



Estimation of aboveground biomass from PolSAR and PolInSAR using regression-based modelling techniques

Ritwika Mukhopadhyay, Shashi Kumar, Hossein Aghababaei & Anurag Kulshrestha

To cite this article: Ritwika Mukhopadhyay, Shashi Kumar, Hossein Aghababaei & Anurag Kulshrestha (2021): Estimation of aboveground biomass from PolSAR and PolInSAR using regression-based modelling techniques, Geocarto International, DOI: [10.1080/10106049.2021.1878289](https://doi.org/10.1080/10106049.2021.1878289)

To link to this article: <https://doi.org/10.1080/10106049.2021.1878289>



Published online: 05 Feb 2021.



Submit your article to this journal [↗](#)



Article views: 170



View related articles [↗](#)



View Crossmark data [↗](#)



Estimation of aboveground biomass from PolSAR and PolInSAR using regression-based modelling techniques

Ritwika Mukhopadhyay^{a,b} , Shashi Kumar^a , Hossein Aghababaei^b  and Anurag Kulshrestha^b 

^aIndian Institute of Remote Sensing (IIRS), Indian Space Research Organisation (ISRO), Dehradun, India; ^bFaculty of Geo-information Science and Earth Observation, University of Twente, Enschede, The Netherlands

ABSTRACT

In the field of forestry studies, microwave remote sensing has broad applications due to the penetration into the semi-transparent media. This feature is used for the estimation of biophysical parameters and monitoring of deforestation. Therefore, the estimation of biophysical parameters is essential for assessing carbon stock management. Hence, the aboveground biomass (AGB) using synthetic aperture radar (SAR) data is recognized as typical approaches in forest application. However, the integrated use of polarimetric (PolSAR) and interferometric (PolInSAR) data might be more efficient tools for AGB mapping. Accordingly, in this study with the integrated data, the efficiency of machine learning techniques including random forest regression (RFR) and multiple linear regression (MLR) model were assessed and compared for the prediction of AGB. The analyses were performed using an image pair of fully polarimetric Radarsat-2 C-band data set and the related field data of Malhan Forest Range, Dehradun Forest Division, which were collected using the systematic sampling technique. Particularly, the training and testing of the models were done using the field sample plots. The experimental results showed that the RFR algorithm provided a better prediction result of AGB than the MLR model. The correlation coefficient (R^2) and root-mean-square error (RMSE) for the RFR algorithm was estimated to be around 0.65 and 24.33 Mg/ha, respectively, while for the MLR model, R^2 and RMSE are estimated as 0.54 and 33.05 Mg/ha, respectively. Therefore, it was concluded that the prediction of AGB through the machine learning technique using PolSAR and PolInSAR data has a significant advantage for accurate estimation of the AGB.

ARTICLE HISTORY

Received 4 September 2020
Accepted 29 December 2020

KEYWORDS

Random forest regression (RFR); multiple linear regression (MLR); PolSAR; PolInSAR; aboveground biomass (AGB); ground-to-volume scattering ratio

Introduction

Forests have a significant contribution in maintaining a balance in the environment and also in providing tangible and intangible resources (FAO 2018). Forests sequester carbon in its trunks, branches, leaves and roots which accumulates to be the biomass of forests.

Therefore, the forest quality and health assessment is of significant concern and can be measured from biophysical parameters (Brandeis et al. 2006; Hogg et al. 2008; Adams 2012; Chi et al. 2016). Biomass is the essential biophysical parameter for indicating the health of the forest as it represents the potential amount of carbon stored in the trees (Brown 1997; Adams 2012). AGB can be estimated through destructive sampling methods (Khanna and Chaturvedi 1994; Husch et al. 2003; Kershaw et al. 2016), non-destructive sampling methods (Husch et al. 1972; Khanna and Chaturvedi 1994), remote-sensed data-like, microwave, Light Detection and Ranging (LiDAR) and optical data (Mangla et al. 2016; Kumar et al. 2018)

Microwave remote sensing is advantageous over other techniques as it can penetrate through cloud cover and can acquire data both in day and night time (Rencz and Ryerson 1999; Moreira et al. 2013). Synthetic Aperture Radar (SAR) sensors transmit microwave signals and receive the backscattered signal from the targets on the earth's surface (Richards 2009; Cloude 2010; Moreira et al. 2013). Polarimetric SAR (PolSAR) is an extended technique of SAR that retrieves the scattering information from each resolution cell with multi-polarization information, which helps in the characterization of the objects (Kumar et al. 2017; Tomar et al. 2019). PolSAR is sensitive to different parameters of objects such as shape, orientation, dielectric properties, as well as permittivity, material constituents of the target and also the ensemble average entropy, which can be used for measuring forest biophysical parameters (Rencz and Ryerson 1999; Papathanassiou and Cloude 2001; Krieger et al. 2005). The PolSAR scattering mechanisms contributed from each SAR resolution cell can be retrieved separately after decomposition (Shafai and Kumar 2020; ME and Kumar 2021). The scattering mechanisms are odd-bounce, volume, double-bounce, helix and wire scatterings (Zhang et al. 2008). There are various decomposition models such as three-component decomposition models (Freeman and Durden 1998; Van Zyl et al. 2011), Yamaguchi four-component decomposition model (Yamaguchi et al. 2005) and multiple component scattering model (Zhang et al. 2008). Polarimetric SAR Interferometry (PolInSAR) is a further extended technique of SAR. It uses the benefits of polarimetry and interferometry coherently (Kumar et al. 2017; Kumar et al. 2020). From the PolInSAR-based coherency matrices, both the polarimetric and interferometric properties of images can be retrieved separately (Papathanassiou and Cloude 2001). PolInSAR coherence values indicate the degree of correlation between the images acquired at different times or from different locations (Neumann 2009).

Previous works have been done for the estimation of biophysical parameters of the forest using X, C, L and P bands of the microwave range (Mette et al. 2004; Santoro et al. 2006; Garestier et al. 2009; Carreiras et al. 2013; Behera et al. 2016; Duncanson et al. 2020; Khati et al. 2020). The integration of models with remote sensing techniques for the estimation of tree height and aboveground biomass for the entire vegetation has shown huge applicability (Agrawal et al. 2019; Joshi and Kumar 2017a, 2017b; Kumar et al. 2017). The different types of models used are WCM (Attema and Ulaby 1978; Kumar et al. 2012; van Leeuwen 2014; Behera et al. 2016), IWCM (Santoro 2003; Kumar 2009; Chandola 2014) and EWCM (Agrawal 2015; Bharadwaj et al. 2015; Tomar 2015; Kumar et al. 2019). Mangla (2015) did a comparative analysis between RFR and MLR for the estimation of AGB using PolSAR data. From the study, it was concluded that AGB could be estimated with higher accuracy using the RFR algorithm compared with MLR. Neumann et al. (2010) in his study compared the efficiency of linear regression, support vector regression (SVR) and random forest (RF) algorithms for the estimation of AGB using PolInSAR data. In this study, linear regression was observed to predict AGB more accurately but further investigation on the efficiency of the RF algorithm was suggested.

The PolSAR- and PolInSAR-based retrieved image parameters, backscatter component and coherence values can be integrated with the field estimated parameters and can be, therefore, utilized to predict AGB using machine learning (ML) algorithms. ML is an artificial intelligence technique implementing supervised and unsupervised learning of the data (Shataee et al. 2012; Mangla 2015). The ML algorithms used for classification or regression are support vector machine (SVM), RF, rotation random forest (RoRF) and classification and regression tree (CART) (Vapnik 2000; Shataee et al. 2012; Mangla 2015; Trisasongko et al. 2019). The regression algorithms adopted for this work is the RFR followed by MLR.

The estimation of biophysical parameters of forest using PolSAR backscatter and PolInSAR coherences has proved to be efficient individually, as discussed above. Therefore, the main objective of this present work was to test and explore the integration of PolSAR backscatter parameters and PolInSAR coherences retrieved from fully polarimetric Radarsat-2 C-band data sets. These parameters were then used to estimate AGB of a forest stand using a ML algorithm, RFR and a conventional regression model, MLR. Finally, a comparative analysis was done between the two regression models on the basis of their efficiency to predict AGB. These two regression models were chosen in order to overcome the rigidity of conventional models (WCM, IWCM, EWCM) for biophysical parameter estimation. The second objective is to optimize the RFR algorithm with different train and test data subsets, followed by the validation and accuracy assessment.

The contribution of this work in this field is the coherent use of the PolSAR backscatter parameters and PolInSAR coherence values and field data for estimating AGB using regression algorithms, RFR and MLR. The developed model will enable us to identify the most sensitive parameters for the prediction of AGB.

Materials and methods

The work was carried out in several stages as explained in the following sections, and the summary of the methodology adopted in this study is presented in Figure 1.

Study area

The study area for this present work is the Malhan Forest Range, Dehradun Forest Division (30° 20' N and 77°49' E) located in the south-westernmost part of the Uttarakhand, India, as shown in Figure 2. Malhan Forest covers an area of 75.9 sq. Km approximately with a sub-tropical type of climate. The major forest types found in the region are Moist Bhabar Dun Sal Forest, Moist Shivalik Sal Forest, Mixed Deciduous Forest and Deciduous Scrubs (Champion and Seth 1968). The forest mainly comprised of homogenous vegetation with *Shorea robusta* being the dominant and the keystone species (Champion and Seth 1968), as shown in Figure 3a,b.

