (S2) and textural metrics (S1), reflected the sensors' sensitivity to canopy structure and foliage abundance. DAP showed limitations, possibly due to its low canopy penetration. It is concluded that although the data fusion between DAP and S2 produced the best model for biomass mapping, S2 alone proved more advantageous for population estimates in forest fragments with limited sampling

Keywords: Forest inventory; Forest mensuration; Forest mapping

How to cite:
Matos, T. B. de, Moura, A. C. de, Paulo, M. V., Santos, V. S., Azevedo, L. R. de, Lorenzon, A. S., Torres, C. M. M. E., Fernandes Filho, E. I., Possato, E., Gleriani, J. M., & Cosenza, D. N. (2025). Assessing multiple remotely sensed data and model-assisted inference for biomass estimation of an Atlantic Forest fragment. *Revista Árvore, 50*(1). https://doi.org/10.53661/1806-9088202650263981

Revista ÁRVORE
ISSN 1806 - 9088

USO DE SENSORIAMENTO REMOTO DE MÚLTIPLAS FONTES E ESTIMADOR DE REGRESSÃO PARA QUANTIFICAR O ESTOQUE DE BIOMASSA DE UM FRAGMENTO DE MATA ATLÂNTICA

RESUMO A quantificação da biomassa florestal (Mg ha^{-1}) é fundamental para o monitoramento de ecossistemas, especialmente em áreas sob pressão antrópica, como os fragmentos da Mata Atlântica. Este estudo teve como objetivo comparar sensores remotos no mapeamento da biomassa de um fragmento de Mata Atlântica e na estimativa do estoque populacional. Dez parcelas de 0,1 ha foram distribuídas aleatoriamente em um fragmento de 17 ha. Foram avaliados dados do Sentinel-1 (S1), Sentinel-2 (S2), fotogrametria aérea digital (DAP) e a fusão entre eles para a construção de modelos preditivos. Modelos lineares com dois preditores foram ajustados: um para cada sensor e outro com fusão de dados, selecionando os melhores preditores entre todos. Os modelos foram aplicados para estimar a biomassa do povoamento usando estimador de regressão. O modelo baseado em fusão de dados apresentou o melhor desempenho preditivo (RMSE = 41%), enquanto o baseado em DAP teve o maior erro (RMSE = 64%). Contudo, a estimativa populacional mais precisa foi obtida com o modelo baseado no S2 (SE = 21 Mg ha^{-1}), com eficiência relativa de 7% comparado com o inventário tradicional (SE = 22 Mg ha^{-1}). As estimativas baseadas em DAP, S1 e fusão foram menos precisas que a do inventário de campo. As métricas selecionadas, como os índices de vegetação (S2) e as métricas texturais (S1), refletiram a sensibilidade dos sensores à estrutura do dossel e à abundância foliar. A DAP apresentou limitações, possivelmente por sua baixa penetração no dossel. Conclui-se que, embora a fusão de dados entre DAP e S2 ter gerado o melhor modelo para mapeamento de biomassa, o S2 isoladamente foi mais vantajoso para estimativas populacionais em fragmentos florestais com amostragem limitada.

Palavras-Chave: Inventário florestal; Mensuração florestal; Caracterização florestal

1. INTRODUCTION

Tropical regions contain approximately 45% of the world's forest cover (FAO, 2020). Among these, the Atlantic Forest stands out from other tropical forests due to its high diversity of endemic species (Souza et al., 2021) and its location in a densely populated region, 58% of the Brazil's inhabitants (IBGE, 2024). Despite its ecological importance, the Atlantic Forest is under severe threat, with only 26% of its original vegetation remaining (MapBiomas, 2024), with 83% composed of small forest fragments (< 50 ha) scattered across the country (Ribeiro et al., 2011). Because of this, the Atlantic Forest is recognized as a critical biodiversity hotspot, that requires monitoring to design and assess effective conservation strategies (Myers et al., 2000).

Growing stock biomass ("biomass") is a forest attribute closely related to carbon stocks and the forest's ecological complexity (Brown & Lugo, 1984). The biomass increase is frequently used as indicators of biodiversity gains, regeneration progress, and improvements in ecosystem quality (Houghton et al., 2009). Biomass is also a fundamental variable for estimating carbon sequestration (Ma et al., 2024). Traditionally, biomass assessments rely on sample-based forest inventories. In this process, the biomass is measured using sample units (field plots) that are spread over the area following some sampling design. Then the population parameters are estimated along with a confidence interval where the true population value lies (Soares, 2017). Because the uncertainty of the estimation is due to the sampling design (the values of the population units are assumed to be fixed), this framework is referred to as design-based inference. However, the traditional measurement approach is costly and logistically complex as it requires field expeditions and access to remote areas (Chukwu & Dau, 2020). As an alternative, remote sensing technologies can be used to enhance the design-based inference. Metrics

derived from remotely sensed data might be tidily correlated to forest structural attributes so they can be used both for attribute mapping and population parameter estimation (McRoberts et al., 2013a). While the forest mapping allows for precision monitoring, knowing the population parameter is important for stand-level assessments.

Forest inventory following a model-assisted framework is a strategy that use the information of auxiliary data to reduce the uncertainty of design-based estimation. First, a model linking the remotely sensed data to the forest attribute is built to predict values to the population units, and then the population parameter is estimated by aggregating the many predictions (the synthetic estimator) plus a bias correction term based on the field plot samples (Gregoire et al., 2016). The estimator is referred to as the Generalized REGression estimator (GREG) when the data used to calibrate the model and compute the correction term originates from the same target population (Särndal et al., 1992, Chapter 6). Such a framework is advantageous because it is asymptotically unbiased and requires only that the sample units follow a consistent probability design (e.g. simple random sample), so there is no constraint regarding the model consistency (McRoberts et al., 2022).

The GREG estimator has been tested with remotely sensed data worldwide (e.g., Saarela et al., 2015; Ståhl et al., 2016; McRoberts et al., 2022). In the Atlantic Forest, however, few studies have explored model-assisted estimation. Existing research has used satellite images and State or National Forest Inventory plots to estimate forest cover (Vibrans et al., 2013; McRoberts et al., 2016) or biomass in large and small areas (David et al., 2024). In some cases, these studies addressed small areas, a statistical concept of a subunit (also known as a domain) of a larger population where there are few or no sample units (Rao & Molina, 2015). This requires specific estimators, the small area estimators, as the ones used in David et al. (2024). Nevertheless, there are also cases where forest inventories would target a specific forest fragment, so even if it has a small coverage area (< 50 ha) it is considered the population itself and not a domain of a larger area. The performance of the GREG estimator in such a case is still unclear.

Forest biomass monitoring using remote sensing can be conducted using different sensors. Active systems based on light detection and ranging (LiDAR) and aerial platforms (airborne laser scanning, ALS) (Maltamo et al., 2014) are considered the most efficient for forest attribute mapping (Fassnacht et al., 2014; Bergseng et al., 2015). The cost of ALS data collection is high, especially for assessments involving small forest areas. In this case, there is the alternative of using freely available data sources from orbital systems or digital aerial photogrammetry (DAP). Among free-access orbital data, the Sentinel-1 (S1) and Sentinel-2 (S2) systems have gained popularity for forest monitoring. The S1 is a synthetic aperture radar (SAR) system (see Moreira et al., 2013) that actively emits C-band microwave energy (5.5 cm). It has the capability of penetrating the forest canopy to capture information about the forest structure with a resolution of 5-20 m. The S2 system is a passive optical sensor collecting data of 13 spectral bands to produce images of 10-60 m pixel resolution. The DAP systems aboard unmanned aerial vehicles (UAV) usually collect data from a few spectral bands, but with the advantage of producing stereo images of centimetric resolutions. This feature is particularly useful for precision assessments in forest fragments with a reduced cost compared to ALS data.

The integration of multiple remote sensing data sources can further improve biomass estimation models. Several studies have shown the integration of SAR and optical data derive more accurate biomass estimates than a single data source (Ghosh & Behera, 2018; Navarro et al., 2019; Zhang et al., 2023; Moghimi et al., 2024; Wang et al., 2024). The lack of field data availability makes it difficult to obtain accurate estimates of biomass and other forest attributes in heterogeneous conditions, such as tropical forests (Verly et al., 2023). It is therefore important to minimize the errors associated with sensor-based estimates, though we are not aware of studies assessing remote sensing data fusion for biomass estimation in the Atlantic Forest. We hypothesize that fusing

multiple data sources might contribute to improve accuracy of biomass maps and estimates.

The aim of this study is to assess the performance of multiple remote sensing data sources in the mapping and estimation of growing stock biomass of an Atlantic Forest fragment. We compare models based on single data sources of S1, S2, and DAP systems, as well as their fusion. In addition, we assessed the impact of the multiple data in the fragment biomass estimation.

