Estimating stem volume of *Eucalyptus* sp. and *Pinus* sp. plantations in Brazil, using Sentinel-1B and ALOS-2/PALSAR-2 data

Juliana Maria Ferreira de Souza Diniz,[a,*] **Fabio Furlan Gama**,[a] **Aliny Aparecida dos Reis**,[b] **Cleber Gonzales de Oliveira**,[c] **and Eduardo Resende Girardi Marques**[d]

[a]National Institute for Space Research – INPE, Earth Observation Coordination, São Paulo, Brazil

[b]University of Campinas – UNICAMP, Interdisciplinary Center of Energy Planning – NIPE, Campinas, Brazil

[c]VISIONA Tecnologia Espacial, São Paulo, Brazil

[d]KLABIN S.A., Telêmaco Borba, Paraná, Brazil

Abstract. Multifrequency synthetic aperture radar (SAR) data have been applied to discriminate subtle differences in the vegetation and to better characterize its structural properties, since each SAR frequency will interact with the different sections of the vegetation canopy. In this study, our main objective was to evaluate the use of multifrequency Sentinel-1 and ALOS-2/PALSAR-2 data for stem volume estimations in *Eucalyptus* sp. and *Pinus* sp. plantations using three different machine learning algorithms: random forest (RF), support vector regression (SVR), and extreme gradient boosting (XGB). Different experiments were carried out using combinations of predictor variables derived from both SAR sensors: backscattering, polarimetric decompositions, and interferometry data, and field data considering specific models for *Eucalyptus* sp. and *Pinus* sp. and a generic model comprising all forest plantations data. The machine learning models using predictor variables derived from SAR data achieved moderately high accuracy to predict stem volume, mainly when SAR data were used in combination with stand age (Experiment iv). In the best prediction scenario (Experiment iv), the RF, SVR, and XGB models were able to explain 81.7%, 68.5%, and 81.8% [coefficient of variation (R^2) values] of stem volume variability considering the generic models, respectively. Our results pointed out that the RF algorithm showed the best performance in predicting stem volume with significant good results and easier implementation in comparison with the other two algorithms (SVR and XGB). [DOI: 10.1117/1.JRS.17.014513]

Keywords: multifrequency; machine learning; polarimetry; synthetic aperture radar.

Paper 220373G received Jun. 30, 2022; accepted for publication Jan. 16, 2023; published online Feb. 6, 2023.

1 Introduction

Using remote sensing data for estimating forest biophysical parameters has increased significantly in the last decades.[1–3] Particularly, synthetic aperture radar (SAR) data allow obtaining information about roughness, tree morphology, canopy density, and soil moisture, and as SAR data present great penetration in the forest canopy,[1] they can be successfully applied for stem volume estimation.[4]

The forest canopy interaction with electromagnetic radiation of SAR data is dependent on the number of polarizations and the operation frequency of the sensor.[5–7] SAR sensors with shorter wavelength, as X- and C-bands, show a great attenuation of backscatter through the tree canopy, interacting with branches and small leaves and thereby being restricted to the upper part of the

*Address all correspondence to Juliana Maria Ferreira de Souza Diniz, julianadinizengflorestal@gmail.com

1931-3195/2023/$28.00 © 2023 SPIE

canopy,[8,9] which reduces the sensors' sensitivity in relation to the vertical structure of the forest.[10,11] Longer wavelength sensors, such as L- and P-bands, have high penetration capacity in the canopy and greater sensitivity to the vertical structure of the forest. Furthermore, L- and P-bands sensors have been more used to estimate forest volume and biomass.[9,12–15] However, until now, there are no SAR sensors in orbit operating in the P-band.

Thus, there is the possibility of using the frequency information to obtain better accuracy in estimating forest biophysical parameters by exploring the different operation frequencies of the available SAR sensors.[16] Through the synergism between multifrequency SAR sensors, subtle differences in the vegetation canopy can be discriminated and its structural properties can be better characterized, as each frequency will interact with the different sections of the vegetation canopy.[16] For example, Santoro et al.[17] analyzed the complementarity of radar backscattering data from the X (TerraSAR-X), C (Sentinel-1), and L (ALOS-2/PALSAR-2) bands to estimate stem volume in a boreal forest in Sweden. Nizalapur et al.[18] used backscattering coefficient of C, L, and P bands in VV polarization of experimental synthetic aperture radar (ESAR) airborne for aboveground biomass estimation in a tropical dry deciduous forest in India. Englhart et al.[14] used SAR backscattering of X and L bands of TerraSAR-X and ALOS-PALSAR, respectively, for aboveground biomass estimation in tropical forests in Indonesia. These studies demonstrate that using multifrequency SAR data leads to an improvement in biomass and stem volume estimation in different forests.

The use of multisensors data demands robust algorithms to deal with non-linear and complex relationships between spectral and biophysical variables.[19–21] Hence, machine learning algorithms have been widely used in remote sensing studies for such tasks. These algorithms can learn from data and deal with the most complex classification and/or regression problems using remote sensing data.[21–23]

Random forest (RF) and support vector machine (SVM) are two machine learning algorithms commonly applied to retrieve forest biophysical parameters using different remote sensing data.[21,23,24] In addition to these algorithms, gradient boosting decision tree algorithms, and more specifically the extreme gradient boost (XGB) algorithm, have demonstrated superior performance in estimating biophysical parameters of different vegetation types.[25,26] RF is a nonlinear and non-parametric ensemble decision-tree method[27] that provides flexible, robust, and accurate predictive capabilities for high dimensional datasets.[22] RF models have been widely and successfully used for forest biophysical parameter estimation.[23,28]

Support vector machine (SVM) is also known as a flexible and robust-to-noise machine learning method. Ataee et al.[29] tested different kernel functions in the SVR models to estimate ground stock volume (GSV), and all SVM models showed good prediction results. Souza et al.[21] compared the results of SVR, multiple linear regression (MLR), artificial neural network (ANN), and RF for predicting *Eucalyptus* stem volume in Minas Gerais using ALOS AVNIR-2 and ALOS PALSAR data. The authors observed that the SVR models showed the best prediction results.

Recently, the XGB regression[30] has been shown to be powerful to solve not only classification but also regression problems.[28] According to Sun et al.,[31] the XGB is a particularly efficient tool to be applied to solve environmental problems. Pham et al.[25] compared different machine learning algorithms, such as XGB, RF, and SVR, for estimating mangrove aboveground biomass using optical (Sentinel-2) and SAR (Sentinel-1 and ALOS-2/PALSAR-2) data. The authors observed that XGB showed the best prediction performance.

Although SAR data has been used for estimating forest biophysical parameters, most studies use SAR data in combination with optical data.[4,21,29] Currently, there is a lack of studies using solely SAR data as well as investigating the combination of multifrequency SAR data with machine learning algorithms to estimate forest biophysical parameters in forest plantations. Thus, our study introduces a novel and unique case study that explores the combination of several advanced polarimetric and interferometric variables derived from multifrequency SAR data for estimating stem volume of forest plantations using robust machine learning algorithms.

The main objective of our study was to evaluate the use of multifrequency Sentinel-1 and ALOS-2/PALSAR-2 for stem volume estimations in *Eucalyptus* sp. and *Pinus* sp. plantations using machine learning algorithms. This study conducted investigations to answer the following questions: (i) How do the different sensor datasets perform in predicting stem volume of

Eucalyptus sp. and *Pinus* sp. plantations?, (ii) Would the use of specific models for *Eucalyptus* sp. and *Pinus* sp. plantations improve the prediction results when compared with using a generic model for both forest plantations?, (iii) Would the use of forest inventory variables improve the prediction results?, (iv) How would the machine learning algorithms RF, support vector regression (SVR), and XGB perform in estimating stem volume, considering the different prediction datasets?

The remainder of this paper is organized as follows. Section 2 describes the study area, the field and SAR based data, the preprocessing workflow of Sentinel-1 and ALOS-2/PALSAR-2 data, and the stem volume estimation using the RF, SVR, and XGB algorithms. Section 3 presents the experiment results of stem volume estimation using specific and generic models and analyzes the performance of the machine learning algorithms. Section 4 provides a discussion about the results and presents comparisons with previous works. Section 5 draws the conclusions and briefly discusses future work.