Field inventory

The field data used in this study were collected following the systematic sampling technique. Firstly, the entire study area was divided into regular grids of 500 sq. m, as represented by the red lines shown in Figure 4a, followed by the implementation of the systematic sampling technique for laying the sample plots. The sample plots were laid systematically within the regular grids at an equal interval. Therefore, the first sample plot location was chosen randomly, while the successive sample plots were laid systematically

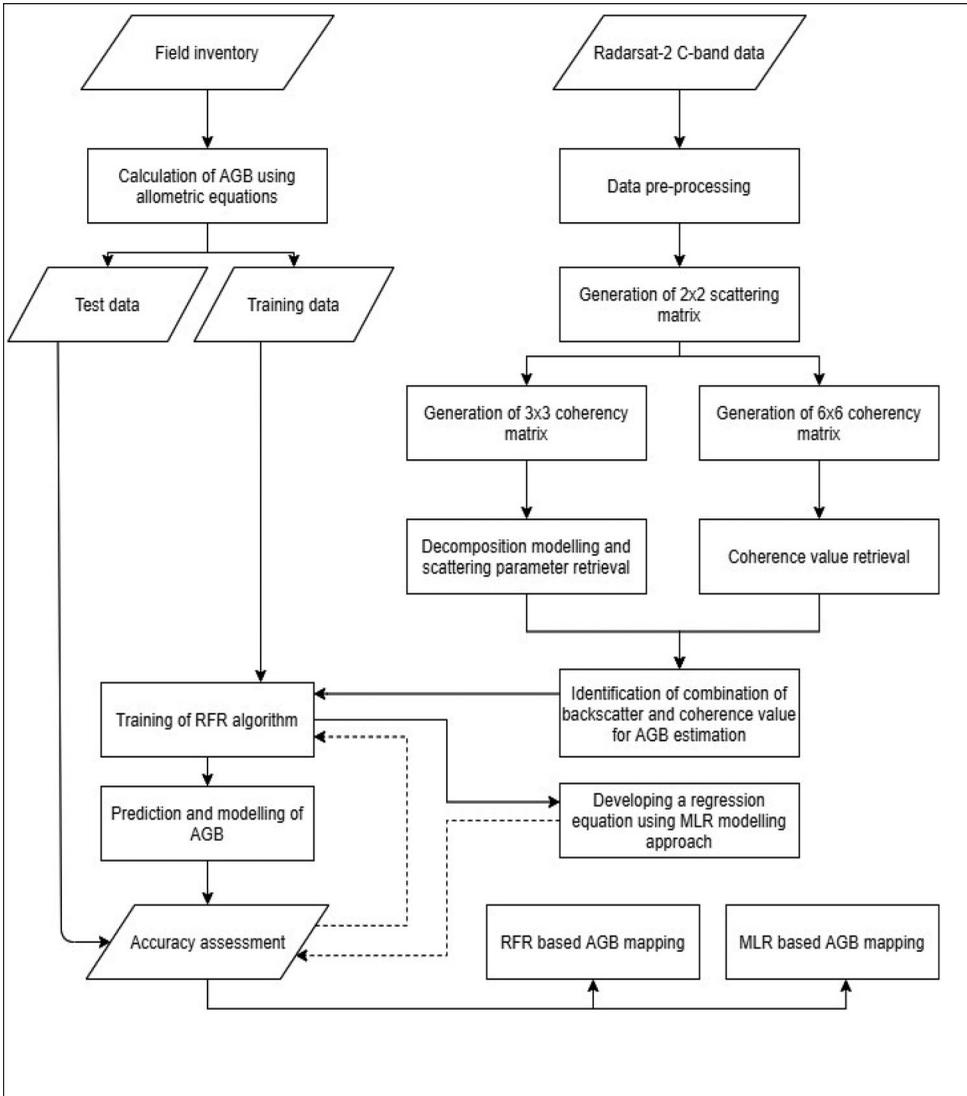


Figure 1. Flowchart of the methodology used in this study.

in both the directions, as represented by the red dots in Figure 4a. The sample plot locations were measured using an Real-time Kinematics (RTK) for recording the latitude and longitudes of the sample plots, as shown in Figure 5a,b. It consisted of a GNSS 10 receiver, GNSS antenna, a Leica GPS controller and a tripod. RTK provides higher positioning accuracy as compared with handheld GPS systems. The accuracy is as high as 2 to 3 cm. Each sample plot was square-shaped with dimensions of 31.62×31.62 sq. m covering an area of 0.1 ha, as shown in Figure 4b. A total of 38 sample plots were laid in the study area for field inventory, out of which one plot was observed to be an outlier during the data analysis. This sampling technique enables an unbiased representation of the study area (Khanna and Chaturvedi 1994; Husch et al. 2003).

Measurements of girth at breast height (gbh) over bark (ob) and height were taken for each tree in each plot. The gbh (ob) was measured using a flexible measuring tape of

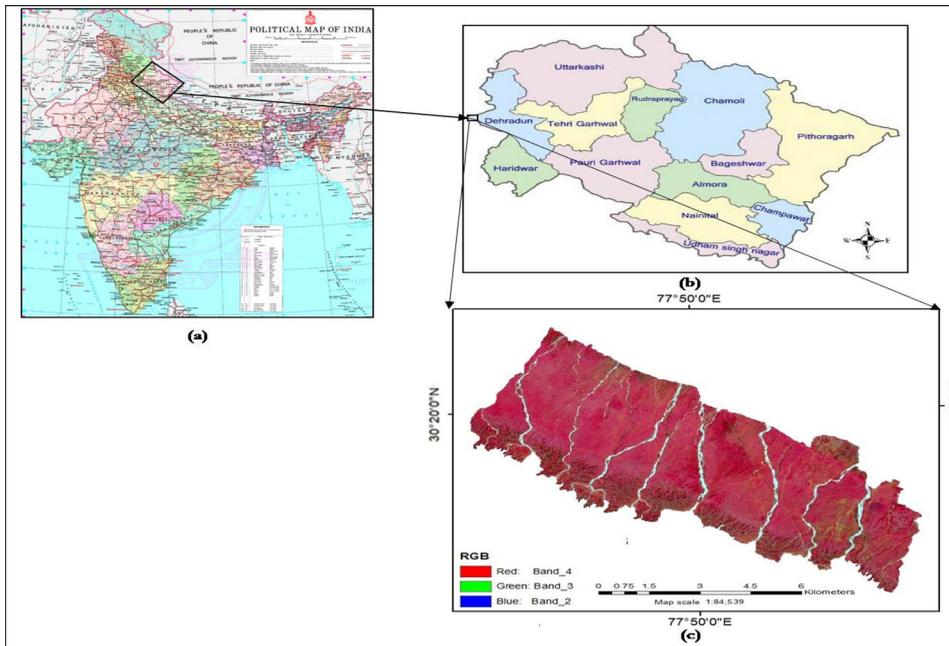


Figure 2. The study area: (a) Political map of India (Survey of India 2020), (b) District level map of Uttarakhand and (c) Malhan Forest Range.



Figure 3. (a)–(b) Field photograph of the Malhan forest range showing the dominant species as Sal (*Shorea robusta*).

30 m length, as shown in [Figure 6a,b](#). The height of trees was measured using a Nikon Forestry Pro hypsometer laser range-finder, as shown in [Figure 7](#). The accuracy of this instrument was ± 1 m. From the field measured gbh (ob), the diameter at breast height (dbh) was estimated by the formula represented in [Equation \(1\)](#), followed by the calculation of stem volumes. The stem volumes were calculated using the allometric equations for the particular species, as provided in [Table 1](#) (Forest Survey of India 1996). The distribution of the stem volumes of each sample plot is shown in [Figure 8](#). The stem volume was observed to range from 214.86 cu.m to 616.75 cu.m. From the stem volumes, the AGB of each tree was estimated using the formula represented in [Equation \(2\)](#), finally leading to the plot-level AGB estimation.

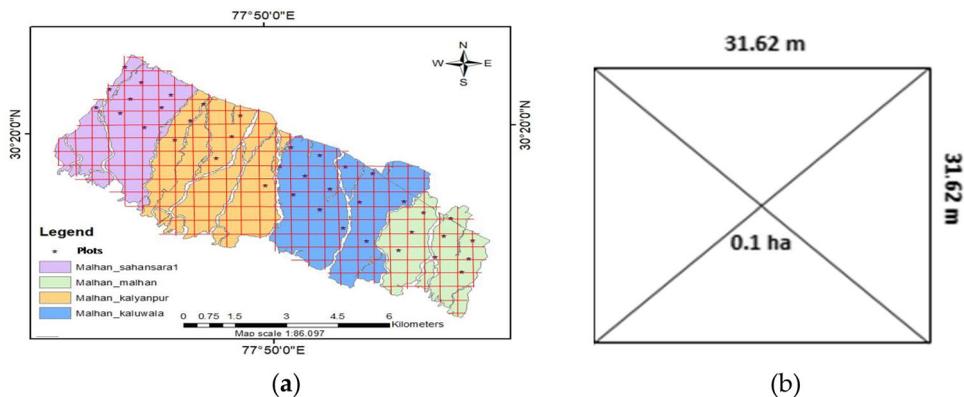


Figure 4. (a) Layout of the systematic sampling plots in field; (b) a sampling unit of 0.1 ha.



Figure 5. (a)–(b) RTK (Real-time Kinematics) being used for recording the latitude and longitudes of the sample plots.



Figure 6. (a)–(b) A flexible measuring tape of 30 m length being used in field to measure the tree gbh.

$$dbh = \frac{gbh}{\pi} \quad (1)$$

$$AGB = v \times g \quad (2)$$



Figure 7. Nikon Forestry Pro hypsometer laser range-finder being used in field for measuring the tree heights.

Table 1. Allometric equations provided by FSI.