2. MATERIAL AND METHODS

2.1 Study Area

The study was conducted in a forest fragment of approximately 17 ha, located in Viçosa, Minas Gerais, Brazil. The area belongs to the Federal University of Viçosa (UFV) (Figure 1). According to the Köppen classification, the climate is of the Cwa type, with an average temperature of 19.9 °C, relative humidity of 79.9%, and 1,269 millimeters (mm) of precipitation (UFV, 2024). The vegetation is classified as Semideciduous Seasonal Forest (IBGE, 2012). The fragment is a secondary forest in an intermediate regeneration stage, dominated by tree species of 5-12 m height and diameters ranging from 10 to 20 cm.

2.2 Field data collection and description

Field reference data were collected from ten permanent plots of 20 m x 50 m (0.1 ha) between July 5 and August 1, 2024. These plots were established following a simple random sampling (SRS) design in the 90s for the purpose of continuous forest monitoring every four years (Mariscal Flores, 1993; Meira Neto, 1997). Each plot corner had its position recorded using a Javad Triumph-1 GNSS receiver which provides sub-meter precision. We collected the tree circumference at breast height (cbh, cm) with

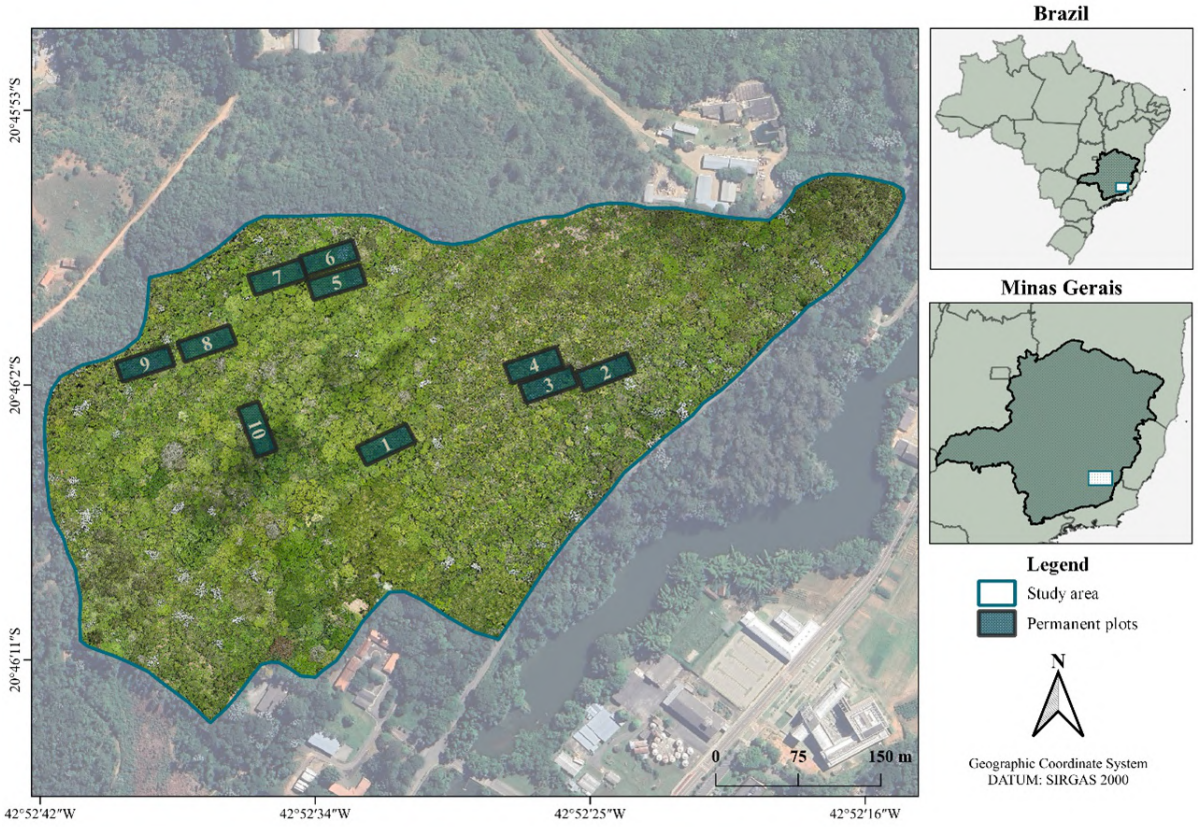

Figure 1. Forest fragment and plot distribution. The background image is from Google Earth, the orthomosaic was generated from the DAP data, and the administrative boundaries are from IBGE (2020)
Figura 1. Fragmento florestal e distribuição das parcelas. A imagem de fundo é do Google Earth, o ortomosaico foi gerado a partir dos dados da DAP, e os limites administrativos são do IBGE (2020)

bark of all living and dead trees with cbh ≥ 15 cm using a measuring tape and later converted the values to diameter at breast height (dbh, cm). The relative position of each tree was also recorded, and total height (h) was measured using a Vertex IV digital hypsometer. The biomass of each tree (b, kg) was estimated using Equation 1 fitted for the same area by Amaro (2010). The values were converted to megagrams per hectare and aggregated at the plot level to obtain the growing stock biomass (B, Mg ha^{-1}). The plot biomass distribution is depicted in Table 1.

$$b = 0.02453\, h^{0.423602} \cdot dbh^{2.443356} \quad \text{(Eq. 1)}$$

2.3 Auxiliary data collection and preparation

2.3.1 DAP data

The DAP images were captured on November 19, 2024, using the M3M RGB camera integrated into a UAV DJI Mavic 3M (DJI, 2023). The camera has a 4/3 CMOS sensor of 20 megapixels that reduces motion blur and minimizes image distortion. The UAV is equipped with a real-time kinematic (RTK) module for differential correction. It utilizes data from a reference station, IBGE-RBMC (Brazilian Continuous Monitoring Network) stations, and GNSS/NTRIP technology (Costa et al., 2008), eliminating the need for ground control points. The flight was carried out under clear skies and moderate wind conditions (5 m/s), at an altitude of 120 m above ground, following a pre-established autonomous plan.

Image processing was performed in Agisoft Metashape Professional 2.1.3, following the software's default. Initial image alignment used structure-from-motion (SfM) algorithms to detect homologous points automatically. Camera position optimization was based on the onboard RTK navigation data. A dense point cloud was then generated increasing the number of points in possible empty spaces and capturing more details of the study area. The point cloud was then exported in LAS and LAZ formats for manipulation.

The entire point cloud processing was performed in R (R Core Team, 2024) using the lidR package (Roussel et al., 2020). Data pre-processing began with noise point removal and ground point filtering. For this procedure, we use the Cloth Simulation Filtering (CSF) algorithm in the default setting (Zhang et al., 2016). A 1-m-pixel digital terrain model (DTM) was generated by triangulation to convert the point altitude values to above-ground high values (i.e. point cloud normalization). From the normalized point cloud, canopy height and cover metrics were calculated for each plot using all points above 2 m from the ground to use them as potential predictor variables (Table 2).

2.3.2 Orbital data

S1 and S2 orbital data were obtained and processed using Google Earth Engine (GEE) platform within a 30-day interval of the DAP data acquisition. The number of available images was 2 for the S1 system (16 and 28 November 2024) and 1 for the S2 (28 November 2024). Therefore, for S1, we used the mean of the pixel values from the available images. We used the Level 1 S1 data, which corresponds to 10 m spatial resolution and dual polarization (vertical-vertical – VV, and vertical-horizontal – VH) (ESA, 2024). These images are preprocessed by the provider for thermal noise removal, terrain correction, and radiometric calibration. The calibration converts image values to backscatter coefficients (σ^0) in decibels (dB). Speckle filtering (Mullissa et al., 2021) and outlier detection were also applied to smooth out noise.

Table 1. Biomass distribution across sample plots (Mg ha^{-1})
Tabela 1. Distribuição da biomassa nas parcelas amostradas (Mg ha^{-1})

PLOT	BIOMASS (Mg ha^{-1})	PLOT	BIOMASS (Mg ha^{-1})
1	255.95	6	256.69
2	104.55	7	425.19
3	101.89	8	161.79
4	89.12	9	146.90
5	91.67	10	229.16

Table 2. Digital aerial photogrammetry metrics extracted from the normalized point cloud with type, name, and description
Tabela 2. Métricas da fotogrametria aérea digital extraídas da nuvem de pontos normalizada com tipo, nome e a descrição

TYPE	NAME	DESCRIPTION
Height	H_{mean}	Mean
	H_{med}	Median
	H_{min}	Minimum
	H_{max}	Maximum
	H_{var}	Variance
	H_{sd}	Standard deviation
	H_{cv}	Coefficient of variation
	H_{iqr}	Interquartile range
	H_{05-95}	5^{th} to 95^{th} percentiles
Canopy Cover	CCH	Number of points below 4 m
	CCH_{mean}	Number of points below H_{mean}
	CCH_{med}	Number of points below H_{med}
	CCH_{P}	% points below 4 m
	CCH_{meanP}	% points below H_{mean}
	CCH_{medP}	% points below H_{med}

S2 data corresponds to the harmonized Level 2A collection, with atmospherically corrected bands, detection of shadows, clouds, water cover, and adjustment of topographic distortions. These preprocessing steps correspond to the standard correction applied in the Sentinel collections available in GEE (see Google (2024b) for S1, and Google (2024a) for S2). All spectral bands were used except B1 (Aerosols) and B9 (Water vapor). The S2 data with 20 m spatial resolution were resampled to 10 m using the nearest neighbor method.