2 Methodology

2.1 *Study Area*

The study area is located in Telêmaco Borba municipality, Midwest of Paraná state, south region of Brazil (Fig. 1). The area is characterized by having a humid subtropical climate conditions,

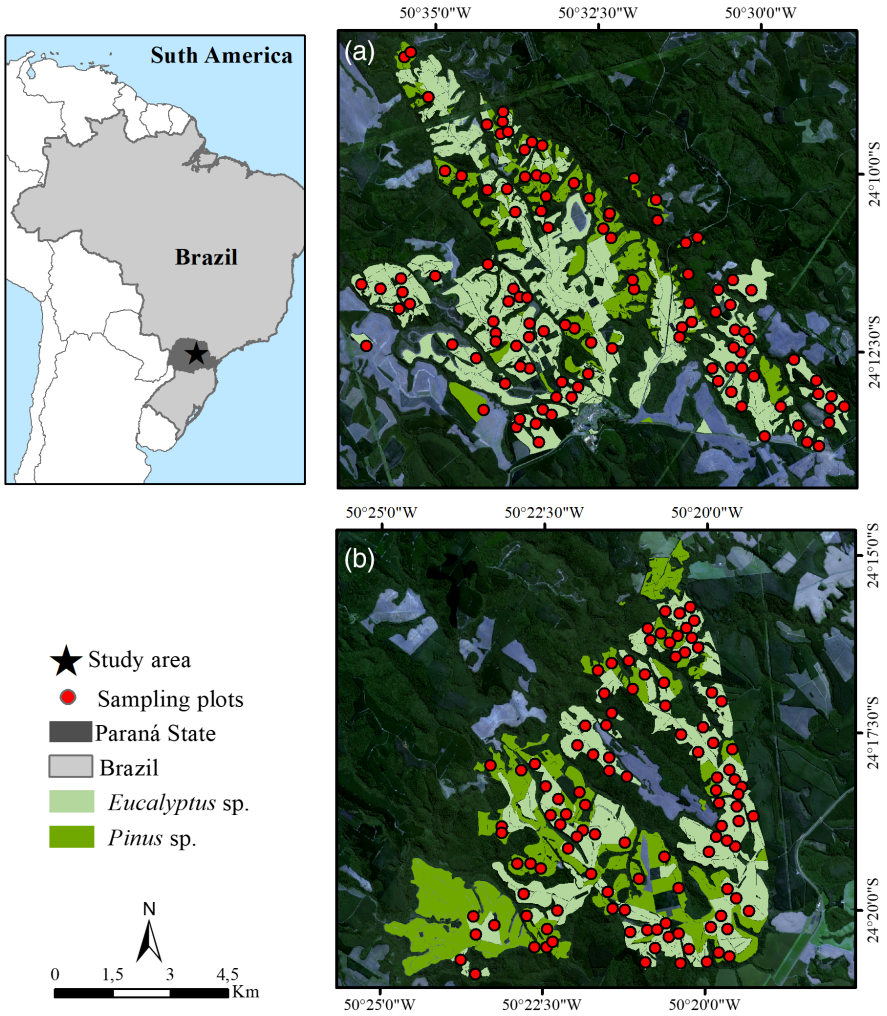

Fig. 1 Location of the study area in Paraná State, Brazil. *Eucalyptus* sp. and *Pinus* sp. stands distribution, and the location of the 220 sampling plots of the forest inventory. (a) Study Area 1; (b) Study Area 2. Image base: Sentinel-2 (432) (April 18, 2019).

corresponding to the transition between Cfa (humid subtropical) and Cfb (oceanic) climatic types, according to the Köppen classification.[32] The average temperature is 23.2°C and 16.3°C in the hottest and coldest months of year, respectively.[33,34] The area has regular rainfall, ranging from 1478 to 1700 mm per year, with average annual humidity between 70% and 75%. The soil is predominantly classified as a dark red latosol with a clayey texture and the relief varies from flat to gently undulating.[32,35]

The study area comprises *Eucalyptus* sp. and *Pinus* sp. plantations, which correspond to a total area of 7150.3 ha. The *Eucalyptus* sp. plantations are composed of stands planted with the following species: *E. saligna*, *E. grandis* × *E. urophylla*, *E. grandis* × *E. camaldulensis*, *E. grandis,* and *E. dunni*. For *Pinus* sp. plantations, the stands were planted with the species *P. taeda* and *P. maximinoi*.

The total area consisted of 508 stands, with 250 stands (4727.1 ha) of *Eucalyptus* sp. and the remaining 258 stands (2423.2 ha) of *Pinus* sp. (Fig. 1). The average age of *Eucalyptus* sp. was 4.6 years old, with the minimum age of 1.4 years and the maximum age of 7.9 years, with spacing ranging from 2.4×3.75 to 4.1×2.19 m. The *Pinus* sp. had the average age of 11.14 years, with the minimum age of 5.2 years and the maximum age of 17.7 years, and the spacing ranging from 2.5×2.4 to 4.0×1.15 m.

2.2 *Fields Measurements*

Field data collection was carried out in 220 sampling plots, 156 plots located at the *Eucalyptus* sp. stands, and 64 plots located at the *Pinus* sp. stands (Fig. 1). The field measurements were carried out between March and April 2019.

In each plot, the diameter at breast height (DBH) of all trees and the height of the 15 dominant trees were measured. The area of the plots varied from 308 to 870 m^2, with an average area of 513 m^2, with only one plot with an area >800 m^2. The stem volume (m^3/ha) was estimated based on the information collected in the plots, using standard forest mensuration equation for this region. For estimating individual tree stem volume, equations based on DBH and total height for each measurement year and DBH class were used. Table 1 summarized the descriptive statistics from the forest inventory.

2.3 *SAR Dataset*

For this study, Sentinel-1B and ALOS-2/PALSAR-2 SAR data were used. Sentinel-1B data were provided from the European Space Agency (ESA). This sensor is characterized by operating in the C-band. The data were acquired in slant range geometry, interferometric wide Swath (IW) mode, with a range of 250 km and a nominal spatial resolution of 5×20 m.[36] Sentinel-1 data have two polarizations VV and VH (Table 2) and were obtained on May 14th and 26th, 2019.

The ALOS-2/PALSAR-2 data, that operates in L-band, were acquired in StripMap (high sensitive) mode, full polarimetric and single look complex (Table 2).

Table 1 Descriptive statistics of the variables obtained from the forest inventory collected in the 220 plots.

	Eucalyptus sp.			*Pinus* sp.		
Variable	Mean	SD	Amplitude	Mean	SD	Amplitude
Age (years)	4.58	2.12	1.40–7.90	11.14	4.30	5.20–17.70
Total height (m)	23.03	7.31	6.70–35.10	15.85	3.57	8.40–24.40
DBH (cm)	14.91	3.44	7.10–21.50	19.45	2.50	15.00–25.10
Volume (m^3/ha)	241.04	130.89	15.00–540.00	323.43	131.68	98.00–791.00

Note: DBH: Diameter at breast height (cm); SD: Standard Deviation.

Table 2 Sentinel-1B and ALOS-2/PALSAR-2 characteristics.

Attribute	Sentinel-1B	ALOS-2/PALSAR-2
Acquisition dates	April 26, 2019 and May 8, 2019	April 28, 2019-04-28 and May 12, 2019
Acquisition mode	Interferometric Wide Swath - IW	StripMap (High sensitive)
Orbit	Descending	Ascending
Swath	250 km	50 to 70 km
Spatial resolution (nominal)	5×20 m	6 m
Wavelength	5.6 cm	~24 cm
Polarizations	VV + VH	HH + HV + VV + VH
Format	Single look complex (SLC) and slant range	

2.4 *Preprocessing*

2.4.1 *Sentinel-1B data*

The Sentinel-1B data, IW mode, were preprocessed using the SNAP software version 7.0 to obtain the SAR predictor variables (Fig. 2). Sentinel-1B images have three subswaths. In this study, the subswath IW2 was selected to cover the study extension and to decrease processing time. The *Apply Orbit File* operation was performed to update the information presented in the image metadata. Then, the *Deburst* processing was applied to join the subswaths. The *Multilooking* processing with 4×1 looks was applied to produce a regular pixel. Next, the *Lee* filter[37] using a 3×3 window was applied to reduce the speckle. The backscattering coefficients and polarimetric decompositions were calculated using the image of April 28, 2019.

Some studies pointed out that the intensity images of Sentinel-1 are affected by additive noise.[38] So, the thermal noise removal procedure was applied to reduce noise effects in the inter-subswaths, normalizing the backscattering signal in the scene.[39] Afterward, the radiometric calibration [Eq. (1)] was carried out to obtain the backscattering parameters (σ^0_{VH} and σ^0_{VV})

$$\sigma_i^0 = \frac{\mathrm{DN}_i^2}{A_i^2}, \tag{1}$$

where DN_i is the digital number values of the pixels, A_i is an absolute calibration constant found in the look up tables (LUTs), and i represents the pixels.

To obtain the H-α polarimetric decomposition variables, the coherence matrix 2×2 [Tdual] was created using dual-polarized data[40]

$$H = \sum_{i=1}^{2} (-P_i \log_2 P_i), \tag{2}$$

$$\alpha = \sum_{i=1}^{2} P_i \alpha_i, \tag{3}$$

where P_i represents the relative importance of the eigenvalues of the coherence matrix (λ_i) and is obtained using

$$P_i = \frac{\lambda_i}{\sum_{j=1}^{2} \lambda_j}. \tag{4}$$

To obtain the interferometric coherence, the Sentinel-1 images were coregistered and the interferometric coherence was calculated using a 10×3 window