Species	Volume equation	Specific gravity
<i>Shorea robusta</i>	$(0.16306 + 4.8991 * D - 1.57402 * (D \circ 0.5) \circ 2$	0.726
<i>Tectona grandis</i>	$0.08847 - 1.4693 * D + 11.98979 * (D) \circ 2 + 1.970560 * (D) \circ 3$	0.578
<i>Mallotus philippinensis</i>	$0.14749 - 2.87503 * D + 19.61977 * (D) \circ 2 - 19.11630 * (D) \circ 3$	0.571
<i>Erhetia laevis</i>	$((0.16609 / (D) \circ 2) - (2.78851 / D) + 17.22127 - (11.60248) * D) * (D) \circ 2$	0.571
<i>Lagerstromia parviflora</i>	$0.10529 - 1.68829 * D + 10.29573 * (D) \circ 2$	0.71
<i>Ficus religiosa</i>	$(0.03629 + 3.95389 * D - 0.84421 * D \circ 2) \circ 2$	0.523
<i>Bauhinia variegata</i>	$((-0.00342 / ((D) \circ 2)) - 0.0922 / D + 2.28178 + 9.46641 * D) * D \circ 2$	0.629
<i>Terminalia tomentosa</i>	$((0.18149 / D \circ 2) - (2.85865 / D) + 18.60799) * D \circ 2$	0.696
<i>Terminalia alata</i>	$0.08658 - 2.04096 * D + 13.28405 * D \circ 2 - 3.58047 * D \circ 3$	0.63

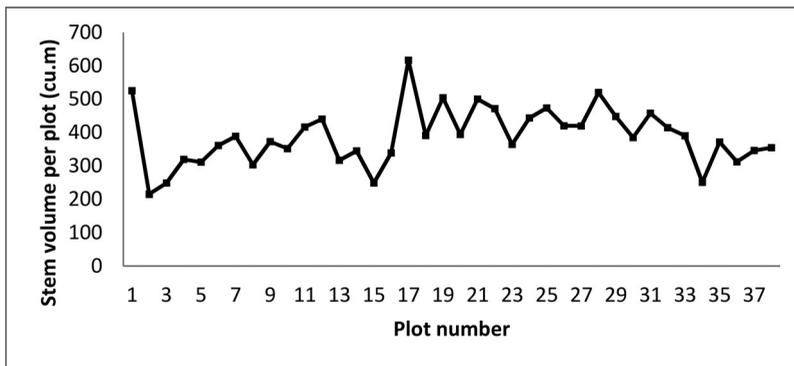


Figure 8. Graph representing the distribution of the stem volumes of 38 sample plots.

where v is stem volume and g represents the specific gravity. The specific gravity was obtained from the standard species specific values provided in Table 1 (Forest Survey of India 1996).

SAR data sets

The images used for this study were a pair of fully polarimetric Radarsat-2 C-band data, as described in Table 2. The images were acquired on 25 January 2019 and 18 February 2019 with a temporal baseline of 24 days, and the related Pauli images are shown in

Figure 9. The image pair was obtained from MDA Geospatial Services. The data are not freely available. The first acquisition has been considered as the master image, while the second acquisition as the slave image.

Backscatter parameter retrieval using PolSAR data

The Radarsat-2 data were first calibrated using the sigma nought look-up table to obtain radar backscatter coefficient values (σ^0). The pre-processing steps were done in polSARpro 5.0 software. The backscatter coefficient values were used to generate the scattering matrices [S], as shown in Equation (3). The scattering matrices were then vectorized in Pauli basis in order to generate the coherence matrix [T₃] for further polarimetric analysis, as shown in Equations (4) and (5).

$$[S] = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \quad (3)$$

The diagonal elements in the [S] matrix are the co-polarized wave channel amplitudes, whereas the off-diagonal elements are the cross-polarized wave channel amplitudes.

$$K_p = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ 2S_{HV} \end{bmatrix} \quad (4)$$

where K_p denotes the Pauli feature basis vector. The 3×3 coherence matrix [T₃] is generated as shown below;

$$[T] = \langle K_p \cdot K_p^\dagger \rangle \quad (5)$$

Table 2. Description of the Radarsat-2 image pair.

Description	1st acquisition	2nd acquisition
Satellite	Radarsat-2	Radarsat-2
Date of acquisition	25-01-2019	18-02-2019
Pass	Descending	Descending
Product Id	PDS_07004040	PDS_07004050
Acquisition mode	Fine quad polarisation	Fine quad polarisation
Range Resolution	4.733 m	4.733 m
Azimuth resolution	5.486 m	5.486 m

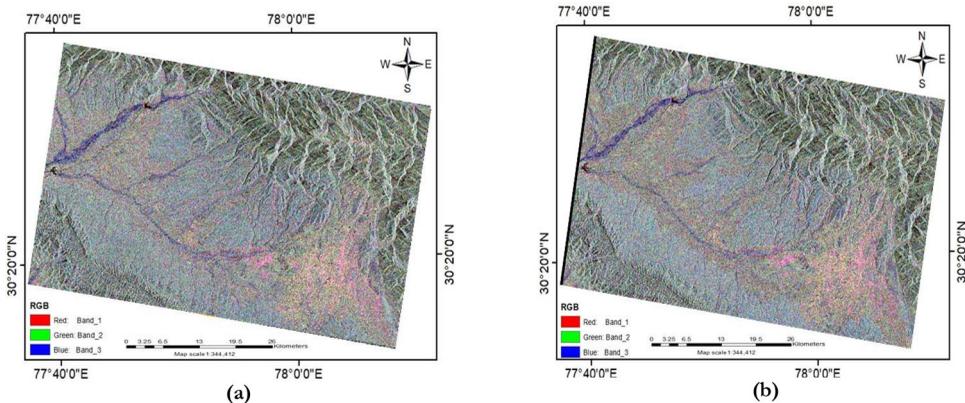


Figure 9. Represent the Pauli colour composite of the (a) master and (b) slave acquisitions.

$$\langle [T] \rangle = \begin{bmatrix} \frac{1}{2} \langle |S_{HH} + S_{VV}|^2 \rangle & \frac{1}{2} \langle (S_{HH} + S_{VV})(S_{HH} - S_{VV})^* \rangle & \langle (S_{HH} + S_{VV})S_{HV}^* \rangle \\ \frac{1}{2} \langle (S_{HH} - S_{VV})(S_{HH} + S_{VV})^* \rangle & \frac{1}{2} \langle |S_{HH} - S_{VV}|^2 \rangle & \langle (S_{HH} - S_{VV})S_{HV}^* \rangle \\ \langle S_{HV}(S_{HH} + S_{VV})^* \rangle & \langle S_{HV}(S_{HH} - S_{VV})^* \rangle & \langle 2|S_{HV}|^2 \rangle \end{bmatrix}$$

The Yamaguchi four-component decomposition modelling approach was then implemented for the retrieval of the backscatter components. In Yamaguchi four-component model, the reflection symmetry is not assumed and can be therefore denoted as, $\langle S_{HH}S_{HV}^* \rangle \neq 0$ and $\langle S_{VV}S_{HV}^* \rangle \neq 0$. The estimated coherence matrix can be represented as linear combination of four canonical scattering mechanisms (Yamaguchi et al. 2005; Yamaguchi et al. 2006) as shown in Equation (6).

$$[T] = \frac{f_s}{1 + |\beta|^2} \begin{bmatrix} 1 & \beta^* & 0 \\ \beta & |\beta|^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{f_d}{1 + |\alpha|^2} \begin{bmatrix} |\alpha|^2 & \alpha & 0 \\ \alpha^* & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{f_v}{4} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \frac{f_c}{2} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \pm j \\ 0 & \mp j & 1 \end{bmatrix} \quad (6)$$

where, f_s , f_d , f_v and f_c are expansion coefficients, and α , β are complex observables. The coefficients can be expressed as follows:

$$\begin{aligned} f_v &= 8 \langle |S_{HV}|^2 \rangle - 4 |\text{Im} \langle S_{HV}^* (S_{HH} - S_{VV}) \rangle| \\ f_s &= B - \frac{|C|^2}{A} \\ f_d &= \frac{1}{2} \langle |(S_{HH} - S_{VV})|^2 \rangle - 2 \langle |S_{HV}|^2 \rangle \\ f_c &= 2 |\text{Im} \langle S_{HV}^* (S_{HH} - S_{VV}) \rangle| \end{aligned} \quad (7)$$

where

$$\begin{aligned} A &= \frac{1}{2} \langle |(S_{HH} - S_{VV})|^2 \rangle - 2 \langle |S_{HV}|^2 \rangle \\ B &= \frac{1}{2} \langle |(S_{HH} + S_{VV})|^2 \rangle - 4 \langle |S_{HV}|^2 \rangle + 2 |\text{Im} \langle S_{HV}^* (S_{HH} - S_{VV}) \rangle| \\ C &= \frac{1}{2} \langle (S_{HH} + S_{VV})(S_{HH} - S_{VV})^* \rangle \end{aligned}$$

In this study, the polarimetric features that have been used in AGB estimation includes the double-bounce, surface, volume and helix backscatter coefficients retrieved from the Yamaguchi four-component modelling approach. Additionally, the entropy, anisotropy, sigma-nought HV, sigma-nought HH and sigma-nought VV backscatter coefficients were extracted from the H-alpha decomposition approach (Lee and Pottier 2009).

Coherence value retrieval using PolInSAR data

Firstly, the master and the slave images were co-registered. From our analysis, a shift of 2 rows and 43 columns was observed and compensated. Then, the $[T_6]$ matrix was generated from the Pauli vectors k_1 and k_2 vectors, as shown in the Equation (8) (Lavalle et al. 2009; Cloude 2010).

```

1 Algorithm: Random Forest Regression
2
3 Input: The PolSAR backscatter parameters and PolInSAR coherence values as independent variables and field estimated AGB as the target variable
4 Output: Predicted AGB values and predicted AGB map
5
6 Preparation of the input data
7 Load the data in .csv format
8 drop the predictor attribute [pred_attr = 'AGB'] and save as y dataset
9 drop extra attributes [drop_attr = 'Plots']
10 save the independent variables as X dataset
11 [drop_columns = [pred_attr] + [attr for attr in drop_attrs]
12 X = dataset.drop(columns=drop_columns)
13 Splitting the data into training and test data subsets
14 X and y datasets splitted in X_train, y_train, X_test, y_test
15 mention the train size = 0.703, (70:30 ratio)
16 implement the train_test_split() function
17 Fitting the RFR model
18 mention the number of regression trees
19 n_estimators = 250 for 250 regression trees
20 fit the model with the training data subset and predict AGB
21 regressor = RandomForestRegressor()
22 regressor.fit(X_train, y_train)
23 y_pred = regressor.predict(X_test)
24 Feature Importance test
25 to check the sensitivity of the independent variables
26 use: feature_imp = ".join(str(np.round(i, 2)) for i in regressor.feature_importances_)
27 Validation of the model results
28 Check for the r2, rmse, rpiq values
29 Mapping of AGB
30 Load the input images as arrays using gdal_array.Loadfile() function.
31 stack the input images
32 Predict the AGB image from the stacked input images
33 plot the predicted AGB map
34 save the predicted AGB map