Pixel values from S1 and S2 were assigned to the field plots based on their geolocation using the SIRGAS 2000 UTM Zone 23S coordinate reference system. We used the plot polygons to compute the average of the intercepting pixel values weighted by the interception area within plots. The terra package (Hijmans, 2023) was used for data manipulation. We computed S1 metrics based on VV and VH bands, the relationships between them, and texture measures (Table 3) (see Fang et al., 2023). The texture measures were calculated using the glcm package (Zvoleff, 2020) with a 3 x 3 moving window and gray-level co-occurrence matrices. For S2, in addition to the spectral bands we calculated 13 spectral indices (Table 3).

2.4 Data analysis

Biomass models were developed using metrics derived individually from DAP, S1, and S2 datasets, as well as from the combined dataset (DAP + S1 + S2), here referred to as data fusion. Due to the limited number of observations available for model fitting, a conservative modeling approach was adopted to minimize the risk of overfitting (Cosenza et al., 2021b). In this case, the models were fitted using ordinary least squares (OLS) regression using two predictor variables. For each dataset, the candidate metrics were filtered by removing those with no variations among plots and grouping those highly correlated metrics (Pearson's $r > 0.85$ in magnitude). From each group we retained the metric having the highest correlation with the biomass.

Variable selection was performed using an exhaustive searching algorithm from the leaps package (Lumley, 2024). This algorithm compares all possible combinations of two predictors in the model and ranks them based on the coefficient of determination (R^2). We ensured the selected model met the assumptions of residual normality (Shapiro-Wilk test) and homoscedasticity (Breusch-Pagan test) at a 5% significance level, and the variance

Table 3. Bands and metrics extracted from Sentinel-1 and Sentinel-2 described by name and equation to perform the calculation
Tabela 3. Bandas e métricas extraídas do Sentinel-1 e Sentinel-2 descritas por nome e equação para realizar cálculo

SOURCE	TYPE	INDEX	EQUATION/DESCRIPTION	REFERENCES
Sentinel-1	Band	VV	Vertical-Vertical - 5.5 cm	-
		VH	Vertical-Horizontal - 5.5 cm	-
	Relationships built	PR	$\frac{VV}{VH}$	-
		PS	$VV + VH$	-
		PD	$VV - VH$	-
	Textural	$S1_{mean}$	$\frac{\sum_{k=1}^{N} x_k}{N}$	-
		$S1_{var}$	$\frac{\sum_{k=1}^{N}(x_k - \overline{x})}{N-1}$	-
		$S1_{con}$	$\sum_{n=0}^{N=1} n^2 \left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} p(i,j) \right\}$	-
		$S1_{ent}$	$-\sum_{i} \sum_{j} p(i,j) \log(p(i,j))$	-
		$S1_{hom}$	$\sum_{i=1}^{N} \sum_{j=1}^{N} \frac{p(i,j)}{1 + \lvert i-j \rvert}$	-
		$S1_{diss}$	$\sum_{i=1}^{N} \sum_{j=1}^{N} \lvert i-j \rvert \times p(i,j)$	-
		$S1_{sm}$	$\sum_{i=1}^{N} \sum_{j=1}^{N} [p(i,j)]^2$	-
		$S1_{cor}$	$\frac{\sum_{i}^{\sum_{j}^{(ij)}} p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$	-
Sentinel-2	Band	B2	Blue - 497 nm	-
		B3	Green - 560 nm	-
		B4	Red - 665 nm	-
		B5	RE 1 - 704 nm	-
		B6	RE 2 - 740 nm	-
		B7	RE 3 - 783 nm	-
		B8	NIR - 835 nm	-
		B8A	RE 4 - 865 nm	-
		B11	SWIR 1 - 1,614 nm	-
		B12	SWIR 2 - 2,202 nm	-
	Spectral indices	SRI	$\frac{B8}{B4}$	(Jordan, 1969)
		SRII	$\frac{B4}{B5}$	(Chappelle et al., 1992)
		SRIII	$\frac{B6}{B3}$	(A. Gitelson & Merzlyak, 1994)
		RVI	$\frac{B4}{B8}$	(Broge & Leblanc, 2001)
		DVI	$2{,}4 \times B8 - B4$	(Richardson & Wiegand, 1977)
		SAPI	$\frac{B8A}{(B3 \times B5)}$	(Datt, 1998)
		RECI	$\left(\frac{B6}{B5}\right) - 1$	(A. A. Gitelson et al., 2003)
		NDVI	$\frac{(B8 - B4)}{(B8 + B4)}$	(Rouse et al., 1974)
		GNDVI	$\frac{(B8 - B3)}{(B8 + B3)}$	(A. A. Gitelson & Merzlyak, 1998)

Cont...

Cont...

SOURCE	TYPE	INDEX	EQUATION/DESCRIPTION	REFERENCES
		NBRI	$\frac{(B8 - B12)}{(B8 + B12)}$	(García & Caselles, 1991)
		SAVI	$1{,}5 \times \frac{(B8 - B4)}{(B8 + B4 + 0{,}5)}$	(Huete, 1988)
		EVI	$2{,}5 \times \frac{(B8 - B4)}{(B8 + 6 \times B4 - 7{,}5 \times B2 + 1)}$	(H. Q. Liu & Huete, 2019)
		REP	$705 + 35 \times \left[\frac{\left(\frac{B7 + B4}{2}\right)}{B6 - B5}\right]$	(Guyot et al., 2021)

*Where: x_k= Gray tone per k pixel value; N = Number of gray-tone values; p = Marginal probability of normalized x; μ = p mean; σ = p standard deviation

inflation factor (VIF) was below 10 to avoid multicollinearity (Myers, 1990). All tests were computed using the performance package (Lüdecke et al., 2021). We used models with the log-transformed response (Equation 2) to avoid negative predictions in the maps. We added the model's residual variance (σ^2) when back-transforming the estimates to the original scale (Equation 3) as per Beauchamp & Olson (1973).

$$\ln(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \quad \text{(Eq. 2)}$$

$$\hat{y} = exp\left(\widehat{\ln(y)}\right) + exp\left(\frac{\sigma^2}{2}\right) \quad \text{(Eq. 3)}$$

Where: y and $\hat{y}$ are, respectively, the observed and estimated biomass value; β_1 and β_2 are the model's parameters; x_1 and x_2 are the predictors; ε is the random error; and σ^2 is the residual variance.

The selected models were compared using accuracy measures based on the square Pearson's correlation (r^2, Equation 4), root mean square error relative (RMSE%, Equation 5) and mean absolute error relative (MAE%, Equation 6) to the sample mean observed values computed using leave-one-out cross-validation (LOOCV). The LOOCV is an iterative validation process where one observation is omitted at a time to have its value predicted by the model calibrated using the remaining observations. All observations had their values predicted at the end so the accuracy measures (r^2, RMSE% and MAE%) can be computed. As an additional assessment, we compute a naïve root mean square error (nRMSE%, Equation 7) to evaluate the improvement provided by the predictive models.

$$r^2 = \frac{\left[\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})\right]^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2} \quad \text{(Eq. 4)}$$

$$RMSE\% = 100 \cdot \frac{1}{\bar{y}} \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad \text{(Eq. 5)}$$

$$MAE\% = 100 \cdot \frac{1}{\bar{y}} \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \quad \text{(Eq. 6)}$$

$$nRMSE\% = 100 \cdot \frac{1}{\bar{y}} \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n}} \quad \text{(Eq. 7)}$$

Where: y_i and $\hat{y}_i$ are, respectively, the observed and estimated biomass values (Mg ha^{-1}) for the plot i = 1, ..., n; $\bar{y}$ and $\bar{\hat{y}}$ are, respectively, the observed and estimated mean biomass of the sample.

2.5 Estimation

Although the forest fragments might be considered as a domain from large area forest inventories, in the present study the fragment was considered as the target population. Field samples plots were used to estimate the population mean, $\hat{\mu}$, and variance, $\widehat{Var}(\hat{\mu})$, so the confidence interval (CI) can be constructed.

The confidence intervals were estimated using Equation 8.

$$CI = \hat{\mu} \pm t_{(1-\alpha/2,\, df)} \sqrt{\widehat{Var}(\hat{\mu})} \quad \text{(Eq. 8)}$$

Where $t_{(1-\alpha/2, df)}$ is the critical value of the Student's t distribution at 95% confidence level with df degrees of freedom.