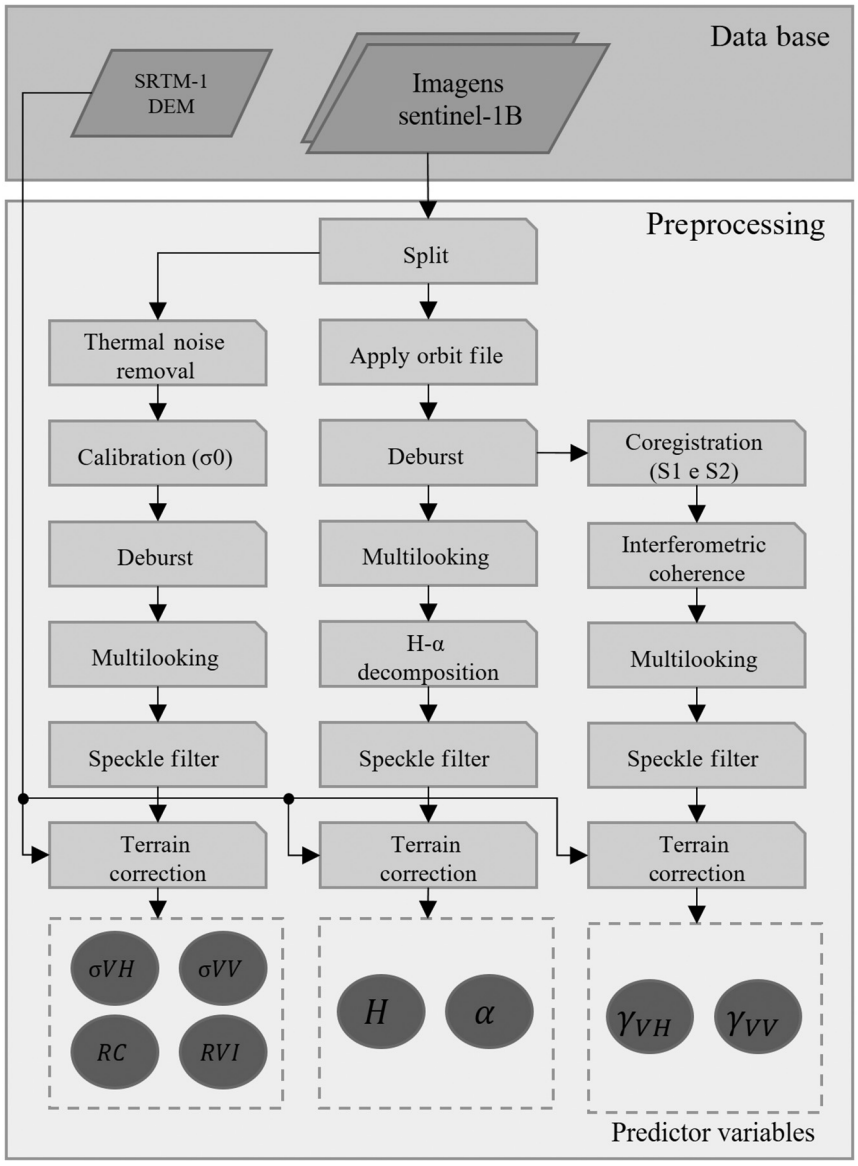

Fig. 2 Sentinel-1B preprocessing flowchart.

$$\gamma = \frac{|\langle S_1(x)S_2(x)^*\rangle|}{\sqrt{\langle |S_1(x)|^2|S_2(x)|^2\rangle}}, \tag{5}$$

where γ represents interferometric coherence; S_1 (April 28, 2019) and S_2 (May 08, 2019) are the two coregistered complex images, and are the average relative to the size of the window.

The *Range Doppler Terrain Correction* was applied to all variables for topographic correction[41] using the SRTM-1 digital elevation model (DEM), with 30 m of spatial resolution, automatically downloaded and used through the SNAP toolbox.

The cross-polarized ratio was obtained through the ratio between σ^0_{VH} and σ^0_{VV} geometrically corrected. The radar vegetation index (RVI) was calculated using Eq. (6) with σ^0_{VH} and σ^0_{VV} geometrically corrected[42,43]

$$\mathrm{RVI} = \frac{4\sigma^0_{\mathrm{VH}}}{(\sigma^0_{\mathrm{VV}} + \sigma^0_{\mathrm{VH}})}. \tag{6}$$

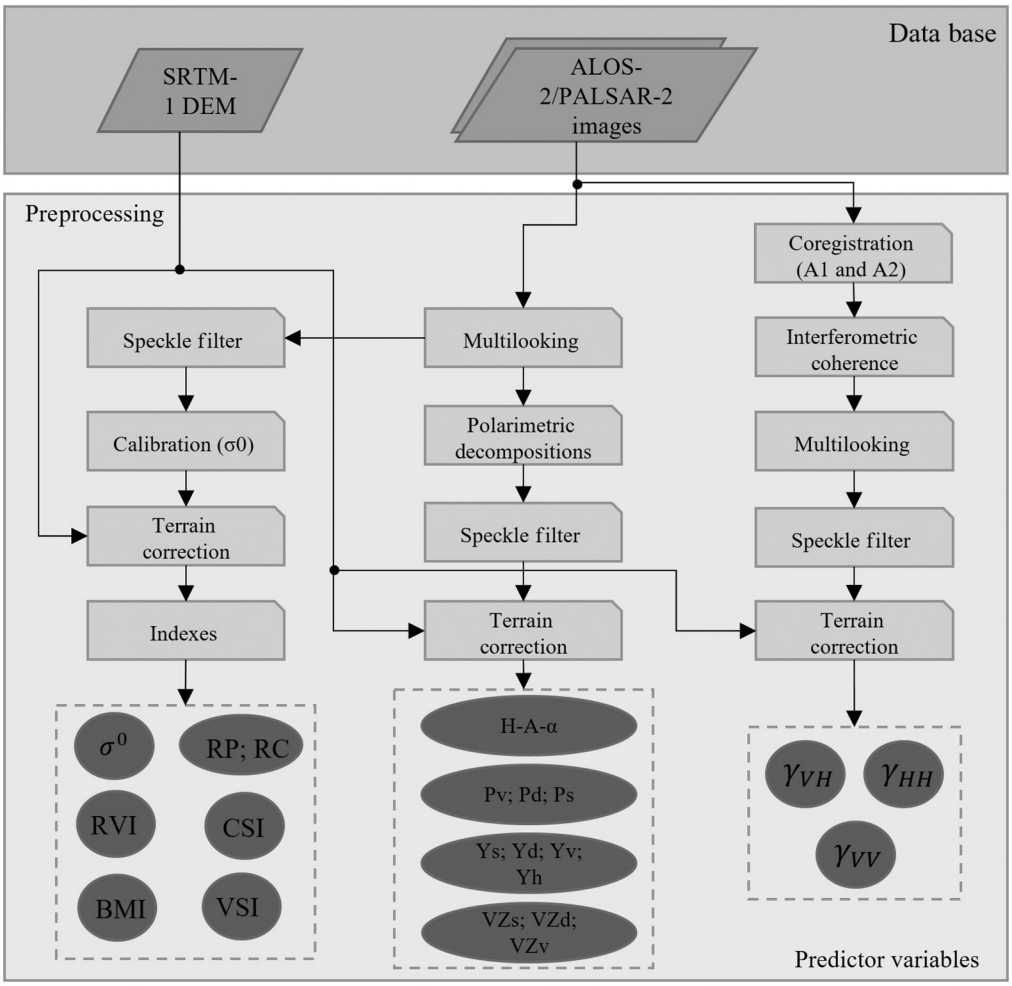

Fig. 3 ALOS-2/PALSAR-2 preprocessing flowchart. The symbols are described in Table 2.

2.4.2 *ALOS-2/PALSAR-2 data*

The ALOS-2/PALSAR-2 SLC image of April 28, 2019, was processed in SNAP software to obtain backscattering and polarimetric decomposition variables. The interferometric coherence was obtained using the SARScape software[44] and the pair of images April 28, 2019 (A_1) and May 12, 2019 (A_2).

The first processing step was to apply the multilooking with two looks in range and five looks in azimuth to produce a regular pixel (Fig. 3). The Lee filter with 3×3 window was also applied to reduce the speckle noise. The backscattering coefficients (σ^0_{HH}, σ^0_{HV}, σ^0_{VV}, and σ^0_{VH}) were obtained based on[45]

$$\sigma^0 = 10.\log_{10}\langle I^2 + Q^2\rangle + \text{CF-}A, \tag{7}$$

where I is the intensity, Q is the quadrature, CF is the calibration factor, and A is the conversion factor equal to 32.

Considering the backscattering coefficients, the ratios between the backscattering (RP and RC) and the indices proposed by Pope et al.[46] – biomass index (BMI), canopy structure index (CSI), and volume scattering index (VSI) – were also calculated (Table 3). To obtain the polarimetric variables, four polarimetric decompositions were carried out: Cloude–Pottier, Freeman–Durden, Yamaguchi, and Van Zyl as shown in Table 3.

Table 3 Description of the polarimetric variables used in this study.

Acronym	Variable	Description	References
		Backscattering	
RP	Copolarization ratio	Highlights the differences between the VV and HH polarizations from vegetation structural aspects.	Henderson and Lewis[47]
RC	Cross-polarization ratio	Sensitive to forest volumetric scattering to support classification and reduce topographic effects on backscattering	
RVI	Radar vegetation index	Parameter sensitive to the biomass level	Kim and Van Zyl[48]
BMI	Biomass index	Relative amount of woody component compared with leaf biomass.	Pope et al.[46]
CSI	Canopy structure index	Relative presence of vertical scatterers (trunks and stems) of vegetation.	
VSI	Volume scattering index	Indicator of canopy thickness or relative density.	
		Polarimetric decomposition	
H	Entropy	Number of dominant scattering mechanisms. Ranges from 0 to 1.	Cloude and Pottier[49]
A	Anisotropy	Measures the relative importance of the second and third types of scattering.	
α	Alpha angle	Indicates the dominant scattering mechanism. Ranges from 0 deg to 90 deg.	
Ps	Surface scattering	Portion of surface scattering.	Freeman and Durden[50]
Pd	*Double-bounce* scattering	Modeled from orthogonal surfaces.	
Pv	Volumetric scattering	Modeled by a cloud of randomly distributed thin cylindrical spreaders.	
Ys	Surface scattering	Portion of surface scattering.	Yamaguchi et al.[51]
Yd	*Double-bounce* scattering	Modeled from orthogonal surfaces.	
Yv	Volumetric scattering	Modeled by a cloud of randomly distributed thin cylindrical spreaders.	
Y_H	Helix scattering	Helix scattering	
VZs	Surface scattering	Simulated surface type scattering from eigenvalues/eigenvectors decomposition	Van Zyl[52]
VZd	*Double-bounce* scattering	Double-bounce scattering simulated from eigenvalues/eigenvectors decomposition	
VZv	Volumetric scattering	Scattering of the simulated volumetric type from the decomposition of eigenvalues/eigenvectors	

2.5 *Modeling and Accuracy Assessment*

To analyze the potential of the predictor variables derived from each SAR sensor for estimating stem volume, four experiments were defined (Table 4). The first three experiments (i, ii, and iii) use only variables extracted from remote sensing data. To observe the contribution of field data in the models, the stand age was added to the previous experiments, resulting in experiment iv.