```

Figure 10. Algorithm of the RFR model for predicting and mapping of AGB.

$$\begin{aligned}
 k_1 &= \begin{bmatrix} S_{HH}^1 + S_{VV}^1 \\ S_{HH}^1 - S_{VV}^1 \\ 2S_{HV}^1 \end{bmatrix} & k_2 &= \begin{bmatrix} S_{HH}^2 + S_{VV}^2 \\ S_{HH}^2 - S_{VV}^2 \\ 2S_{HV}^2 \end{bmatrix} \\
 [T_6] &= < \begin{bmatrix} S_{HH}^1 + S_{VV}^1 \\ S_{HH}^1 - S_{VV}^1 \\ 2S_{HV}^1 \\ S_{HH}^2 + S_{VV}^2 \\ S_{HH}^2 - S_{VV}^2 \\ 2S_{HV}^2 \end{bmatrix} \begin{bmatrix} S_{HH}^{1*} + S_{VV}^{1*} & S_{HH}^{1*} - S_{VV}^{1*} & 2S_{HV}^{1*} & S_{HH}^{2*} + S_{VV}^{2*} & S_{HH}^{2*} - S_{VV}^{2*} & 2S_{HV}^{2*} \end{bmatrix} \\
 [T_6] &= \begin{bmatrix} [T_{11}] & [\Omega_{12}] \\ [\Omega_{12}]^{*T} & [T_{22}] \end{bmatrix}
 \end{aligned} \tag{8}$$

The subscripts 1 and 2 denote the two images acquired at each end of the temporal baseline, that is the master and the slave image. $[T_{11}]$ and $[T_{22}]$ represent the Hermitian coherency matrix of each acquisition and, $[\Omega_{12}]$ is a non-hermitian complex polarimetric interferometric matrix (Cloude 2010).

```

1 Algorithm: Multiple Linear Regression
2
3 Input: The PolSAR backscatter parameters and PolInSAR coherence values as independent variables and field estimated AGB as the target variable
4 Output: Predicted AGB values and predicted AGB map
5
6 Preparation of the input data
7 Load the data in .csv format
8 drop the predictor attribute [pred_attr = 'AGB'] and save as y dataset
9 drop extra attributes [drop_attr = 'Plots']
10 save the independent variables as X dataset
11 [drop_columns = [pred_attr] + [attr for attr in drop_attr]]
12 X = dataset.drop(columns=drop_columns)
13 Splitting the data into training and test data subsets
14 X and y datasets splitted in X_train, y_train, X_test, y_test
15 mention the train_size = 0.703, (70:30 ratio)
16 implement the train_test_split() function
17 Fitting the MLR model
18 fit the model with the training data subset and predict AGB
19 regressor = linear_model.LinearRegression()
20 regressor.fit(X_train, y_train)
21 y_pred = regressor.predict(X_test)
22 Validation of the model results
23 Check for the r2, rmse, rpiq values
24 Mapping of AGB
25 Load the input images as arrays using gdal_array.Loadfile() function.
26 stack the input images
27 Predict the AGB image from the stacked input images
28 plot the predicted AGB map
29 save the predicted AGB map
30

```

Figure 11. Algorithm of the MLR model for predicting and mapping of AGB.

From the definition of Ω_{12} , the complex coherence can be obtained as (Cloude 2010).

$$\mu_1 = \omega_1^\dagger K_{p1} \quad (9)$$

$$\mu_2 = \omega_2^\dagger K_{p2} \quad (10)$$

where ω_1 and ω_2 are the unitary projection vectors. μ_1 and μ_2 are complex scalars due to the projection of the scattering vectors on ω_1 and ω_2 . Complex coherence can be expressed as the vectorized form of the coherence. It is the product of the coherence magnitude and the phase factor.

$$Y = |Y|e^{i\phi} = \gamma_{(\omega_1\omega_2)} \quad (11)$$

$$\gamma_{(\omega_1\omega_2)} = \frac{\langle \mu_1 \mu_2^\dagger \rangle}{\sqrt{\langle \mu_1 \mu_1^\dagger \rangle \langle \mu_2 \mu_2^\dagger \rangle}} = \frac{\omega_1^\dagger \Omega_{12} \omega_2}{\sqrt{\omega_1^\dagger T_{11} \omega_1 \omega_2^\dagger T_{22} \omega_2}} \quad (12)$$

Prediction of AGB using RFR algorithm

All the backscatter components and the coherence values were taken as inputs for the RFR algorithms represented in Figure 10. The algorithm was optimized by tuning the hyper-parameters, namely, random subset of predictors (*'mtry'*) and the number of regression trees (*'ntree'*) used to form the RF (Ghosh and Behera 2018; Stelmaszczyk-Górska et al. 2018; Brigot et al. 2019). RFR trees were generated using the following independent variables, PolSAR- and PolInSAR-derived variables, namely, entropy, anisotropy, sigma-nought HV, sigma-nought VV, sigma-nought HH backscatter coefficients, volume, helix, surface, double-bounce backscatter mechanisms and HH, HV, VV, HH-VV, HH + VV and HV + VH coherences. Two-third data (26 samples) were used for training the RFR algorithm, while the remaining one-third data (11 samples) were used for testing and validation purpose. The algorithm predicts the target variable using various input parameters

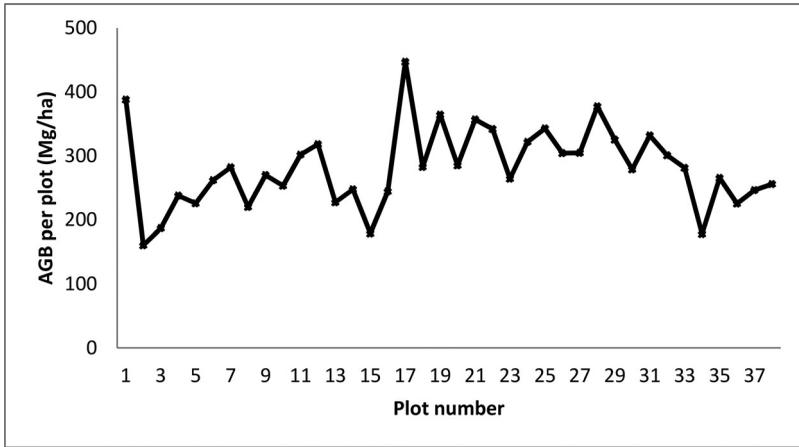


Figure 12. Graph representing the distribution of the aboveground biomass of 38 sample plots.

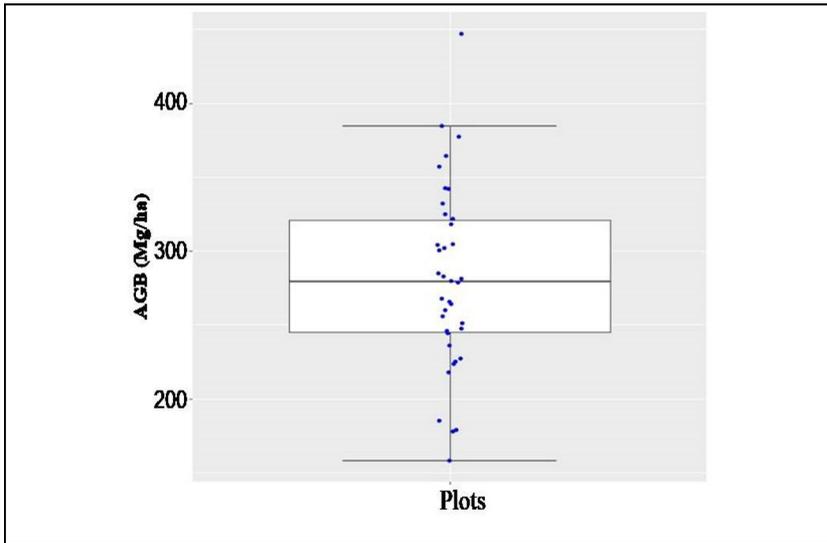


Figure 13. Boxplot showing the distribution of each plot-level AGB with the maximum and minimum values. Plot 17 can be identified as an outlier of the data set.

as an ensemble average of the predictions of each regression trees as represented in Equation (13) (Shataee et al. 2012; Esteban et al. 2019). Each tree is built randomly based on the training sample subset of the data (Dhanda 2013).

$$\widehat{AGB}_R = \frac{1}{k} \sum_{k=1}^k k_{\text{th tree responses}} \quad (13)$$

where k is the number of regression trees used for the prediction and \widehat{AGB} is the estimated AGB from the SAR data using the RFR algorithm.