First, we construct intervals following simple random sampling (SRS) estimation, where Equation 9 and 10 were used to compute the population mean and variance (Cochran, 1977, Chapter 2.5).

$$\hat{\mu}_{SRS} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \text{(Eq. 9)}$$

$$\widehat{Var}(\hat{\mu}_{SRS}) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{\sum_{i=1}^{n}(y_i - \hat{\mu}_{SRS})}{n-1} \quad \text{(Eq. 10)}$$

Where: μ_{SRS} is the simple random sampling population mean; $\widehat{Var}(\hat{\mu}_{SRS})$ is the estimated variance associated with μ_{SRS}; y_i is the observed value of the sampling unit i; n is the number of sample units.

For model-assisted estimation, the biomass maps were generated with a spatial resolution of 31.62 m, equivalent to the square root of the sample plot area (1,000 m²). Later, the mean and variance were computed using Equation 11 and 12.

$$\hat{\mu}_{MA} = \frac{1}{N}\sum_{i=1}^{N} \hat{y}_i - \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \quad \text{(Eq. 11)}$$

$$\widehat{Var}(\hat{\mu}_{MA}) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n-1} \quad \text{(Eq. 12)}$$

Where: $\hat{\mu}_{MA}$ is the model-assisted estimator of the population mean; $\widehat{Var}(\hat{\mu}_{MA})$ is the variance estimator associated with $\hat{\mu}_{MA}$; $\hat{y}_i$ is the value estimated by the model for the unit of population i; y_i is the observed value of the response variable for sampling unit i; N is the population size, for example, the number of pixels with information from the sensors; n is the number of sample units.

Although we do not have the true biomass of the population, we compared the different estimates with the SRS design. We used the standard error, $SE=\sqrt{\widehat{Var}(\hat{\mu})}$, and the relative efficiency (RE, Equation 13) to assess the estimates. While the SE gives an indicator of the estimate uncertainty, the RE informs about the benefit of model-assisted estimates over SRS, where RE > 1 indicates increased effectiveness (McRoberts et al., 2013b).

$$RE = \frac{\widehat{Var}(\hat{\mu}_{SRS})}{\widehat{Var}(\hat{\mu}_{MA})} \quad \text{(Eq. 13)}$$

3. RESULTS

3.1 Model performance

The models differed considerably regarding prediction accuracy (Table 4 and Figure 2). The best model performance was achieved using the data fusion, followed by the S2, S1, and DAP datasets. This rank is reflected in the accuracy metrics. RMSE increased from 41% to 64%, while correlations (r^2) decreased from 0.49 to 0.15 and MAE increased from 27% to 48% from the best to the worst model. The naïve accuracy (nRMSE%) was 54%, and it was higher than all model accuracies, excepting the one based on DAP data, which indicates superior model performances compared to the simple mean model. The S1 based model selected textural metrics ($S1_{var}$ and $S1_{cor}$), whereas the S2 based model relied on vegetation indices derived from red and NIR bands (DVI), and red-edge bands (RECI). The DAP model incorporated metrics related to the point height dispersion (H_{var}) and range (H_{95}). The model based on data fusion used different metrics from the other models but with similar characteristics; one metric was related to the DAP's point height range (H_{iqr}) and the other to S2 indices based on the red and NIR band (SRI).

3.2 Forest maps and estimation

The biomass forest maps varied substantially across models (Figure 3). The map based on S1 data had the highest minimum biomass value (43 Mg ha^{-1}) and lowest range (464 Mg ha^{-1}). In contrast, the other maps had a minimum biomass of 3-5 Mg ha^{-1} and broader ranges exceeding 600 Mg ha^{-1}, with the S2-based map ranging the most (1,156 Mg ha^{-1}).

The growing stock biomass of the forest fragment following SRS estimation was 186 Mg ha^{-1}. Model-assisted estimates differed from this by up to ~50 Mg ha^{-1} in magnitude, with estimate from data fusion differing the most (137 Mg ha^{-1}) (Figure 4). However, all

Table 4. Predictive model performance based on Sentinel-1 (S1), Sentinel-2 (S2), digital aerial photogrammetry (DAP), and fused sensor data
Tabela 4. Desempenho dos modelos preditivos com base nos dados dos sensores Sentinel-1(S1), Sentinel-2 (S2), fotogrametria aérea digital (DAP) e fusão de dados

DATASET	FITED MODEL	σ^2	r^2	RMSE%	MAE%
S1	$\ln(B) = 6.709 - 0.004\ S1_{var} - 0.996\ S1_{cor}$	0.103	0.20	51.06	39.78
S2	$\ln(B) = 2.845 - 8.695\ DVI + 4.186\ RECI$	0.090	0.33	45.16	35.79
DAP	$\ln(B) = 2.853 - 0.037\ H_{var} + 0.145\ H_{95}$	0.082	0.15	64.29	48.21
Fusion	$\ln(B) = 2.753 - 0.142\ H_{iqr} + 0.253\ SRI$	0.032	0.49	41.27	27.21

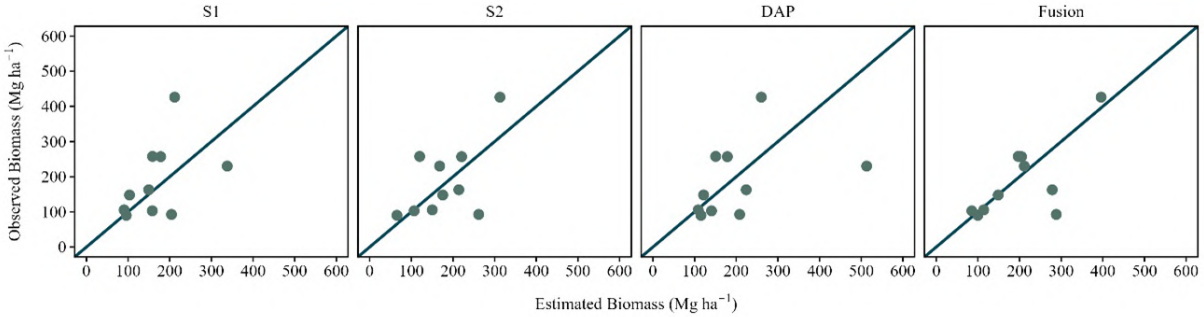

Figure 2. Scatterplots of estimated versus observed biomass values (Mg ha^{-1}) for each modeling approach: Sentinel-1 (S1), Sentinel-2 (S2), digital aerial photogrammetry (DAP), and data fusion. The diagonal line represents the 1:1 relationship
Figura 2. Dispersão dos valores estimados versus observados de biomassa (Mg ha^{-1}) para cada abordagem de modelagem: Sentinel-1 (S1), Sentinel-2 (S2), fotogrametria aérea digital (DAP) e fusão de dados. A linha diagonal representa a relação 1:1

estimated confidence intervals overlapped and encompassed the many estimated means. The largest interval was associated with S1 data (SE = 34 Mg ha^{-1}), while the lowest was found for the S2 data (SE = 21 Mg ha^{-1}), which had comparable accuracy to the SRS estimate (SE = 22 Mg ha^{-1}). The standard errors associated with the DAP- and data-fusion-based models were intermediary and similar (26-27 Mg ha^{-1}).

The certainty gains (REs) from model-assisted estimates, were only observed with S2 model (RE = 1.07), this estimate was 7 percentual points more certain than the SRS estimate. The estimates derived from the other datasets were considered less certain, nevertheless, with REs of 0.69, 0.63, and 0.40 for DAP data, data fusion, and S1 data, respectively.

4. DISCUSSION

In general, the produced models developed in the study achieved accuracy comparable to those reported in previous studies involving multilayered tropical forests (e.g., Bispo et al., 2020; Zimbres et al., 2021). The attribute prediction in structurally complex forests tends to show lover accuracy (e.g., higher RMSE) compared to single-layer forests, such as even-aged plantations (Cosenza et al., 2021a) This reduction in accuracy is explained by the sensor saturation, i.e., a limited sensibility of the predictor variables as forest structural complexity increases (Joshi et al., 2017). Additionally, the relatively small number of plots as used in our study might also impose difficulties in capturing the pattern between the predictors and the response variable (Bouvier et al., 2019). This might explain the larger bias (MAE values) found in our models. Nonetheless, this sample size is not uncommon in tropical forest studies, where logistical constraints often restrict the number of field plots (Proisy et al., 2002; Santos et al., 2003;