Table 4 Experiments to predict stem volume using Sentinel-1B, ALOS-2/PALSAR-2, and field data.

Experiment	Data source	Variables	Number of predictor variables
(i)	Sentinel-1B	Backscattering, polarimetric decomposition, and interferometric coherence	8
(ii)	ALOS-2/PALSAR-2	Backscattering, polarimetric decomposition, and interferometric coherence	26
(iii)	Sentinel-1B + ALOS-2/PALSAR-2	Sentinel-1B and ALOS-2/PALSAR-2 variables	34
(iv)	Sentinel-1B + ALOS-2/PALSAR-2 + forest inventory	Sentinel-1B variables + ALOS-2/PALSAR-2 variables + age	35

For each experiment, both specific models of *Eucalyptus* sp. and *Pinus* sp., and a generic model considering both *Eucalyptus* sp. and *Pinus* sp. data were developed.

The data extraction from the predictor variables was performed by calculating the mean of the pixel values located within the polygons of the plots measured in the field, for each one of the predictor variables.

Three nonparametric machine learning algorithms were tested to estimate the stem volume of forest plantation: RF, SVR, and XGB. RF algorithm was proposed as an ensemble method of decision trees by Breiman.[27] The decision trees are created independently, through a subset of training samples. The combination of decision trees reduces the error in regression tasks thanks to the use of bootstrap aggregation or bagging.[53] There are two main parameters to be tuned, Mtry and Ntree. Mtry is the number of prediction variables used in each node and Ntree is the number of regression trees. The ensemble prediction for regression problems given by a forest is obtained as an average of the prediction results of the individual trees.[54] Further, the RF can calculate the relative importance of each predictor variable in the model performance.[28]

The SVM was first developed to solve problems related to classification.[55] However, it can also be applied to regression problems, being referred to as SVR. This method allows solving nonlinear problems with linear solutions by developing a kernel function to project the feature space into high-dimensional space. The kernel functions include linear, polynomial, sigmoid, and radial basis function (RBF). The algorithm controls the model's complexity from two parameters: cost (C) and gamma (g). The higher the cost value, the lower the error tolerance, which may create a model that does not generalize correctly.[56] The "g" parameter influences the separation form of the support vectors.[57]

The XGB[30] develops strong models through an additive training process. This algorithm is based on the augmentation of decision trees, which uses an efficient second-order expression. This model is generalizable and avoids overfitting the data as well as underfitting it. The main parameters to be tuned during the modeling process are learning rate, number of trees, minimum number of observations in terminal nodes, tree depth, and gamma, which represents the minimum loss of additional partitioning of a node in the tree.[58]

For each model, the optimized parameters were determined using the cross-validation K-fold method ($k = 5$). In the fivefold cross-validation, the training data was separated into 5 subsets, with each subset treated as the validation and the other four subsets used as training samples. This process was repeated until all the subsets had been used.[28] The total field sampling plots (220) were divided randomly into 70% for training and 30% for validation of the specific and generic models (Table 5).

Modeling and accuracy assessments were conducted using R open-source software,[59] using the following packages: *rgdal, e1071, randomForest, xgboost, caret*, *mlr*, *dplyr,* and *ggplot2*. The accuracies of predicted stem volume were evaluated using the root-mean-square error (RMSE) and the coefficient of determination (R^2) to compare the prediction performance of the different experiments (Table 4) and the tree machine learning algorithms.

Table 5 Sampling plots for each model used for training and testing.

Models	Training	Testing
Eucalyptus sp.	110	46
Pinus sp.	45	19
Generic	154	66

3 Results

Table 6 shows the RMSE (%) and R^2 (%) values obtained with the RF, SVR, and XGB models for *Eucalyptus* sp. and *Pinus* sp. specific models and for the generic model for stem volume estimation.

For the *Eucalyptus* sp. specific models, the RF algorithm showed the best result for experiment (iv), with 89.89% of R^2. On the other hand, SVR showed best performance compared with other algorithms with experiment (ii) ($R^2 = 13.44\%$). The XGB showed similar performance to RF in terms of R^2; however, RF presented low RMSE (%) results. Comparing the performance of using Sentinel-1 and ALOS-2/PALSAR-2 data combined or individually, the combination of Sentinel-1 and ALOS-2/PALSAR-2 in experiment (iii) improved the results compared with when these data were used individually, with an increase of 1.10% and 5.46% in R^2 for experiments (i) and (ii), respectively, considering the RF model. Experiment (iv) also combined stand age with SAR data, which shows a great improvement in the results [an increase of 74.80% in terms of R^2 in comparison with experiment (iii) for the RF model].

Table 6 Stem volume estimation with SAR variables for *Eucalyptus* sp. and *Pinus* sp. using the specific models and the generic model.

	RF		SVR		XGB	
Experiments	RMSE %	R^2 %	RMSE %	R^2 %	RMSE %	R^2 %
			Eucalyptus sp. Models			
(i)	52.34	13.99	53.8	13.03	53.70	9.56
(ii)	55.10	9.63	56.2	13.44	57.22	9.55
(iii)	52.90	15.09	57.4	11.68	56.33	15.27
(iv)	17.89	89.89	21.4	85.79	18.94	89.52
			Pinus sp. models			
(i)	40.25	28.98	43.25	13.01	37.80	62.56
(ii)	39.19	28.92	45.20	7.28	39.92	25.95
(iii)	39.51	29.24	45.56	9.64	41.86	23.38
(iv)	33.89	46.94	27.81	64.22	32.06	51.81
			Generic models			
(i)	48.06	15.35	54.12	3.32	48.95	11.68
(ii)	42.90	32.35	46.27	19.92	42.96	30.40
(iii)	42.48	33.56	47.25	18.55	41.79	34.68
(iv)	21.99	81.79	28.99	68.53	22.00	81.81

Considering the *Pinus* sp. specific models, the experiment (i) with Sentinel-1 variables shows better results with SVR and XGB models in comparison with experiments (ii) and (iii). The use of ALOS-2/PALSAR-2 data does not show expressive improvement in the stem volume estimation, resulting in a worse performance of SVR and XGB models in the experiments (ii) and (iii). The addition of stand age (iv) improved the results using the RF and SVR models. The best specific model for *Pinus* sp. was obtained in the experiment (iv) using the SVR algorithm, with $R^2 = 64.2\%$.

The generic model, which considers *Eucalyptus* sp. and *Pinus* sp. data together, showed an improvement using ALOS-2/PALSAR-2 variables (ii), which represents an increase of 17% in terms of R^2 in comparison with using just Sentinel-1 variables (i) in the RF model. The experiment (iii), using Sentinel-1 and ALOS-2/PALSAR-2 variables, do not show significant improvement when compared with the use of only ALOS-2/PALSAR-2 variables (ii). The use of stand age, in experiment (iv), played an important role in the generic models of stem volume estimation, improving the results of all algorithms, with an increase of 48.23% in terms of R^2 compared with experiment (iii) for the RF model.

The scatterplots of observed versus predicted stem volume for the experiment (iv) and the tree machine learning algorithms for each prediction model are illustrated in Fig. 4. The RF and XGB algorithms presented similar results of R^2 (%) and RMSE (%) for the *Eucalyptus* sp. and