Prediction of AGB using MLR model

The MLR model enables to develop a statistical relationship to explain the variation of the target variable with respect to the variations of the independent variables, as explained

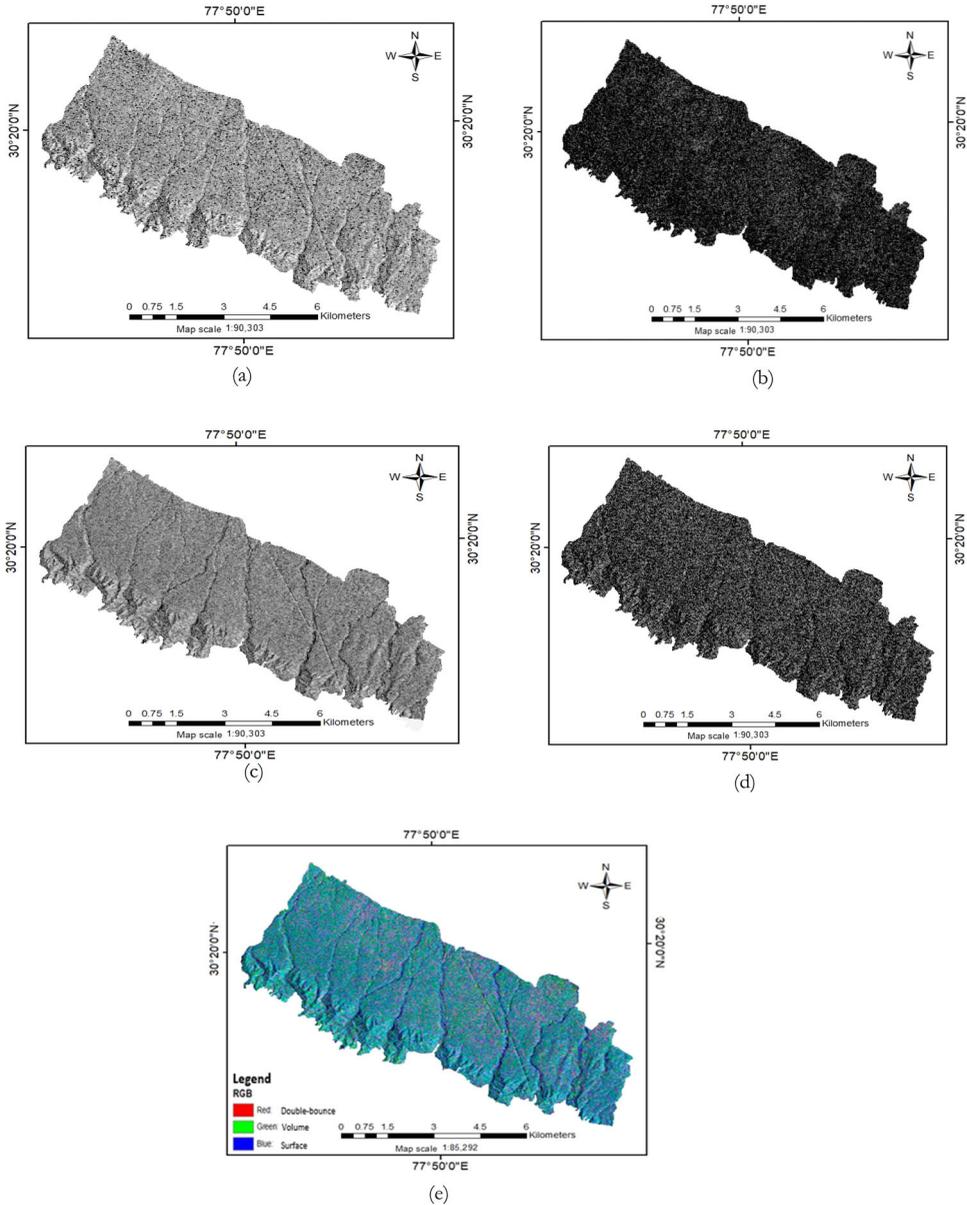


Figure 14. The products of the Yamaguchi four-component decomposition modelling: (a) surface backscatter, (b) double-bounce backscatter, (c) volume backscatter, (d) helix backscatter, and (e) RGB colour composite (R – double-bounce backscatter, G – volume backscatter, B – surface backscatter).

in [Figure 11](#) (Stein et al. 1999; Mangla 2015). The independent variables can be more than one, as shown in [Equation \(14\)](#).

$$\widehat{AGB}_M = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_n \cdot x_n \quad (14)$$

where β_0 is the intercept, $\beta_1, \beta_2, \dots, \beta_n$ are the coefficients, x_1, x_2, \dots, x_n are the independent variables and \widehat{AGB}_M is the dependent variable, that is the estimated AGB from the SAR data using the MLR algorithm.

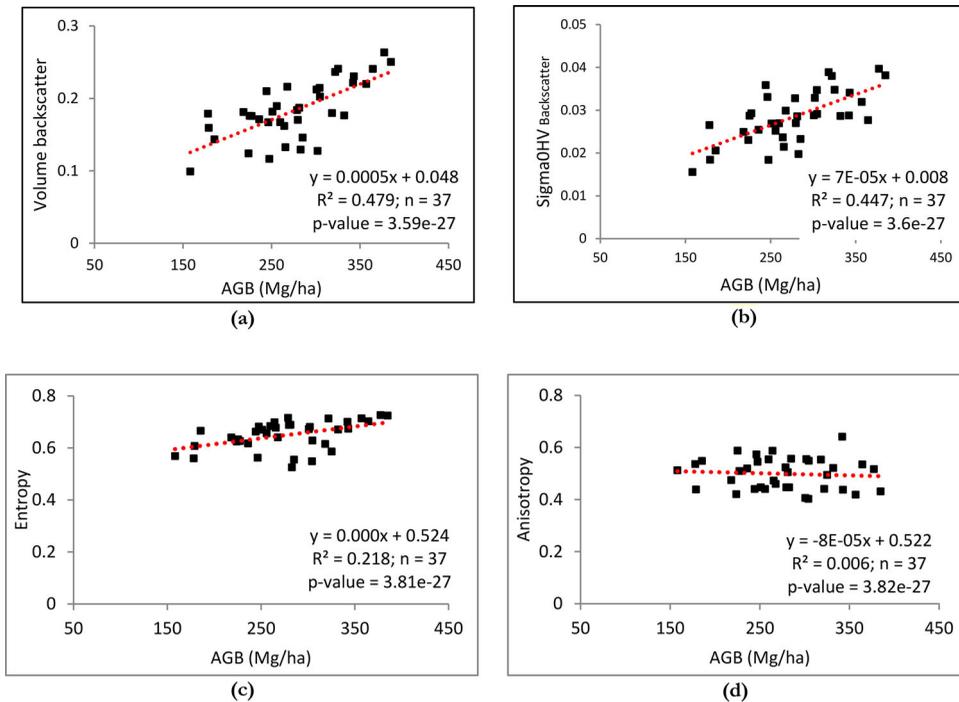


Figure 15. The graphs represent the relation between the field estimated AGB and the PolSAR backscatter components like (a) volume backscatter; (b) sigma-nought HV; (c) entropy; (d) anisotropy.

In this study, the independent variables of RFR, with higher sensitivity to the target variable were selected to serve as the independent variables for the MLR model. The multi-collinearity test was performed to check the correlation of the independent variables of the developed regression model. Also, the variance inflation factors (VIFs) were calculated to indicate the degree of collinearity among the independent variables (O'Brien 2007). The validation and accuracy assessment for both the regression algorithms were done by comparing the predicted AGB with respect to the test data subset, that is the field estimated AGB.

Results

Field-estimated AGB

The tree-level field-estimated AGB contributed to the plot-level estimation of AGB. The AGB for the entire study area was observed to range from 158.14 to 384.84 Mg/ha, as shown in Figure 12. The AGB for plot 17 was seen to fall as an outlier from the complete data set having a value of 447.23 Mg/ha, as shown in Figure 13. The rest of the plots were observed to be almost evenly spread within the inter-quartile ranges.

Backscatter parameter retrieval using PolSAR data for AGB estimation

The backscatter images were retrieved from the Yamaguchi four-component decomposition modelling approach. The surface, double-bounce, volume and helix scattering parameters were extracted from the respective images for all the sample plots as shown in

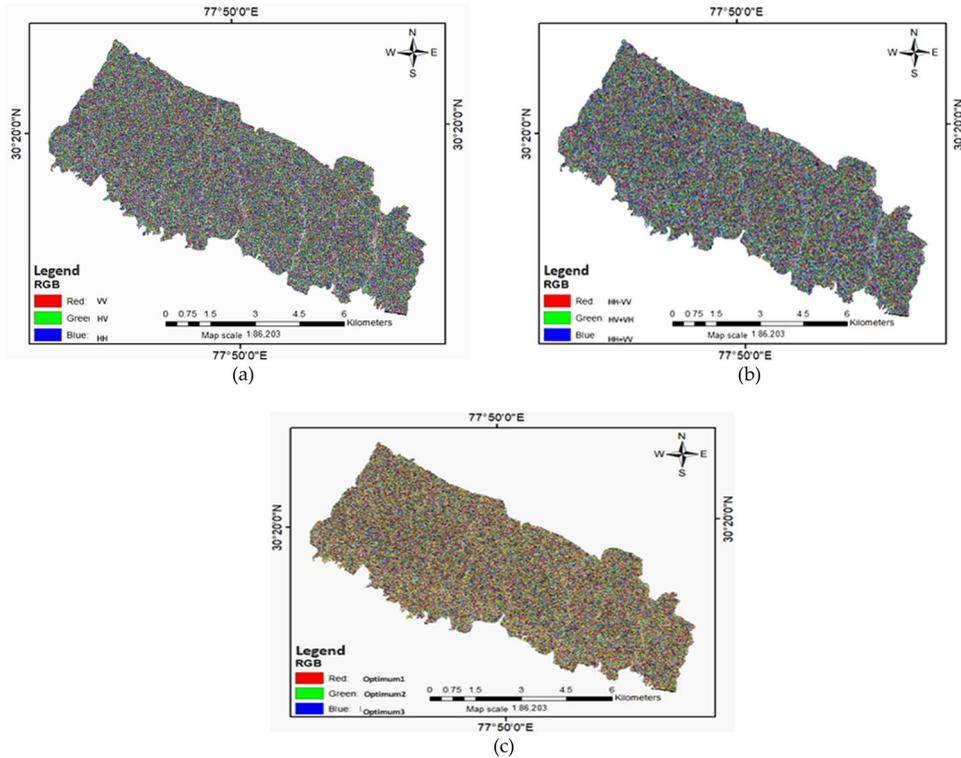


Figure 16. The RGB colour composites of the coherence images in (a) linear basis representing RGB as VV, HV and HH band combinations, (b) Pauli basis representing RGB as HH-VV, 2HV and HH + VV band combinations and (c) optimal basis representing RGB as Opt-1, Opt-2 and Opt-3 band combinations.

Figure 14a–d. The colour composite image of the Yamaguchi four-component decomposition is shown in **Figure 14e**. The red band represents double-bounce scattering from stable scatterers, the green band represents volume scattering from the vegetated area and blue band represents surface scattering from top surface canopy and coherent scatterers. The H/A/alpha (Entropy/Anisotropy/alpha)decomposition modelling approach was also implemented to extract the entropy, anisotropy and sigma-nought HV scattering parameters.