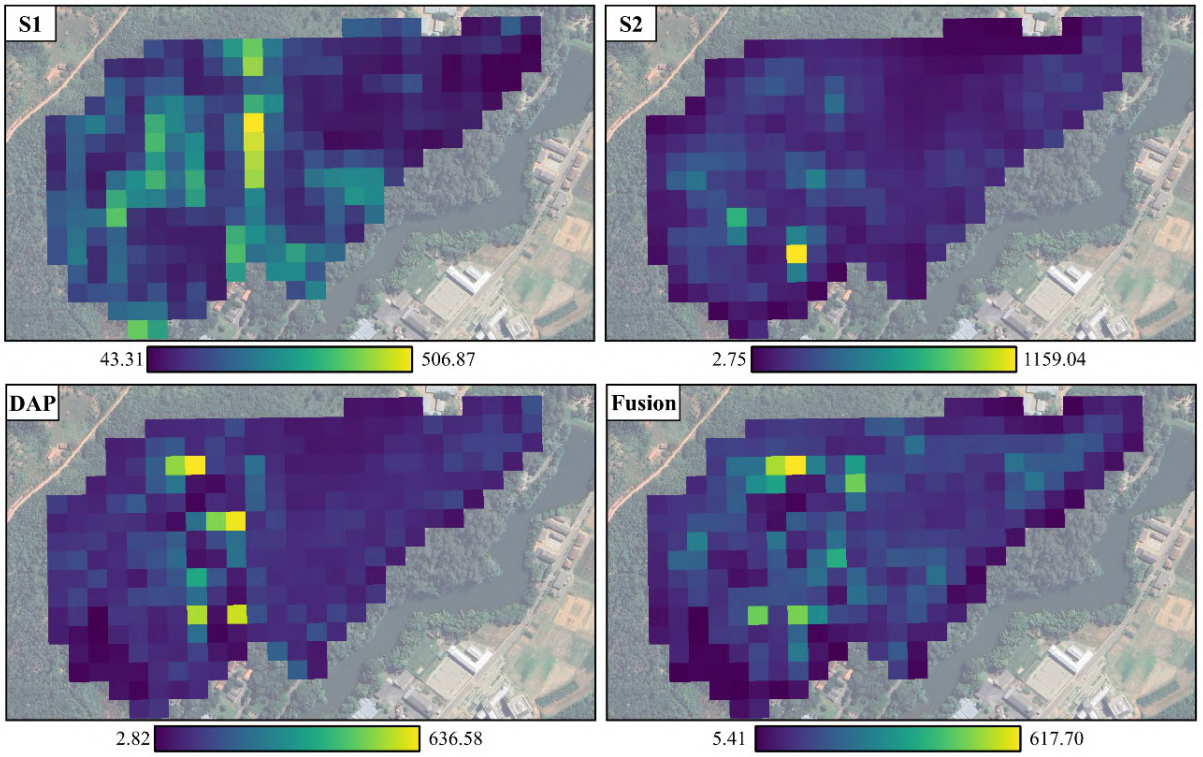

Figure 3. Maps of forest growing stock biomass (Mg ha^{-1}) predicted by fitted models based on Sentinel-1 (S1), Sentinel-2 (S2), digital aerial photogrammetry (DAP), and data fusion. Biomass estimates are shown at the pixel level over the study area
Figura 3. Mapas da biomassa em volume de madeira em pé (Mg ha^{-1}) previstos por modelos ajustados com base nos dados do Sentinel-1 (S1), Sentinel-2 (S2), fotogrametria aérea digital (DAP) e fusão de dados. As estimativas de biomassa são apresentadas em nível de pixel na área de estudo

Mauya et al., 2015). Nevertheless, the use of conservative approaches, such as two-predictor OLS models, proved effective in producing reasonably accurate predictions despite the small sample size.

The models derived from data fusion showed the best performance. This reinforces the complementarity between spectral information from S2 and the three-dimensional structural metrics from DAP. In contrast, the DAP-based model alone exhibited the poorest performed, consistent with findings from other studies conducted in tropical forests (Kachamba et al., 2016). Unlike LiDAR, the DAP systems cannot penetrate the forest canopy and may therefore miss information from understory or dominated trees. Besides, DAP point clouds might miss the ground surface and underestimate the tree height, especially in dense vegetation covers (Dandois & Ellis, 2013; Goodbody et al., 2019). Some studies suggest that the use of LiDAR-based DTM or DTM-independent DAP metrics can improve the reliability of DAP-based analyses (Giannetti et al., 2018), but their effectiveness for steep slope terrains as in our areas must still be assessed.

Models based on S1 and S2 had intermediary performance, and they used predictors combining multiple bands or radar-based textural information. Such a strategy has been proved effective to increase data sensitivity to the forest structure and reduce saturation effects in both optical (Frampton et al., 2013) and radar data (Fang et al., 2023). Specifically, vegetation indices using red, red-edge, and NIR bands are sensitive to the leaf abundance (Askar et al., 2018), while SAR textural metrics allow for capturing canopy structural complexity (Chen et al., 2019).

The pattern found in model performances was not strictly followed in the estimation. The estimate derived from S2 data was considered the more precise

Revista ÁRVORE
ISSN 1806 - 9088

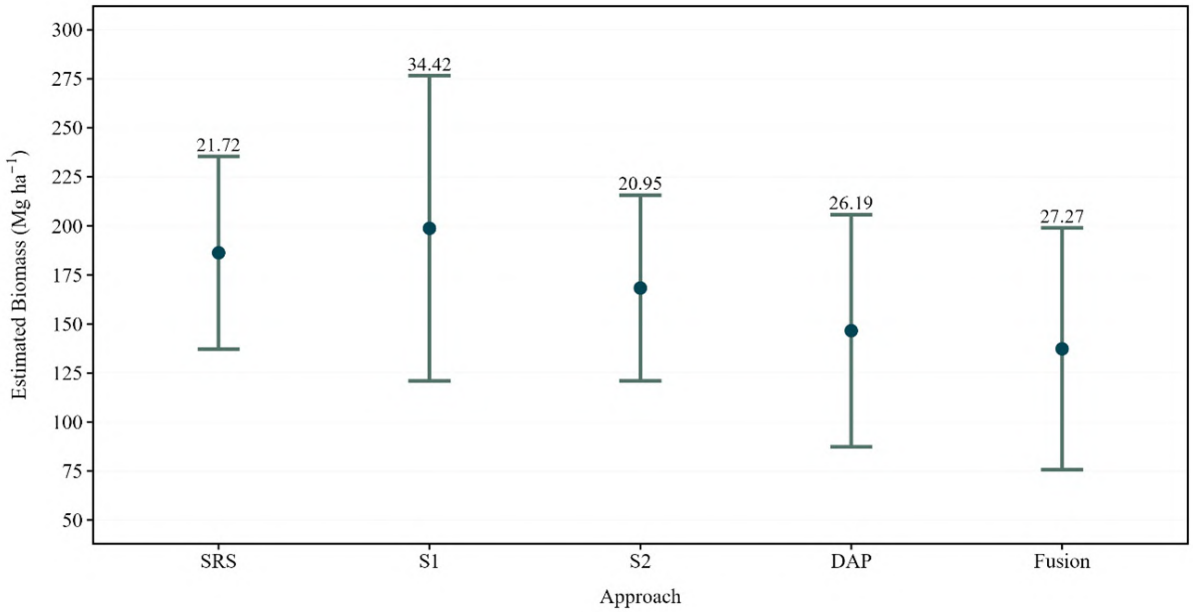

Figure 4. Forest growing stock biomass estimation following simple random sampling (SRS), and model-assisted framework using Sentinel-1 (S1), Sentinel-2 (S2), digital aerial photogrammetry (DAP), and data fusion. Midpoints and whiskers indicate the estimated population mean and confidence intervals, while the standard errors (SE) are indicated in the numbers above each whisker
Figura 4. Estimativa da biomassa do estoque de madeira em pé com base na amostragem aleatória simples (SRS) e em abordagem com auxílio de modelos utilizando os sensores Sentinel-1 (S1), Sentinel-2 (S2), fotogrametria aérea digital (DAP) e fusão de dados. Os pontos centrais e os limites do intervalo indicam a média populacional estimada e os intervalos de confiança, enquanto erros padrões (SE) são representados pelos valores acima de cada limite do intervalo

(narrowest confidence interval), indicating the highest precision in estimating population mean. This is expected in model-assisted estimation, as the regression estimators leverages the relationship between observed values model predictions to reduce the variance of the population mean estimate (McRoberts et al., 2022). The SRS, however, produced slightly larger but still comparable confidence intervals, whereas the uncertainty associated with the other estimates (S1, DAP, and data fusion) was greater. The difference observed between the model results and the population estimates can be attributed to the heterogeneity of the fragment itself, which includes regions in earlier successional stages. The lower vertical structural complexity of these areas likely reduces the accuracy of estimates derived from metrics such as DAP and S1 data. In this context, data fusion performed worse than S2 as DAP metrics add little relevant information for biomass estimation. Studies conducted in boreal forests noted that variables related to vertical structure derived from S1 and ALOS-2 PALSAR increased model complexity without resulting in significant improvements in the estimates, while S2 combined with environmental variables, delivered the best performance (Liu et al., 2024). The boreal forest exhibits low vertical structural complexity, reinforcing the idea that these sensors may have limited effectiveness in environments with less stratified vegetation. Therefore, in situations where S2 data does not exhibit saturation, the sensor shows strong potential for more accurate biomass mapping.