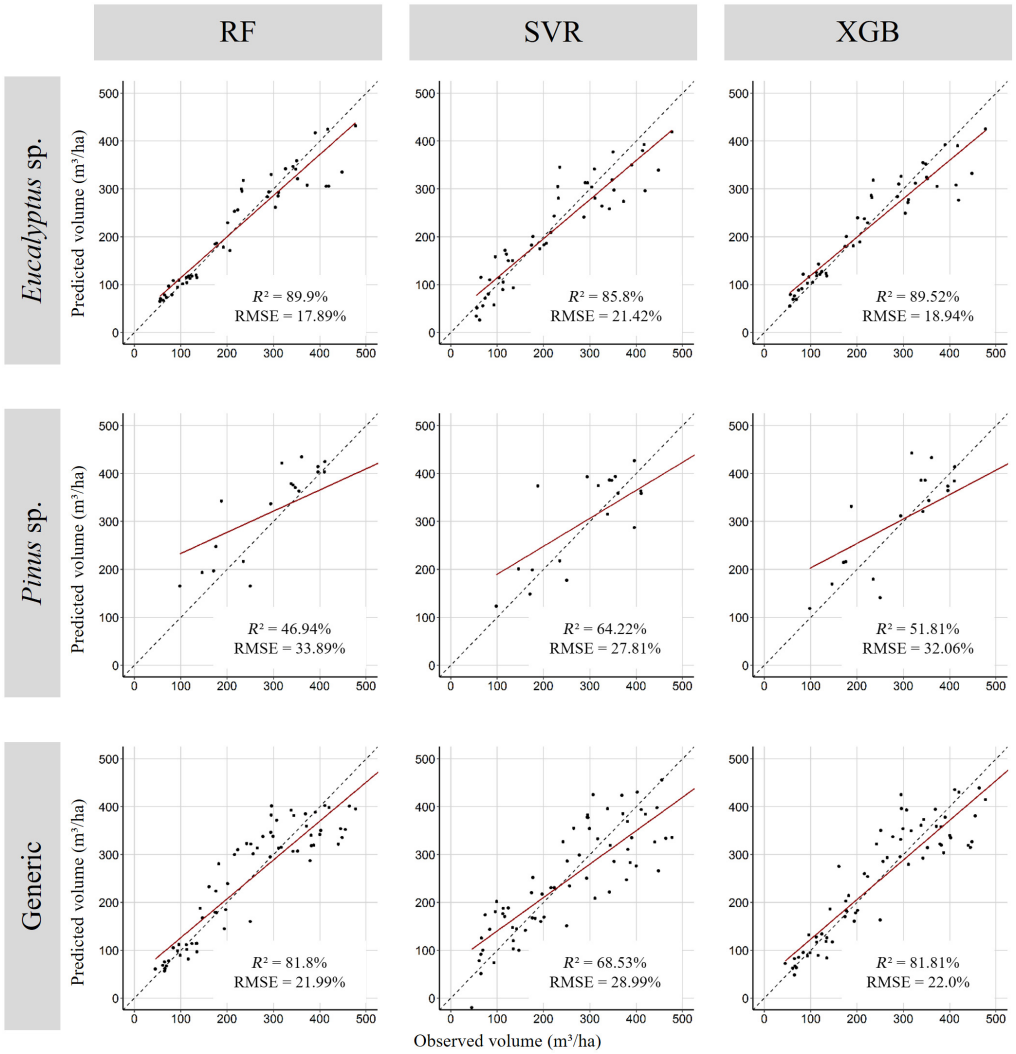

Fig. 4 Scatterplots of observed *versus* predicted stem volume produced by experiment (iv) using the tree machine learning algorithms: RF, SVR, and XGB, considering the specific and generic models.

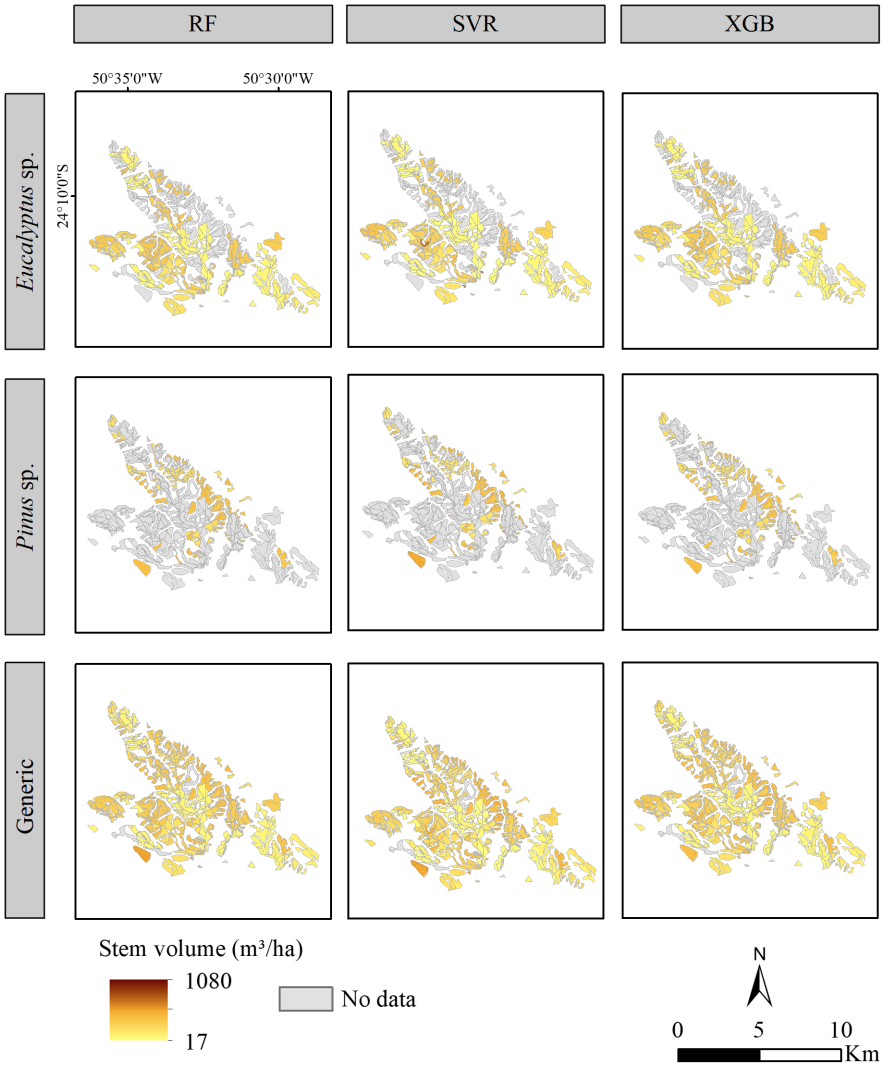

Fig. 5 Predicted stem volume (m^3/ha) for study area 1, considering the specific and generic models for RF, SVR, and XGB algorithms.

the generic models. However, observing the scatterplots, the RF model shows a better fitting of the data in comparison with the XGB algorithm. For the *Pinus* sp. models, the SVR models show a better fitting of the data, which corroborates the results of R^2 and RMSE.

The estimation maps obtained using the tree machine learning algorithms for each model are illustrated in Figs. 5 and 6, for the study area 1 and 2, respectively. Visually the maps generated by the machine learning algorithms show similar results.

For the *Eucalyptus* sp. model, we observed that RF and XGB algorithms resulted in similar predicted stem volume values. Analyzing the statistics values (Table 7 and Fig. 7), the minimum and maximum values are very close (ranging from 33.86 to 476.58 m^3/ha with RF, and 35.05 to 456.01 m^3/ha with XGB), and agree with the field data, which ranges from 15.00 to 432.00 m^3/ha. However, the SVR algorithm showed a different behavior. The minimum and maximum values are very distant from the values observed in the field. The predicted minimum values were too close to zero, and the maximum values reached 1079.43 m^3/ha, which is not consistent with the real behavior of the stem volume for the study area.

The *Pinus* sp. model predicted higher stem volume values compared with the *Eucalyptus* sp. model. In this case, RF and XGB also generated similar results (ranging from 133.52 to 509.47 m^3/ha with RF, and 103.48 to 549.52 m^3/ha with XGB – Table 7 and Fig. 7) The SVR algorithm also generated different prediction results, with the minimum value close to zero and the maximum value of 865.27 m^3/ha, which is more similar to the field data.

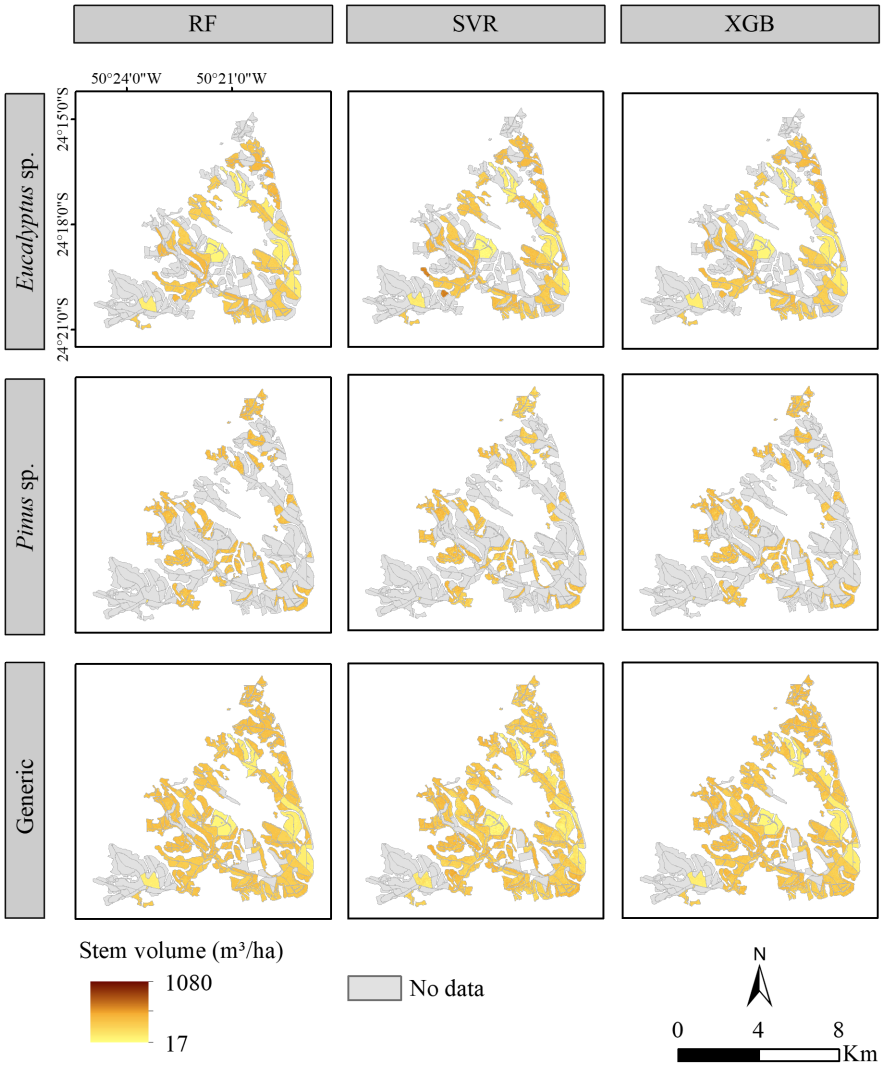

Fig. 6 Predicted stem volume (m^3/ha) for study area 2, considering the specific and generic models for RF, SVR, and XGB algorithms.

Table 7 Basic statistics of predicted stem volume considering the specific and generic models for RF, SVR, and XGB algorithms.