The relationship between the PolSAR backscatter components and the field-estimated AGB for the 37 sample plots were observed, as shown in **Figure 15**. The coefficient of determination for AGB with respect to volume backscatter, sigma-nought HV backscatter coefficient, entropy and anisotropy is 0.479, 0.447, 0.218 and 0.006, respectively. The p values for the volume backscatter, sigma-nought HV backscatter coefficient, entropy and anisotropy are $3.59\text{e-}27$, $3.6\text{e-}27$, $3.81\text{e-}27$ and $3.82\text{e-}27$, respectively.

Coherence value retrieval using PolInSAR data for AGB estimation

The complex coherence values retrieved from the $[T_6]$ matrix are a function of the polarization basis, as shown in **Equation (8)**, which range from 0 to 1. The coherences in the linear polarization basis are HH, HV and VV. In the HH channel, the polarization basis is: $\omega_1 = \omega_2 = \left[\frac{1}{\sqrt{2}} \quad \frac{-1}{\sqrt{2}} \quad 0 \right]^T$ whereas in the VV channel, the polarization basis are $\omega_1 = \omega_2 = \left[\frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}} \quad 0 \right]^T$ (Khati 2014), and the polarization basis for the HV channel

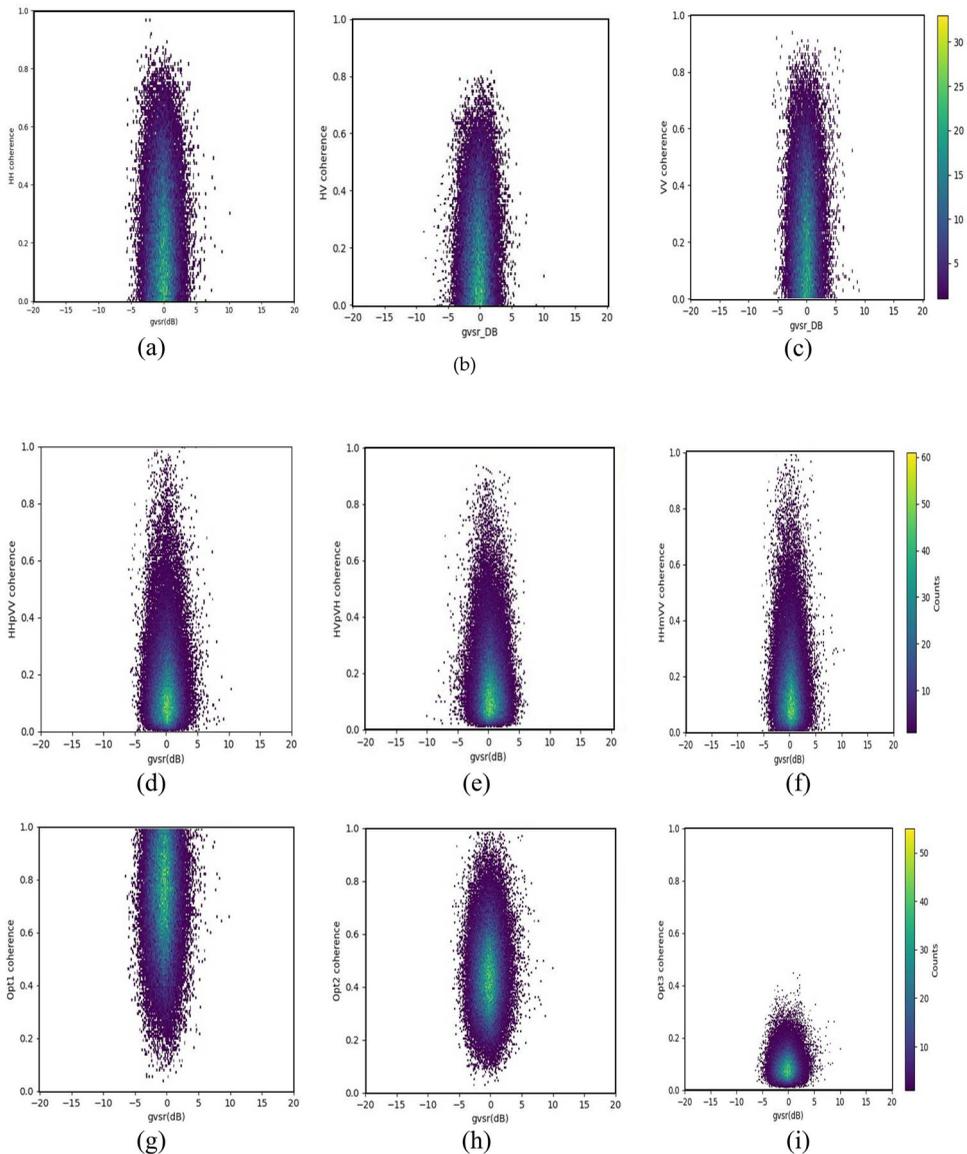


Figure 17. The density scatter plots of the coherences in different polarization basis as a function of the ground-to-volume scattering ratio derived based on double-bounce scattering: (a) HH coherence, (b) HV coherence, (c) VV coherence, (d) HH + VV coherence, (e)HV + VH coherence, (f) HH-VV coherence, (g) Optimum-1 coherence, (h) Optimum-2 coherence and (i) Optimum-3 coherence.

is $\omega_1 = \omega_2 = [0 \ 0 \ 1]^T$ (Khati 2014). The coherences were also retrieved for Pauli basis (HH + VV, HV + VH and HH-VV) and Optimal basis (Opt-1, Opt-2 and Opt-3). The RGB colour composite images of the coherences for linear, Pauli and optimal basis are shown in Figure 16.

The dependency of coherence with respect to ground-to-volume scattering ratio was analysed from the scatter plots shown in Figure 17. The ground-to-volume scattering ratio was derived from the ratio of the double-bounce and volume scattering powers extracted from the Yamaguchi four-component decomposition modelling results. The ratio is observed to range between -20 and $+20$ dB. The positive range, that is 0 to $+20$ dB,

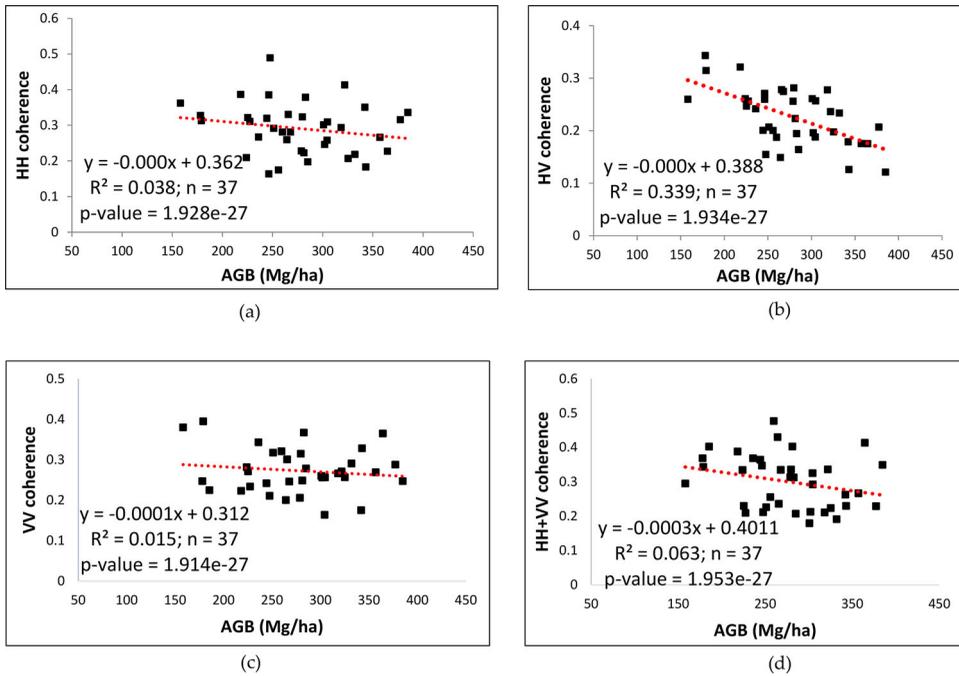


Figure 18. Scatter plot between field AGB and coherence values: (a) HH coherence, (b) VV coherence, (c) HV coherence, (d) HH + VV coherence, (e) HV + VH coherence, (f) HH-VV coherence, (g) Optimum coherence 1, (h) Optimum coherence 2 and (i) Optimum coherence 3.

represents the dominance of ground scattering, whereas the negative range, that is 0 to -20 dB, represents the dominance of volume scattering. The coherences of each polarization basis were plotted against the ground-to-volume scattering ratio, and their dependency was analysed.

The relation between the coherences and the field-estimated AGB for the 37 sample plots was also observed, as shown in Figure 18. The coefficient of determination for AGB and each polarization channel, HH, HV, VV, HH + VV, HV + VH, HH-VV, optimum 1, optimum 2 and optimum 3 were observed to be 0.038, 0.339, 0.015, 0.063, 0.398, 0.007, 0.0001, 0.0002 and 0.018, respectively. The p values for each polarization channel, HH, HV, VV, HH + VV, HV + VH, HH-VV, optimum 1, optimum 2 and optimum 3 were observed to be $1.928e-27$, $1.934e-27$, $1.914e-27$, $1.953e-27$, $1.957e-27$, $1.919e-27$, $2.014e-27$, $1.953e-27$ and $1.887e-27$, respectively.

Prediction of AGB using RFR algorithm

The distribution of RMSE with respect to the size of training data subset and the number of regression trees were plotted, as shown in Figure 19a. Also, the distribution of the ‘test score’, that is the value of the coefficient of determination, was plotted with respect to the size of training data subset as shown in Figure 19b. The feature importance and sensitivity was observed for all the independent variables, as shown in Figure 19c.