Studies focusing on Brazilian tropical forests have shown that SRS estimators perform well in forest fragments (Mello et al., 2015). However, it is possible that the small number of sample plots was not sufficient to correct the synthetic estimator (left hand in Equation 11) for S1, DAP, and data fusion model. As demonstrated in previous works, overfit models might increase estimation uncertainty considerably, leading to unrealistic confidence intervals (Cosenza et al., 2024). Although our models

were rather conservative, it is likely the large bias found for these models (e.g., DAP and S1 models) be the cause of the increased SEs. Still, model-assisted estimators might be considered a good alternative in scenarios with limited sample size (Cheng et al., 2025). Despite the strategies adopted to mitigate overfitting (e.g., OLS with two predictors), the limited number of sample units likely reduced the robustness of the models. While multi-source remote sensing partially mitigated the effects of the reduced sample size, the estimates still exhibited considerable variability and potential bias. Therefore, expanding the sample size remains essential for enhancing the robustness and generalizability of model-assisted estimation in this biome. Thus, our findings should be generalized only to secondary Atlantic Forest fragments with similar structural characteristics, and extrapolations to other forest types or biomes should be made cautiously.

Although our data set does not allow us to infer about estimation biasness, future studies may explore the use of external models to ensure estimation unbiasedness in a model-assisted framework. External models are models built using distinct data from those used to correct the synthetic estimator in Equation 11 (McRoberts et al., 2022). The model-assisted estimates using external models are considered unbiased (Särndal et al., 1992, Chapter 6), though they would require external data for model calibration. Evaluating this approach, particularly in fragmented and topographically complex landscapes such as the Atlantic Forest, remains an important avenue for future research.

5. CONCLUSION

This study demonstrated the potential of using remote sensing data to estimate growing stock biomass in Atlantic Forest fragments, where the reduced sample sizes are common due to logistical constraints. Three conclusions can be taken from this work: First, the model based on DAP and S2 data fusion had the best performance for mapping the biomass, while models based solely on DAP data and standard DAP metric should be discouraged in similar contexts. Second, model-assisted regression estimators based on S2 data contributed to reducing the confidence intervals of population estimates, proving to be an effective approach to improving forest inventory precision, particularly in fragments with limited sampling. Third, estimates based on DAP, S1, and fused (DAP + S2) data were not advantageous (i.e., less certain than SRS) and might suffer from effects of the small samples.

From an applied perspective, the results indicate that S2 data offer a promising approach for monitoring forest fragments, as they are freely available and effective for attribute estimation. DAP data should not be discouraged, though its use requires additional resources for acquisition in forest fragments. These insights may support environmental agencies in rapidly accessing forest inventory information and defining cost-effective monitoring strategies. Increasing sample size remains a fundamental strategy to improve the precision and reliability of model-assisted estimators while enhancing the forest attribute mapping. Besides, improving the treatment of DAP data using LiDAR-derived products and using SAR interferometry (InSAR) or tomography (TomoSAR) might be an alternative to increase the accuracy of DAP- and S1-based maps. These topics require investigation in future studies targeting forest fragments.

AUTHOR CONTRIBUTIONS

Matos, T.B.: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Visualization; Writing – original draft; Writing – review & editing. Moura, A.C.: Formal analysis; Software; Validation; Writing – review & editing. Paulo, M.V.: Formal analysis; Software; Validation; Writing – review & editing. Santos, V.S.: Validation; Writing – review & editing. Azevedo, L.R.: Validation; Writing – review & editing. Lorenzon, A.S.: Resources; Validation; Writing – review & editing. Torres, C.M.M.E.: Data curation; Resources; Validation; Writing – review & editing. Fernandes Filho, E.I.: Validation; Writing – review & editing. Possato, E.L.: Data curation; Resources; Validation; Writing – review & editing. Gleriani, J.M.: Resources; Validation; Writing – review &

Revista ÁRVORE
ISSN 1806 - 9088

editing. Cosenza, D.N.: Conceptualization; Formal analysis; Investigation; Methodology; Resources; Software; Supervision; Writing – review & editing.

DATA AVAILABILITY

The entire dataset supporting the findings of this study has been published within the article.

7. REFERENCES

Amaro, M. A. (2010). *Quantificação do estoque volumétrico, de estacional semidecidual no município de Viçosa-MG* (Tese de doutorado). Universidade Federal de Viçosa.

Askar, Nuthammachot, N., Phairuang, W., Wicaksono, P., & Sayektiningsih, T. (2018). Estimating aboveground biomass on private forest using Sentinel-2 imagery. *Journal of Sensors*, *2018*. https://doi.org/10.1155/2018/6745629

Beauchamp, J. J., & Olson, J. S. (1973). Corrections for bias in regression estimates after logarithmic transformation. *Ecology, 54*(6), 1403–1407. https://doi.org/10.2307/1934208

Bergseng, E., Ørka, H. O., Næsset, E., & Gobakken, T. (2015). Assessing forest inventory information obtained from different inventory approaches and remote sensing data sources. *Annals of Forest Science, 72*(1), 33–45. https://doi.org/10.1007/s13595-014-0389-x

Bispo, P. da C., Rodríguez-Veiga, P., Zimbres, B., do Couto de Miranda, S., Giusti Cezare, C. H., Fleming, S., Baldacchino, F., Louis, V., Rains, D., Garcia, M., Del Bon Espírito-Santo, F., Roitman, I., Pacheco-Pascagaza, A. M., Gou, Y., Roberts, J., Barrett, K., Ferreira, L. G., Shimbo, J. Z., Alencar, A., … Balzter, H. (2020). Woody aboveground biomass mapping of the Brazilian savanna with a multi-sensor and machine learning approach. *Remote Sensing, 12*(17), 2685. https://doi.org/10.3390/rs12172685

Bouvier, M., Durrieu, S., Fournier, R. A., Saint-Geours, N., Guyon, D., Grau, E., & de Boissieu, F. (2019). Influence of sampling design parameters on biomass predictions derived from airborne lidar data. Canadian *Journal of Remote Sensing, 45*(5), 650–672. https://doi.org/10.1080/07038992.2019.1669013

Broge, N. H., & Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Remote Sensing of Environment, 76(2), 156–172. https://doi.org/10.1016/S0034-4257(00)00197-8

Brown, S., & Lugo, A. E. (1984). Biomass of tropical forests: A new estimate based on forest volumes. *Science, 223*(4642), 1290–1293. https://doi.org/10.1126/science.223.4642.1290

Chappelle, E. W., Kim, M. S., & McMurtrey, J. E. (1992). Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll A, chlorophyll B, and carotenoids in soybean leaves. *Remote Sensing of Environment, 39*, 239–247. https://doi.org/10.1016/0034-4257(92)90089-3

Chen, L., Wang, Y., Ren, C., Zhang, B., & Wang, Z. (2019). Optimal combination of predictors and algorithms for forest above-ground biomass mapping from sentinel and SRTM data. *Remote Sensing, 11*(4), 414. https://doi.org/10.3390/RS11040414

Cheng, X., Hou, Z., Kangas, A., Renaud, J. P., Tang, H., Zeng, W., & Xu, Q. (2025). Alleviating small sample problem in continuous forest monitoring with remote sensing-assisted Copulas. *Ecological Indicators, 171*, 113132. https://doi.org/10.1016/j.ecolind.2025.113132

Chukwu, O., & Dau, J. H. (2020). Forest inventory. In P. Editor (Ed.), *Practice, progress, and proficiency in sustainability* (pp. 306–322). IGI Global. https://doi.org/10.4018/978-1-7998-0014-9.ch016

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.

Cosenza, D. N., Korhonen, L., Maltamo, M., Packalen, P., Strunk, J. L., Næsset, E., Gobakken, T., Soares, P., & Tomé, M. (2021a). Comparison of linear regression, k-nearest neighbour and random forest methods in airborne laser-scanning-based prediction of growing stock. *Forestry: An International Journal of Forest Research, 94*(2), 311–323. https://doi.org/10.1093/forestry/cpaa034

Cosenza, D. N., Packalen, P., Maltamo, M., Varvia, P., Räty, J., Soares, P., Tomé, M., Strunk, J. L., & Korhonen, L. (2021b). Effects of numbers of observations and predictors for various model types on the performance of forest inventory with airborne laser scanning. Canadian Journal of Forest Research, 52, 385–395. https://doi.org/10.1139/cjfr-2021-0192

Cosenza, D. N., Saarela, S., Strunk, J., Korhonen, L., Maltamo, M., & Packalen, P. (2024). Effects of model-overfit on model-assisted forest inventory in boreal forests with remote sensing data. Forestry: An International Journal of Forest Research, 98. https://doi.org/10.1093/forestry/cpae055

Costa, S. M. A., Lima, M. A. A., Júnior, N., Abreu, M. A., Silva, A., & Fortes, L. P. S. e. (2008). RBMC em tempo real, via NTRIP, e seus benefícios nos levantamentos RTK e DGPS. II Simpósio Brasileiro de Ciências Geodésicas e Tecnologias Da Geoinformação, 8–11.