	RF			SVR			XGB			Field Data		
Models	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
Eucalyptus sp.	33.86	476.58	222.27	0.00	1079.43	231.76	35.05	456.01	215.18	15.00	432.00	194.90
Pinus sp.	133.52	509.47	298.95	0.18	865.27	328.72	103.48	549.52	291.18	98.00	791.00	323.67
Generic	41.01	600.49	263.45	0.00	744.86	266.17	17.44	563.18	243.27	15.00	791.00	265.08

The generic model showed a similar behavior compared with the specific models, with an improvement of the homogeneity of the predicted stem volume in the forest stands, and a reduction of salt-and-pepper effect in the predicted stem volume maps. The RF predicted stem volume values ranging from 41.01 to 600.49 m^3/ha. The SVR generic model also generated minimum values close to zero, and the maximum value of 744.86 m^3/ha, which was the closest predicted value to the maximum values observed in the field data. The XGB predictions range from 17.44 to 563.18 m^3/ha.

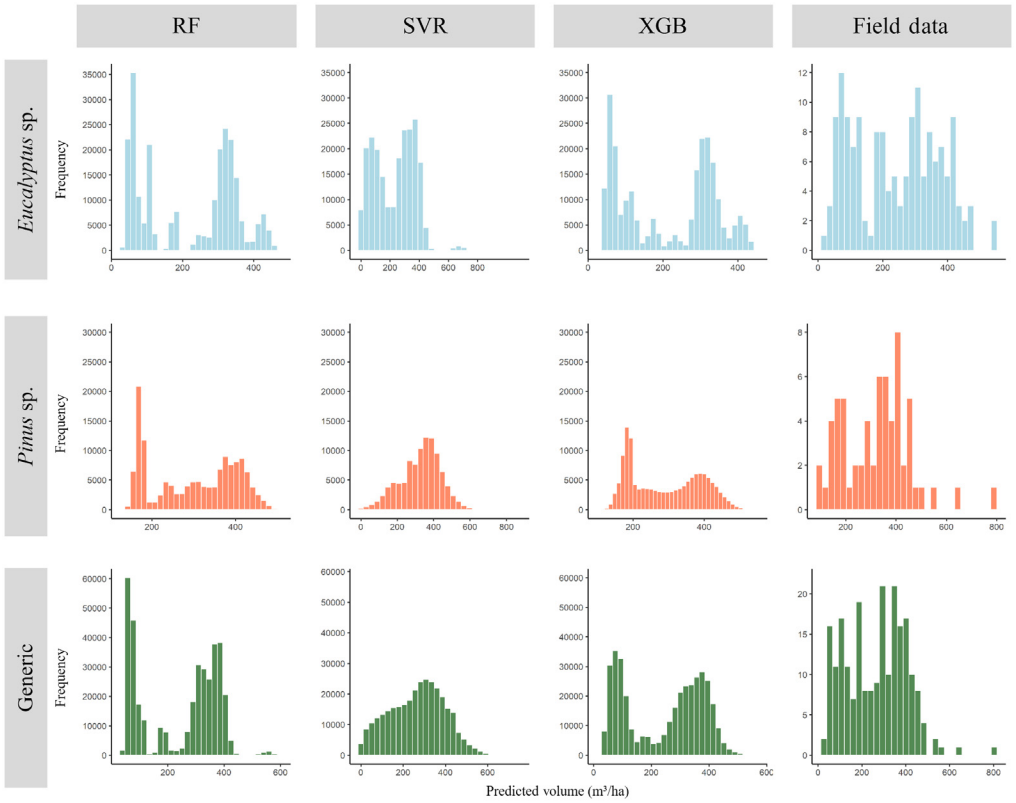

Fig. 7 Histogram of observed and predicted stem volume considering the specific and generic models for RF, SVR, and XGB algorithms.

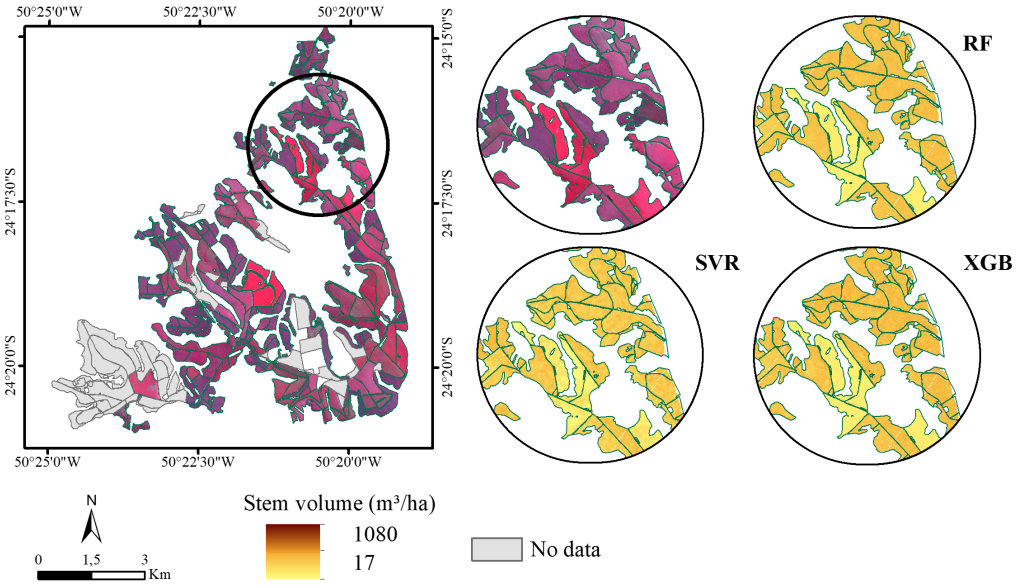

Fig. 8 Highlighted stands in study area 2 comparing the experiment (iv) with each machine learning algorithm using the generic model, with an optical Sentinel-2 image of April 18, 2019 (RGB: 843).

Figure 8 highlighted a subset of the Study Area 2 where there are significant stem volume differences in the stands. The generic model generated different stem volume predictions when each machine learning algorithm was used. The RF was the algorithm that generated more homogeneity in the predicted values, reducing the salt-and-pepper effect in the predicted stem

volume map. On the other hand, the SVR and XGB algorithms generated stem volume predictions with a greater salt-and-pepper effect within the stands. The same pattern was also observed for the stem volume predictions generated by the three machine learning algorithms for the Study Area 1.

4 Discussion

In the present work, we investigated the integration of multifrequency SAR data with machine learning approaches for retrieving stem volume of forest plantations in Brazil at the stand-scale. Although ALOS-2/PALSAR-2 allowed the extraction of a greater number of predictor variables in comparison with Sentinel-1B due the limitation of polarizations in the Sentinel-1B data, our results did not show significant improvement in accuracy of stem volume estimation using only ALOS-2/PALSAR-2 variables [Experiment (ii)] when compared with using just Sentinel-1B variables [Experiment (i)], mainly for the specific models. However, the generic models showed an improvement in comparison with using just Sentinel-1B variables [Experiment (i)] when ALOS-2/PALSAR-2 variables were used [Experiment (ii)] and when combining Sentinel-1B with ALOS-2/PALSAR-2 variables [Experiment (iii)]. In this case, we can infer that the predictor variables derived from individual SAR sensors were better able to predict stem volume in homogeneous forest (specific models) than for a heterogeneous area with stem volume of *Eucalyptus* sp. and *Pinus* sp. (generic model), which shows higher stem variability.

The predictor variables derived from Sentinel-1B data did not result in enhanced results for the generic model. The C-band backscattering saturates at low biomass levels (30 to 40 T/ha), while L-band data show biomass saturation levels between 40 and 150 T/ha.[60–62] Santoro et al.[17] studied the relationship between SAR frequency and stem volume, with a decrease of RMSE for stem volume estimation when longer wavelengths were used. Other study carried out by Stelmaszczuk-Górska et al.[63] showed that the use of multifrequency SAR data, using ALOS-2/PALSAR-2 band-L data and RADARSAT-2 band-C, improved the results for aboveground biomass estimation, when compared with the results using ALOS-PALSAR and RADARSAT-2 data separately. Therefore, combining SAR C-band and L-band can improve the prediction performance of the machine learning models compared with the use of single sensor data, as observed in experiment (iii) with the generic model. Consequently, the stem volume estimates using the combination of the two SAR frequencies achieved a superior accuracy when compared with using just one single SAR frequency.

In general, the R^2 (%) and RMSE (%) of the machine learning models showed better results for the *Pinus* sp. models than for the *Eucalyptus* sp. models when just SAR variables were used [Experiments (i), (ii), and (iii)]. These results can be related to the number of species in the *Eucalyptus* sp. plantations (five different species) while *Pinus* sp. plantations had just two species. The different species could have caused differences in spectral responses of the tree canopies associated with the variation in the growth rate of each species. This fact was also observed by Blanco et al.[64] estimating biomass in *Eucalyptus* sp. (*E. grandis*) and *Pinus* sp. (*P. oocarpa*, *P. patula,* and *P. tecunumanii*) plantations. The authors observed that the RMSE values for *E. grandis* (42.96 T/ha) was smaller than for *Pinus* sp. (150.94 T/ha), which they related to the fact that *Pinus* sp. have three species with different spectral responses.

The experiment (iv) showed that the use of stand age as a predictor variable improved the strength of the models and the accuracy of stem volume estimates. In this case, the *Eucalyptus* sp. models showed better performance in comparison with *Pinus* sp. models. Reis et al.[4] also observed that the use of stand age as a predictor variable in stem volume estimation of *Eucalyptus* sp. (*E. urophylla* × *E. grandis*) had improved the accuracy of the prediction models when used together with remote sensing variables with an R^2 of 0.71 in comparison with the prediction model using just Sentinel-1 variables, which have an R^2 of 0.34.