For validation of the predicted values, a linear regression analysis was done between the field-estimated AGB and the predicted AGB of the RFR model (\widehat{AGB}_R), as shown in Figure 19d followed by the mapping of AGB for the entire study area, as shown in Figure 19e.

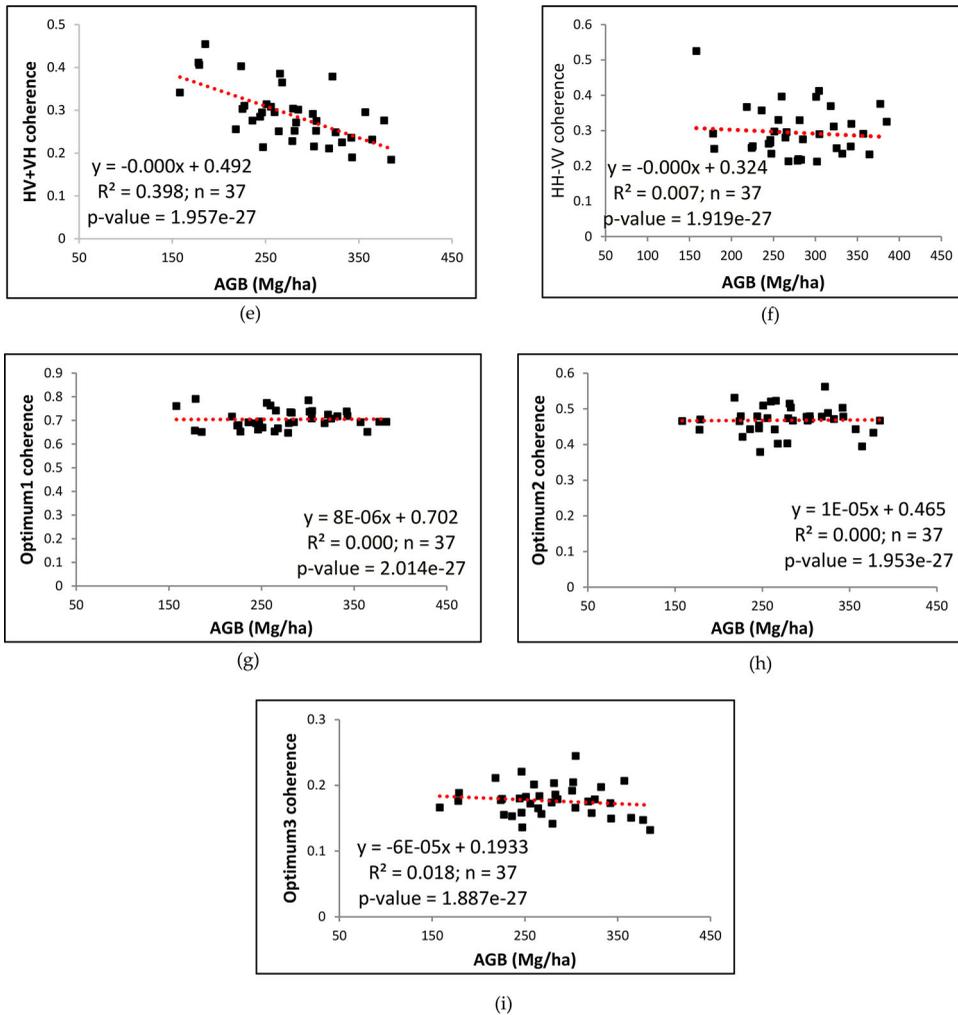


Figure 18. Continued.

Prediction of AGB using MLR model

The most sensitive parameters retrieved from the optimised RFR algorithm were selected for developing a regression equation using the MLR model. The scatter plot matrix, shown in Figure 20, represents the distribution of all the independent variables against each other. The independent variables were volume backscatter component, sigma-nought HV backscatter component, HV coherence and HV + VH coherence, and the target variable was \widehat{AGB}_M . The first regression equation was developed, as shown in Equation (15), followed by the test of multi-collinearity. The multi-collinearity test was done to check the interdependency or the degree of collinearity of the input variables. After discarding the highly interdependent variables, the second regression equation was developed, as shown in Equation (16), followed by the estimation of the variation inflation factors (VIFs).

$$\widehat{AGB}_M = \quad (15)$$

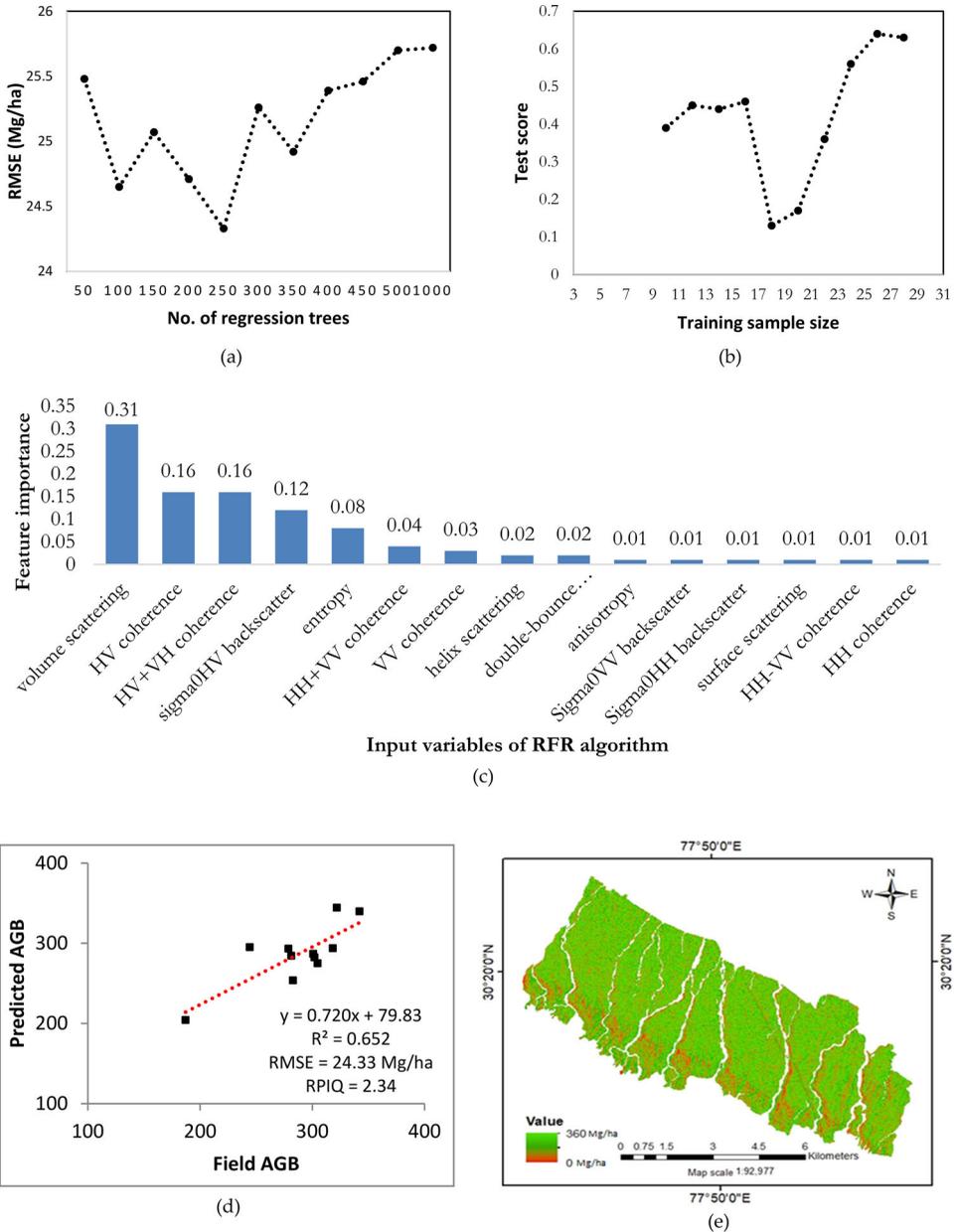


Figure 19. (a) The distribution of RMSE with respect to the variation in the number of regression trees; (b) the distribution of 'test_score' values with respect to the size of the training data set; (c) bar graph indicating the sensitivity of the input variables in the RFR algorithm; (d) relation between field estimated AGB and predicted AGB; (e) mapped AGB using RFR algorithm.

$$222.519 + 0.177 \times (\text{sig0HV}) - 0.115 \times (\text{HV} + \text{VH}) + 0.445 \times (\text{vol}) - 0.362 \times (\text{HV})$$

$$\widehat{AGB}_M = 227.373 - 0.390 \times (\text{HV} + \text{VH}) + 0.589 \times (\text{vol}) \quad (16)$$

For validation, a linear regression analysis was done between the AGB estimated from field inventory and the predicted AGB of the MLR model (\widehat{AGB}_M), using Equation (16),

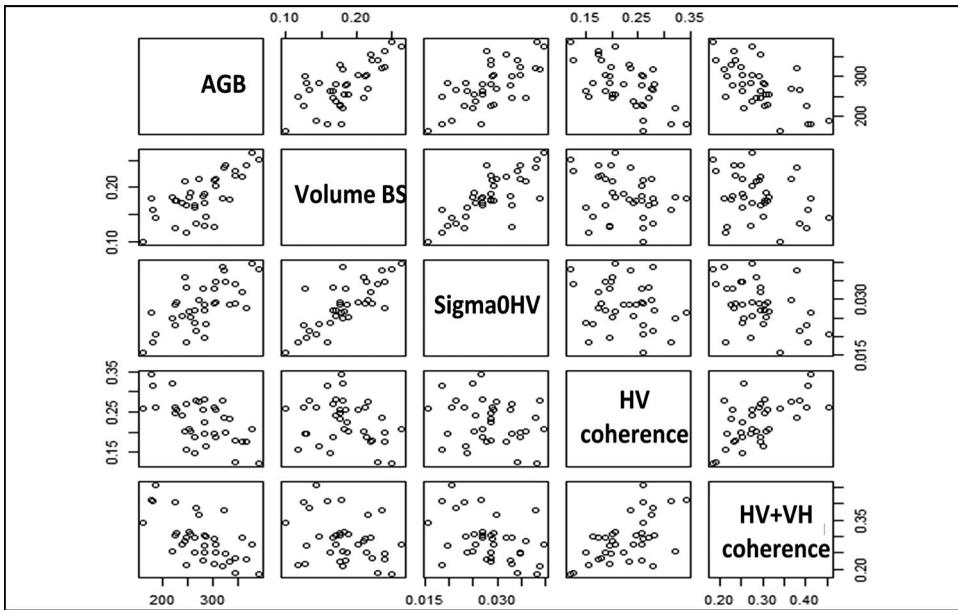


Figure 20. The scatter plot matrix showing the distribution of the input variables with respect to each other and the target variable.

as shown in [Figure 21a](#) followed by the mapping of AGB for the entire study area, as shown in [Figure 21b](#).