Dandois, J. P., & Ellis, E. C. (2013). High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. *Remote Sensing of Environment, 136*, 259–276. https://doi.org/10.1016/j.rse.2013.04.005

Datt, B. (1998). Remote sensing of chlorophyll A, chlorophyll B, chlorophyll A+B, and total carotenoid content in eucalyptus leaves. *Remote Sensing of Environment, 66,* 111–121. https://doi.org/10.1016/s0034-4257(98)00046-7

David, H. C., Vibrans, A. C., Martins-Neto, R. P., Dalla Corte, A. P., & Péllico Netto, S. (2024). Incorporating forest mapping-related uncertainty into the error propagation of wall-to-wall biomass maps: a general approach for large and small areas. *Remote Sensing, 16*(22), 4295. https://doi.org/10.3390/rs16224295

DJI. (2023). DJI Mavic 3M User Manual (Versão 1.2). https://dl.djicdn.com/downloads/DJI_Mavic_3_Enterprise/20230531/DJI_Mavic_3M_UM-pt-br_v1.2.pdf

ESA. (2024). The Sentinel missions. European Space Agency; European Space Agency (ESA). https://www.esa.int/applications/observing_the_earth/copernicus/the_sentinel_missions

Fang, G., Yu, H., Fang, L., & Zheng, X. (2023). Synergistic use of Sentinel-1 and Sentinel-2 based on different preprocessing for predicting forest aboveground biomass. *Forests, 14*, 1615. https://doi.org/10.3390/f14081615

FAO. (2020). Global Forest Resources Assessment 2020: Main report. FAO. https://doi.org/10.4324/9781315184487-1

Fassnacht, F. E., Hartig, F., Latifi, H., Berger, C., Hernández, J., Corvalán, P., & Koch, B. (2014). Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment, 154*(1), 102–114. https://doi.org/10.1016/j.rse.2014.07.028

Frampton, W. J., Dash, J., Watmough, G., & Milton, E. J. (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing, 82*, 83–92. https://doi.org/10.1016/j.isprsjprs.2013.04.007

García, M. J. L., & Caselles, V. (1991). Mapping burns and natural reforestation using Thematic Mapper data. *Geocarto International, 6*, 31–37. https://doi.org/10.1080/10106049109354290

Ghosh, S. M., & Behera, M. D. (2018). Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest. *Applied Geography, 96*, 29–40. https://doi.org/10.1016/j.apgeog.2018.05.011

Giannetti, F., Chirici, G., Gobakken, T., Næsset, E., Travaglini, D., & Puliti, S. (2018). A new approach with DTM-independent metrics for forest growing stock prediction using UAV photogrammetric data. *Remote Sensing of Environment, 213,* 195–205. https://doi.org/10.1016/j.rse.2018.05.016

Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology, 160*, 271–282. https://doi.org/10.1078/0176-1617-00887

Revista ÁRVORE
ISSN 1806 - 9088

Gitelson, A. A., & Merzlyak, M. N. (1998). Remote sensing of chlorophyll concentration in higher plant leaves. *Advances in Space Research, 22*, 689–692. https://doi.org/10.1016/s0273-1177(97)01133-2

Gitelson, A., & Merzlyak, M. N. (1994). Spectral reflectance changes associated with autumn senescence of Aesculus hippocastanum L. and Acer platanoides L. leaves. spectral features and relation to chlorophyll estimation. *Journal of Plant Physiology, 143,* 286–292. https://doi.org/10.1016/s0176-1617(11)81633-0

Goodbody, T. R. H., Coops, N. C., & White, J. C. (2019). Digital aerial photogrammetry for updating area-based forest inventories: A review of opportunities, challenges, and Future Directions. *Current Forestry Reports, 5*(2), 55–75. https://doi.org/10.1007/s40725-019-00087-2

Google. (2024a). Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A | Earth Engine Data Catalog. https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED

Google. (2024b). Sentinel-1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected, log scaling | Earth Engine Data Catalog. https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD

Gregoire, T. G., Næsset, E., McRoberts, R. E., Ståhl, G., Andersen, H. E., Gobakken, T., Ene, L., & Nelson, R. (2016). Statistical rigor in lidar-assisted estimation of aboveground forest biomass. *Remote Sensing of Environment, 173,* 98–108. https://doi.org/10.1016/j.rse.2015.11.012

Guyot, G., Baret, F., Broge, N. H., Leblanc, E., Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., Mochizuki, K., & Basnyat, B. (2021). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Journal of Applied Remote Sensing, 12*, 1. https://doi.org/10.3126/forestry.v18i01.41749

Hijmans, R. J. (2023). _terra: Spatial Data Analysis_. https://cran.r-project.org/web/packages/terra/terra.pdf

Houghton, R. A., Hall, F., & Goetz, S. J. (2009). Importance of biomass in the global carbon cycle. *Journal of Geophysical Research: Biogeosciences, 114*(3). https://doi.org/10.1029/2009jg000935

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment, 25*(3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-x

IBGE. (2012). Manual técnico da vegetação brasileira: sistema fitogeográfico inventário das formações florestais e campestres: técnicas e manejo de coleções botânicas: procedimentos para mapeamentos. https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=263011

IBGE. (2022). Malha Municipal do Brasil 2022. Instituto Brasileiro de Geografia e Estatística. www.ibge.gov.br/geociencias/downloads-geociencias.html

IBGE. (2024). Panorama do Censo 2022. Instituto Brasileiro de Geografia e Estatística. https://censo2022.ibge.gov.br/panorama/

Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology, 50*, 663–666. https://doi.org/10.2307/1936256

Joshi, N., Mitchard, E. T. A., Brolly, M., Schumacher, J., Fernández-Landa, A., Johannsen, V. K., Marchamalo, M., & Fensholt, R. (2017). Understanding "saturation" of radar signals over forests. *Scientific Reports, 7*(1), 1–11. https://doi.org/10.1038/s41598-017-03469-3

Kachamba, D. J., Ørka, H. O., Gobakken, T., Eid, T., Mwase, W., Melgani, F., Nex, F., Moreno, J., Atzberger, C., & Thenkabail, P. S. (2016). Biomass estimation using 3D data from unmanned aerial vehicle imagery in a tropical woodland. *Remote Sensing, 8*(11), 968. https://doi.org/10.3390/rs8110968

Liu, H. Q., & Huete, A. (2019). A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing, 33*(2), 457–465. https://doi.org/10.1109/tgrs.1995.8746027

Liu, P., Ren, C., Yang, X., Wang, Z., Jia, M., Zhao, C., Yu, W., & Ren, H. (2024). Combining Sentinel-2 and diverse environmental data largely improved aboveground biomass estimation in China's boreal forests. Scientific Reports, 14(1), 1–17. https://doi.org/10.1038/s41598-024-78615-9

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), 3139. https://doi.org/10.21105/joss.03139

Lumley, T. (2024). leaps: Regression subset selection. https://cran.r-project.org/package=leaps

Ma, T., Zhang, C., Ji, L., Zuo, Z., Beckline, M., Hu, Y., Li, X., & Xiao, X. (2024). Development of forest aboveground biomass estimation, its problems and future solutions: A review. *Ecological Indicators, 159*, 111653. https://doi.org/10.1016/j.ecolind.2024.111653

Maltamo, M., Næsset, E., & Vauhkonen, J. (2014). Forestry applications of airborne laser scanning. *Managing Forest Ecosystems, 27*, 460.

MapBiomas. (2024). Projeto Mapbiomas – Coleção 9 da série anual de mapas de cobertura e uso da terra do Brasil. https://brasil.mapbiomas.org/infograficos/

Mariscal Flores, E. J. (1993). *Potencial produtivo e alternativas de manejo sustentável de um fragmento de Mata Atlântica secundária, município de Viçosa, Minas Gerais* (Dissertação de mestrado). Universidade Federal de Viçosa.

Mauya, E. W., Hansen, E. H., Gobakken, T., Bollandsås, O. M., Malimbwi, R. E., & Næsset, E. (2015). Effects of field plot size on prediction accuracy of aboveground biomass in airborne laser scanning-assisted inventories in tropical rain forests of Tanzania. *Carbon Balance and Management, 10*(1), 10. https://doi.org/10.1186/s13021-015-0021-x

McRoberts, R. E., Næsset, E., & Gobakken, T. (2013). Inference for lidar-assisted estimation of forest growing stock volume. *Remote Sensing of Environment, 128,* 268–275. https://doi.org/10.1016/j.rse.2012.10.007

McRoberts, R. E., Næsset, E., Heikkinen, J., Chen, Q., Strimbu, V., Esteban, J., Hou, Z., Giannetti, F., Mohammadi, J., & Chirici, G. (2022). On the model-assisted regression estimators using remotely sensed auxiliary data. *Remote Sensing of Environment, 281*, 113168. https://doi.org/10.1016/j.rse.2022.113168

McRoberts, R. E., Tomppo, E. O., Vibrans, A. C., & Freitas, J. V. de. (2013). Design considerations for tropical forest inventories. *Pesquisa Florestal Brasileira, 33*(74), 189–202. https://doi.org/10.4336/2013.pfb.33.74.430

McRoberts, R. E., Vibrans, A. C., Sannier, C., Næsset, E., Hansen, M. C., Walters, B. F., & Lingner, D. V. (2016). Methods for evaluating the utilities of local and global maps for increasing the precision of estimates of subtropical forest area. *Canadian Journal of Forest Research, 46*(7), 924–932. https://doi.org/10.1139/cjfr-2016-0064

Meira Neto, J. A. (1997). *Estudos florísticos, estruturais e ambientais nos estratos arbóreo e herbáceo-arbustivo de uma floresta estacional semidecidual em Viçosa, MG* (Tese de doutorado). Universidade Estadual de Campinas.