When analyzing the machine learning algorithms, it is important to observe how they performed considering the field data (Fig. 7 and Table 7). The RF and XGB algorithms generated similar predictions, with predicted values close to the values of field data for *Eucalyptus* sp. stands. On the other hand, for *Pinus* sp. stands, these algorithms generated predicted maximum values below those observed in the field data. In this case, saturation may be occurring in the

predicted stem volume values for *Pinus* sp. stands. Long et al.[65] summarized and exemplified some studies that demonstrate the problem with saturation of SAR data for the estimation of stem volume in Chinese Fir plantation. Saturation happens when the polarimetric properties of the SAR data are no longer sensitive to forest changes.[66–68] In contrast, the SVR algorithm generated the maximum predicted values closest to the field data for *Pinus* sp. stands. However, the SVR algorithm resulted in a larger range of predicted values for all models, leading to an overfitting and underfitting of the predicted values.

The machine learning algorithms can show different results according to the data used for training the models and study area. Souza et al.[21] compared the results of SVR, MLR, ANN, and RF for predicting *Eucalyptus* stem volume using ALOS AVNIR-2 and ALOS PALSAR and observed that SVR, ANN, and RF models have similar accuracy, with 0.92, 0.91, and 0.90 of R^2 values, respectively, where the SVR showed the best prediction performance. Pham et al.[58] compared five machine learning algorithms: XGB, gradient boosting regression (GBR), SVR, RF, and Gaussian process regression (GPR) to estimate mangrove aboveground biomass using ALOS-2/PALSAR-2 and Sentinel-2 data. The authors observed that XGB generated the best prediction result, with a $R^2 = 0.805$.

After analyzing all our results, we realized that choosing an algorithm to predict stem volume is not an easy task, due to the subtle results differences. In turn, the RF algorithm requires the tuning of just two main hyperparameters and has an easier implementation in comparison with the other two algorithms. Then, the RF is an algorithm that can be used for predicted stem volume with significant good results.

The approaches used in this study provide a framework for integrating multifrequency SAR data and machine learning techniques, highlighting methods that greatly improve the spatial prediction of stem volume in forest plantations of *Eucalyptus* sp. and *Pinus* sp. While this study uses a relatively small number of field sample plots for training the machine learning models, collecting a large amount of field data to develop prediction models generated based on the relationship between field-based measurements and variables derived from multifrequency SAR data is one of the main limitations of the approaches proposed in this study. Another potential limitation is related to data accessibility. ALOS-2/PALSAR-2 is commercial satellite data, and therefore, ALOS-2/PALSAR-2 data are not open access as ESA's Sentinel-1 data. Besides that, the number of polarizations available for Sentinel-1 data can be a limitation. Dual pol data limits the calculations of polarimetric decompositions, which will decrease the number of predictor variables derived from Sentinel-1 data.

5 Conclusions

We demonstrate how multifrequency SAR variables can be used to estimate the stem volume of *Eucalyptus* sp. and *Pinus* sp. using specific models and a generic model. Using multifrequency SAR predictor variables resulted in moderately high accuracy of the stem volume prediction models, mainly when SAR data were used in combination with stand age [Experiment (iv)].

The three machine learning algorithms tested in our study were capable of estimating stem volume with similar results. However, the RF algorithm generated more homogeneous predictions for the entire area of the forest plantations as highlighted in the prediction maps. Moreover, the RF models showed an easier implementation compared with the other two algorithms (SVR and XGB).

The generic models generated similar results compared with the specific models in terms of predicted stem volume maps. Our results corroborate the potential of using multifrequency SAR data to improve the stem volume estimation accuracy in forest plantations with different species.

Based on this research, future works can be developed for reducing inventory costs as well as the frequency of field campaigns. In addition, future works can explore the combination of using SAR and optical variables to improve the estimation accuracy of the prediction models.

Acknowledgments

The authors are grateful for the support of VISIONA TECNOLOGIA ESPACIAL S.A. and FAPESP (Project 2017/25639-5) for providing the ALOS-2/PALSAR-2 images, and KLABIN

S.A. for providing the field data for the research. The first author is also grateful to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Finance Code 001 by the grant received during the investigation.

Code, Data, and Materials Availability

The data that support the findings of this study were available from the Klabin S.A. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the authors with the permission of Klabin S.A.

References

1. I. Ali et al., "Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data," *Remote Sens.* **7**, 16398–16421 (2015).
2. R. Ismail et al., "Assessing the utility of ALOS PALSAR and SPOT 4 to predict timber volumes in even-aged Eucalyptus plantations located in Zululand, South Africa," *South. For.: J. For. Sci.* **77**(3), 203–211 (2015).
3. R. Torres et al., "GMES sentinel-1 mission," *Remote Sens. Environ.* **120**, 9–24 (2012).
4. A. A. Reis et al., "Volume estimation in a Eucalyptus plantation using multi-source remote sensing and digital terrain data: a case study in Minas Gerais State, Brazil," *Int. J. Remote Sens.* **40**(7), 2683–2702 (2019).
5. J. F. Dallemand et al., *Radar Imagery: Theory and Interpretation, Lecture Notes*, Rome (1993).
6. F. M. Henderson and A. J. Lewis, "Radar detection of wetland ecosystems: a review," *Int. J. Remote Sens.* **29**(20), 5809–5835 (2008).
7. T. S. F. Silva et al., "Capturing the dynamics of Amazonian wetlands using synthetic aperture radar: lessons learned and future directions," in *Remote Sensing of Wetlands: Applications and Advances*,R. Tiner, M. Lang, and V. Klemas, Eds., pp. 453–470, CRC Press, Boca Raton, Florida (2015).
8. R. M. Green, "The sensitivity of SAR backscatter to forest windthrow gaps," *Int. J. Remote Sens.* **19**, 2419–2425 (1998).
9. L. Naidoo et al., "Savannah woody structure modelling and mapping using multi-frequency (X-, C- and L-band) Synthetic Aperture Radar data," *ISPRS J. Photogramm. Remote Sens.* **105**, 234–250 (2015).
10. Y. Rauste, "Radar-based forest biomass estimation," *Int. J. Remote Sens.* **15**(14), 2797–2808 (1994).
11. I. H. Woodhouse, *Introduction to Microwave Remote Sensing*, pp. 1–370, CRC Press, Boca Raton, Florida (2006).
12. M. C. Dobson et al., "Dependence of radar backscatter on coniferous forest biomass," *IEEE Trans. Geosci. Remote Sens.* **30**, 412–415 (1992).
13. F. F. Gama, J. R. Dos Santos, and J. C. Mura, "Eucalyptus biomass and volume estimation using interferometric and polarimetric SAR data," *Remote Sens.* **2**(4), 939–956 (2010).
14. S. Englhart, V. Keuck, and F. Siegert, "Aboveground biomass retrieval in tropical forests—the potential of combined X- and L-band SAR data use," *Remote Sens. Environ.* **115**, 1260–1271 (2011).
15. F. F. Gama, J. R. Dos Santos, and J. C. Mura, "Continuous monitoring of biophysical Eucalyptus sp. parameters using interferometric synthetic aperture radar data in P and X bands," *J. Appl. Remote Sens.* **10**(2), 026002 (2016).
16. L. O. Pereira et al., "Multifrequency and full-polarimetric SAR assessment for estimating above ground biomass and Leaf Area Index in the Amazon Várzea Wetlands," *Remote Sens.* **10**(9), 1355 (2018).
17. M. Santoro et al., "Complementarity of X-, C-, and L-band SAR backscatter observations to retrieve forest stem volume in boreal forest," *Remote Sens.* **11**, 1563 (2019).
18. V. Nizalapur, C. S. Jha, and R. Madugundu, "Estimation of above ground biomass in Indian tropical forested area using multi-frequency DLR-ESAR data," *Int. J. Geom. Geosci.* **1**(2), 167–178 (2010).