Mello, J. de M., Scolforo, H. F., Raimundo, M. R., Scolforo, J. R. S., Oliveira, A. de D., & Filho, A. C. F. (2015). Estimating precision of systematic sampling in forest. *Ciência e Agrotecnologia, 39*(1), 15–22. https://doi.org/10.1590/s1413-70542015000100002

Moghimi, A., Tavakoli Darestani, A., Mostofi, N., Fathi, M., & Amani, M. (2024). Improving forest above-ground biomass estimation using genetic-based feature selection from Sentinel-1 and Sentinel-2 data (case study of the Noor forest area in Iran). *Kuwait Journal of Science, 51*(2), 100159. https://doi.org/10.1016/j.kjs.2023.11.008

Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., & Papathanassiou, K. P. (2013). A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine, 1*(1), 6–43. https://doi.org/10.1109/mgrs.2013.2248301

Revista ÁRVORE
ISSN 1806 - 9088

Mullissa, A., Vollrath, A., Odongo-Braun, C., Slagter, B., Balling, J., Gou, Y., Gorelick, N., & Reiche, J. (2021). Sentinel-1 SAR backscatter analysis ready data preparation in Google Earth Engine. *Remote Sensing, 13*, 1954. https://doi.org/10.3390/rs13101954

Myers, N., Mittermeler, R. A., Mittermeler, C. G., Da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature, 403*(6772), 853–858. https://doi.org/10.1038/35002501

Myers, R. H. (1990). *Classical and modern regression with applications.* PWS-Kent Publishing Company.

Navarro, J. A., Algeet, N., Fernández-Landa, A., Esteban, J., Rodríguez-Noriega, P., & Guillén-Climent, M. L. (2019). Integration of UAV, Sentinel-1, and Sentinel-2 aata for mangrove plantation aboveground biomass monitoring in Senegal. *Remote Sensing, 11*(1). https://doi.org/10.3390/rs11010077

Proisy, C., Mougin, E., Fromard, F., Trichon, V., & Karam, M. A. (2002). On the influence of canopy structure on the radar backscattering of mangrove forests. *International Journal of Remote Sensing, 23*(20), 4197–4210. https://doi.org/10.1080/01431160110107725

R Core Team. (2024). R: The R project for statistical computing. https://www.r-project.org

Rao, J. N. K., & Molina, I. (2015). *Small area estimation* (2nd ed.). Wiley. https://doi.org/10.1002/9781118735855

Ribeiro, M. C., Martensen, A. C., Metzger, J. P., Tabarelli, M., Scarano, F., & Fortin, M.-J. (2011). The Brazilian Atlantic Forest: A shrinking biodiversity hotspot. In *Biodiversity hotspots* (pp. 405–434). Springer. https://doi.org/10.1007/978-3-642-20992-5_21

Richardson, A. J., & Wiegand, C. L. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing, 43*(12), 1541–1552. https://www.asprs.org/wp-content/uploads/pers/1977journal/dec/1977_dec_1541-1552.pdf

Rouse, J. W., Jr., Haas, R. H., Deering, D. W., Schell, J. A., & Harlan, J. C. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation (NASA Contractor Report No. NASA-CR-144661). NASA. https://ntrs.nasa.gov/api/citations/19750020419/downloads/19750020419.pdf

Roussel, J. R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J. F., de Boissieu, F., & Achim, A. (2020). lidR: An R package for analysis of airborne laser scanning (ALS) data. *Remote Sensing of Environment, 251*, https://doi.org/10.1016/j.rse.2020.112061

Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordkvist, K., & Hyyppä, J. (2015). Model-assisted estimation of growing stock volume using different combinations of lidar and Landsat data as auxiliary information. *Remote Sensing of Environment, 158*, 431–440. https://doi.org/10.1016/j.rse.2014.11.020

Santos, J. R., Freitas, C. C., Araujo, L. S., Dutra, L. V., Mura, J. C., Gama, F. F., Soler, L. S., & Sant'Anna, S. J. S. (2003). Airborne P-band SAR applied to the aboveground biomass studies in the Brazilian tropical rainforest. *Remote Sensing of Environment, 87*(4), 482–493. https://doi.org/10.1016/j.rse.2002.12.001

Särndal, C.-E., Swensson, B., & Wretman, J. (1992). *Model assisted survey sampling.* Springer-Verlag, Inc.

Soares, C. P. B. (2017). Mensuração Florestal. https://www.mensuracaoflorestal.com.br/

Souza, C. R., Maia, V. A., Aguiar-Campos, N. de, Santos, A. B. M., Rodrigues, A. F., Farrapo, C. L., Gianasi, F. M., Paula, G. G. P. de, Fagundes, N. C. A., Silva, W. B., & Santos, R. M. (2021). Long-term ecological trends of small secondary forests of the atlantic forest hotspot: A 30-year study case. *Forest Ecology and Management, 489*, 119043. https://doi.org/10.1016/j.foreco.2021.119043

Ståhl, G., Saarela, S., Schnell, S., Holm, S., Breidenbach, J., Healey, S. P., Patterson, P. L., Magnussen, S., Næsset, E., McRoberts, R. E., & Gregoire, T. G. (2016). Use of models in large-area forest surveys: comparing model-assisted, model-based and hybrid estimation. *Forest Ecosystems, 3*(1). https://doi.org/10.1186/s40663-016-0064-9

Universidade Federal de Viçosa - UFV. (2024). Estação Climatológica Principal de Viçosa. https://posmet.ufv.br/boletim-meteorologico/

Verly, O. M., Leite, R. V., Tavares-Junior, I. da S., Rocha, S. J. S. S. da, Leite, H. G., Gleriani, J. M., Rufino, M. P. M. X., Silva, V. de F., Torres, C. M. M. E., Plata-Rueda, A., Castro, B. M. de C. e, Zanuncio, J. C., & Jacovine, L. A. G. (2023). Atlantic forest woody carbon stock estimation for different successional stages using Sentinel-2 data. *Ecological Indicators, 146*, 109870. https://doi.org/10.1016/j.ecolind.2023.109870

Vibrans, A. C., McRoberts, R. E., Moser, P., & Nicoletti, A. L. (2013). Using satellite image-based maps and ground inventory data to estimate the area of the remaining Atlantic forest in the Brazilian state of Santa Catarina. *Remote Sensing of Environment, 130,* 87–95. https://doi.org/10.1016/j.rse.2012.10.023

Wang, C., Zhang, W., Ji, Y., Marino, A., Li, C., Wang, L., Zhao, H., & Wang, M. (2024). Estimation of aboveground biomass for different forest types using data from Sentinel-1, Sentinel-2, ALOS PALSAR-2, and GEDI. *Forests, 15*(1), 215. https://doi.org/10.3390/f15010215

Zhang, L., Zhang, X., Shao, Z., Jiang, W., & Gao, H. (2023). Integrating Sentinel-1 and 2 with lidar data to estimate aboveground biomass of subtropical forests in northeast Guangdong, China. *International Journal of Digital Earth, 16*(1), 158–182. https://doi.org/10.1080/17538947.2023.2165180

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., & Yan, G. (2016). An easy-to-use airborne lidar data filtering method based on cloth simulation. *Remote Sensing, 8*(6), 501. https://doi.org/10.3390/rs8060501

Zimbres, B., Rodríguez-Veiga, P., Shimbo, J. Z., da Conceição Bispo, P., Balzter, H., Bustamante, M., Roitman, I., Haidar, R., Miranda, S., Gomes, L., Alvim Carvalho, F., Lenza, E., Maracahipes-Santos, L., Abadia, A. C., do Prado Júnior, J. A., Mendonça Machado, E. L., Dias Gonzaga, A. P., de Castro Nunes Santos Terra, M., de Mello, J. M., … Alencar, A. (2021). Mapping the stock and spatial distribution of aboveground woody biomass in the native vegetation of the Brazilian Cerrado biome. *Forest Ecology and Management, 499*, 119615. https://doi.org/10.1016/j.foreco.2021.119615

Zvoleff, A. (2020). glcm: Calculate textures from grey-level co-occurrence matrices (GLCMs). https://cran.r-project.org/package=glcm