19. S. Abdullahi, F. Kugler, and H. Pretzsch, "Prediction of stem volume in complex temperate forest stands using TanDEM-X SAR data," *Remote Sens. Environ.* **174**, 197–211 (2016).
20. J. Dong et al., "Mapping tropical forests and rubber plantations in complex landscapes by integrating PALSAR and MODIS imagery," *ISPRS J. Photogramm. Remote Sens.* **74**, 20–33 (2012).
21. G. S. A. Souza et al., "Multi-sensor prediction of Eucalyptus stand volume: a support vector approach," *ISPRS J. Photogramm. Remote Sens.* **156**, 135–146 (2019).
22. M. Belgiu and L. Dragut, "Random forest in remote sensing: a review of applications and future directions," *ISPRS J. Photogramm. Remote Sens.* **114**, 24–31 (2016).
23. A. A. Reis et al., "Spatial prediction of basal area and volume in Eucalyptus stands using Landsat TM data: an assessment of prediction methods," *N. Z. J. For. Sci.* **48**(1), 1–17 (2018).
24. B. H. Trisasongko and D. J. Paull, "L-band SAR for estimating aboveground biomass of rubber plantation in Java Island, Indonesia," *Geocarto Int.* **35**(12), 1327–1342 (2020).
25. T. D. Pham et al., "Estimating mangrove above-ground biomass using extreme gradient boosting decision trees algorithm with fused sentinel-2 and ALOS-2 PALSAR-2 data in Can Gio Biosphere Reserve, Vietnam," *Remote Sens.* **12**(5), 777 (2020).
26. A. A. Dos Reis et al., "Monitoring pasture aboveground biomass and canopy height in an integrated crop–livestock system using textural information from PlanetScope imagery," *Remote Sens.* **12**(2534), 1–21 (2020).
27. L. Breiman, "Random forests," *Mach. Learn.* **45**(1), 5–32 (2001).
28. C. Wu et al., "Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery," *J. Appl. Remote Sens.* **10**(3), 035010 (2016).
29. M. S. Ataee et al., "Improving estimation accuracy of growing stock by multi-frequency SAR and multi-spectral data over Iran's heterogeneously-structured broadleaf Hyrcanian forests," *Forests* **10**(641), 1–18 (2019).
30. T. Chen and C. Guestrin, "XGBoost: a scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 13–17 August 2016, San Francisco, California, pp. 785–794 (2016).
31. Y. Sun et al., "Evaluating urban heat island intensity and its associated determinants of towns and cities continuum in the Yangtze River Delta Urban Agglomerations," *Sustain. Cities Soc.* **50**, 101659 (2019).
32. C. A. Alvares, "Köppen's climate classification map for Brazil," *Meteorol. Z.* **22**(6), 711–728 (2013).
33. G. T. Trewartha and L. H. Horn, *An Introduction to Climate*, 5th ed., McGrawHill, New York, 416 p. (1980).
34. C. E. Barbosa et al., "Diversity of regenerating plants in reforestations with Araucaria angustifolia (Bertol.) O. Kuntze of 12, 22, 35, and 43 years of age in Paraná State, Brazil," *Restoration Ecol.* **17**(1), 60–67 (2007).
35. V. T. Silva et al., "Comparação entre parâmetros abióticos e a estrutura florestal de um fragmento de floresta e um reflorestamento abandonado de eucalipto (Eucalyptus saligna Smith) no parque ecológico da Klabin, Telêmaco Borba/PR," *Semina: Cienc, Biol, Saúde.* **31**(1), 3751 (2010).
36. European Space Agency ESA, "Sentinel-1 SAR technical guide," https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-1-sar (accessed 2020).
37. J. Lee et al., "Speckle reduction in multipolarization, multifrequency SAR imagery," *IEEE Trans. Geosci. Remote Sens.* **29**(4), 535–544 (1991).
38. J. W. Park, A. Korosov, and M. Babiker, "Efficient thermal noise removal of Sentinel-1 image and its impacts on sea ice applications," in *Proc. EGU Gen. Assembly Conf. Abstr.*, 23–28 April 2017, Vienna, vol. **19**, p. 12613 (2017).
39. F. Filipponi, "Sentinel-1 GRD preprocessing workflow," *Proceedings* **18**(11), 1–4 (2019).
40. S. R. Cloude, "The dual polarisation entropy/alpha decomposition," in *Proc. 3rd Int. Work Sci. Appl. SAR Polarim. Polarim. Interferom.*, pp. 1–6 (2007).
41. D. Small and A. Schubert, *Guide to ASAR Geocoding*, ESA, Paris (2008).
42. R. Nasirzadehdizaji et al., "Sensitivity analysis of multi-temporal sentinel-1 SAR parameters to crop height and canopy coverage," *Appl. Sci.* **9**(4), 655 (2019).

43. P. Gururaj, P. Umesh, and A. Shetty, "Assessment of spatial variation of soil moisture during maize growth cycle using SAR observations," *Proc. SPIE* **11149**, 1114916 (2019).
44. SARMAP, SARscape technical description, http://www.sarmap.ch/pdf/SARscapeTechnical.pdf (accessed 12 December 2021).
45. M. Shimada et al., "PALSAR radiometric and geometric calibration," *IEEE Trans. Geosci. Remote Sens.* **47**(12), 3915–3932 (2009).
46. K. O. Pope, J. M. Rey-Benayas, and J. F. Paris, "Radar remote sensing of forest and wetland ecosystems in the central American tropics," *Remote Sens. Environ.* **48**, 205–219 (1994).
47. F. M. Henderson and A. J. Lewis, *Principles and Applications of Imaging Radar, Manual of Remote Sensing*, 3rd ed., 866 p., John Wiley & Sons, New York (1998).
48. Y. Kim and J. J. van Zyl, "A time-series approach to estimate soil moisture using polarimetric radar data," *IEEE Trans. Geosci. Remote Sens.* **47**(8), 2519–2527 (2009).
49. S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.* **34**(2), 498–518 (1996).
50. A. Freeman and S. L. Durden, "A three-component scattering model for polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.* **36**(3), 963–973 (1998).
51. Y. Yamaguchi et al., "Four-component scattering model for polarimetric SAR image decomposition," *IEEE Trans. Geosci. Remote Sens.* **43**(8), 1699–1706 (2005).
52. J. J. Van Zyl, "Application of Cloude's target decomposition theorem to polarimetric imaging radar data," *Proc. SPIE* **1748** (1993).
53. E. Izquierdo-Verdiguier and R. Zurita-Milla, "An evaluation of guided regularized random forest for classification and regression tasks in remote sensing," *Int. J. Appl. Earth Obs. Geoinform.* **88**, 102051 (2020).
54. A. Criminisi, J. Shotton, and E. Konukoglu, "Decision forests: a unified framework for classification, regression, density estimation, manifold learning and semisupervised learning," *Found. Trends Comput. Graph. Vision* **7**(2–3), 81–227 (2012).
55. V. Vapnik, *The Nature of Statistical Learning Theory*, Springer (2000).
56. EXELIS. Harris Geospatial Solutions, "Support vector machine background," 2015, https://www.harrisgeospatial.com/docs/BackgroundSVM.html (accessed 15 June 2021).
57. W. Li and Q. Du, "Support vector machine with adaptive composite kernel for hyperspectral image classification," *Proc. SPIE* **9501**, 95010O (2015).
58. T. D. Pham et al., "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 (2020).
59. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna (2018).
60. M. L. Imhoff, "A theoretical analysis of the effect of structure on synthetic aperture radar backscatter and the remote sensing biomass," *IEEE Trans. Geosci. Remote Sens.* **33**, 341–351 (1995).
61. R. M. Lucas et al., "The potential of L-band SAR for quantifying mangrove characteristics and change: case studies from the tropics," *Aquat. Conserv.* **17**, 245–264 (2007).
62. E. T. A. Mitchard et al., "Using satellite radar backscatter to predict above-ground woody biomass: a consistent relationship across four different African landscapes," *Geophys. Res. Lett.* **36**, L2340 (2009).
63. M. Stelmaszczuk-Górska et al., "Retrieval of aboveground biomass using multi-frequency SAR," *Living Planet Symp.* **740**, 194 (2016).
64. A. L. T. Blanco, I. A. L. Salcedo, and N. R. Eraso, "Estimación de biomasa aérea de Eucalyptus grandis y Pinus spp usando imágenes Sentinel1A y Sentinel2A en Colombia," *Colombia Forestal.* **23**(1), 117–141 (2020).
65. J. Long et al., "Mapping growing stem volume of Chinese fir plantation using a saturation-based multivariate method and quad-polarimetric SAR images," *Remote Sens.* **11**(1872), 1–20 (2019).
66. T. A. Chowdhury et al., "Growing stock volume estimation from L-band ALOS PALSAR polarimetric coherence in Siberian forest," *Remote Sens. Environ.* **155**, 129–144 (2014).
67. N. Joshi et al., "Understanding 'saturation' of radar signals over forests," *Sci. Rep.* **7**(3505), 1–11 (2017).

68. C. Thiel and C. Schmullius, "Impact of tree species on magnitude of PALSAR interferometric coherence over Siberian forest at frozen and unfrozen conditions," *Remote Sens.* **6**, 1124–1136 (2014).

Juliana Maria Ferreira de Souza Diniz received her bachelor of forest engineering degree from University of Lavras, Brazil, in 2016 and received an MS degree in remote sensing from National Institute for Space Research (INPE), Brazil, in 2019. She is finishing the PhD in remote sensing from National Institute for Space Research (INPE). She has experience in SAR applications such as forest modeling estimations and land use and land cover mapping.

Fabio Furlan Gama graduated in electronic engineering from Fundação Valeparaibana de Ensino (1986). He received PhD in radar remote sensing application from National Institute for Space Research, in 2007. He currently works at National Institute for Space Research (INPE-Brazil) as senior researcher in the Earth Observation and Geoinformatics Division, with SAR applications, acting on the following subjects SAR, P-band, interferometry, differential interferometry techniques, biomass, and forest inventory.

Aliny Aparecida dos Reis received her PhD in forest management from University of Lavras, Brazil, in 2018. She worked as a visiting scholar at Trent University, Canada, from 2017 to 2018, and as a postdoctoral researcher at University of Campinas, Brazil, from 2019 to 2021. She is currently working as a remote sensing scientist at Farmers Edge Inc. Her research interests include remote sensing and artificial intelligence applications for forest and agricultural monitoring.

Cleber Gonzales de Oliveira received his bachelor of cartographic engineering degree from the Universidade Estadual Paulista Júlio de Mesquita Filho in 1997, his master's degree in remote sensing from the National Institute for Space Research (INPE) in 2005, and his PhD in remote sensing from the National Institute of Space Research in 2011. He is currently chief operating officer – COO at Visiona Tecnologia Espacial S.A.

Eduardo Resende Girardi Marques received his bachelor of forest engineering degree from ESALQ in 2014. Currently pursuing his master's degree in forest management at UFPR. He works in forest inventory and remote sensing at Klabin S.A. focused on statistical modeling applied in biomass predictive models; innovation and carbon inventory. He has previously worked as an intern at the Institute of Research and Forestry Studies and U.W. – School of Environmental and Forest Sciences.