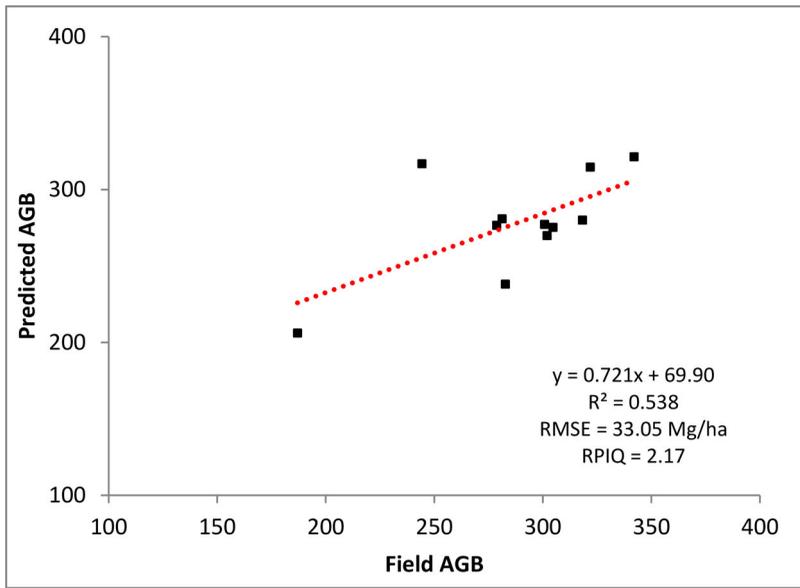
Discussion

In this paper, we investigated the efficiency of the coherent implementation of PolSAR and PolInSAR techniques for parameter retrieval and using them for AGB estimation through regression modelling and ML approach. The field-estimated AGB data showed that biomass had a broad range even in a homogenous forest. Therefore, the PolSAR- and PolInSAR-based parameters were analysed with respect to the different levels of AGB distributed over the entire study area and to examine their sensitivity for the prediction of AGB.

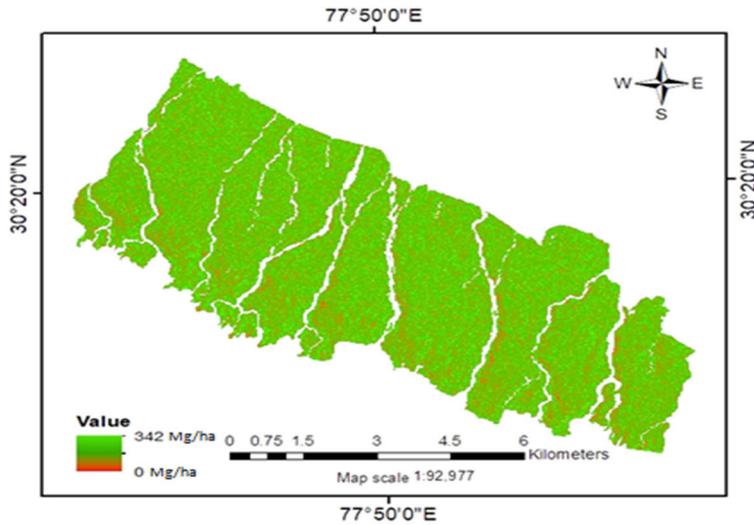
In the RGB colour composite image of the Yamaguchi four-component decomposition model, shown in [Figure 14e](#), where the vegetated areas were observed resulting in volume scattering. Surface scattering was observed mainly from the top surface of the canopy and also from the dry riverbed regions due to the presence of stable scatterers such as rocks, pebbles and sand. Moreover, there are some patches that are resulting in double-bounce scattering between the vegetated areas from the regions with less dense canopy cover.

The relation between the volume backscatter, sigma-nought HV backscatter and entropy against field-estimated AGB showed (in [Figure 15](#)) a positive linear correlation in each case. The densely vegetated areas contributed to higher volume backscatter, higher sigma-nought HV backscatter, as well as higher entropy than the lesser dense regions. And from the p values it was observed that the PolSAR backscatter components are statistically significant for the prediction of AGB with a confidence level of 95%.

In the colour composite image of the coherences in the linear polarization basis shown in [Figure 16a](#), the HV coherence was observed mainly from the forest area as well as compact urban structures. The dry riverbeds were observed to contribute in double-bounce scattering due to the presence of stable scatterers. In the colour composite image



(a)



(b)

Figure 21. (a) Relation between field estimated AGB and predicted AGB (\widehat{AGB}_M); (b) mapped AGB using MLR model.

of the Pauli basis, as shown in Figure 16b, double-bounce scattering is represented by HH-VV, surface scattering is represented by HH + VV and the volume scattering is represented by HV + VH. The riverbeds are represented by cyan colour, which indicates the presence of stable surface (ground or sand) and double-bounce scatterers (stones). The forest regions appear darker due to the volume decorrelation factor in the forest region. The colour composite image of the optimum polarizations as shown in Figure 16c, the coherence for the forest region was observed to be low in all the three polarizations due to high decorrelation as it appears darker, while the riverbeds appear bright indicating low decorrelation.

Table 3. SWOT analysis.

	Strengths and opportunities	Weaknesses and Threats
Internal factors	Integration of PolSAR backscatter components and PolInSAR coherence values Flexibility in terms of considering input variables for the RFR algorithm Attaining an acceptable accuracy between predicted AGB and field estimated AGB Having scope for further improvement and extension of the present study	To check the robustness of the optimized RFR and MLR models Comparative analysis using multi-frequency data sets
External factors	Intensive field inventory Implementation of systematic sampling technique in a sub-tropical forest with dense undergrowth Quadpol C-band Radarsat-2 data sets having the ability of cloud cover penetration and efficiency in the characterization of objects Large-scale estimation of AGB compared to conventional methods Positive cooperation from forest personnel of Dehradun Forest Division	Weather constraints during field inventory Presence of dense undergrowth in the ground strata of the study area Presence of wildlife in the forest area Uncertainty due to the allometric equations provided by Forest Survey of India (FSI) Uncertainty due to instrumental errors

The coherence with respect to ground-to-volume scattering ratio was analysed in order to study the dependency of coherence as a function of the ground-to-volume scattering ratio. For the linear basis polarizations (Figure 17(a–c)), the range of HV in the y axis in both the cases was seen to be lower than that of HH and VV, which represents higher decorrelation from randomly oriented dipole scatterers. For the Pauli basis, the HV + VH and HH-VV polarizations were seen to show a similar distribution pattern as in Figure 17d–f. The HH-VV polarization in both cases was seen to be centred at 0 dB, therefore, representing the contribution from stable scatterers, which were independent of the ground-to-volume scattering ratio. And, HV + VH in both the cases have a lower range of distribution in y axis as they represent higher decorrelation from randomly oriented dipole scatterers. On the other hand, the coherences in the optimal basis (in Figure 17(g–i)) show no dependency on the ground-to-volume scattering ratio as they have an equiprobable distribution with a gradual decrease in the range of coherence values from opt-1 to opt-2 and the least in opt-3. From this analysis, it can be concluded that the HH, VV, HH + VV and HH-VV co-pols are more sensitive to ground scattering than the cross-pol HV and HV + VH for a forest region.

In the linear regression plots of PolInSAR coherence and field-estimated AGB, of Figure 18, a high correlation was observed for using a C-band data set as it has a low frequency and, hence, less penetrating power. Therefore, the major percentage of contribution to coherence is from the canopy cover. The plots with dense canopy cover contributed to lower coherence than the plots with less dense canopy cover. The highest correlation was observed between coherence values in HV + VH polarization and field biomass, followed by HV polarization. Therefore, from the analysis of coherence, it was seen that scattering media, that is the vegetation, interacts differently with different combinations of coherences in different polarization basis and also with different back-scattering components. Also, from the p values it was observed that the PolInSAR coherences are statistically significant for the prediction of AGB with a confidence level of 95%.

The RFR algorithm was optimised at 250 regression trees and with 70:30 training and testing data ratio. From the analysis of \widehat{AGB}_R , it was found that the highest ranked variables in the feature importance graph, shown in Figure 19c, are the most sensitive parameters for AGB estimation. The ranking of the variables was volume backscatter

component followed by HV and HV + VH coherences, and sigma-nought HV backscatter coefficient. Therefore, it can be said that the most sensitive parameters make the best combination of variables for the prediction of AGB with highest accuracy and yields the lowest RMSE. From the RFR algorithm (Figure 19(d)), the value of the coefficient of determination obtained was 0.65, the RMSE value was 24.33 Mg/ha and the RPIQ (ratio of performance to inter-quartile distance) value was 2.34. And from the MLR model (Figure 21(a)), the value of the coefficient of determination obtained was 0.54, the RMSE value was 33.05 Mg/ha and the RPIQ value was 2.17. The values for coefficient determination and RPIQ were observed to be greater for RFR model, whereas the RMSE value was observed to be lower for RFR model as compared with the MLR model. Therefore, the prediction of AGB using RFR was considered and seen to be better than that of MLR. The mapping of AGB was done for both the regression models, as shown in Figures 19e and 21b. The non-vegetated areas such as riverbed boundaries, power-line and fire lines showed AGB value of 0 Mg/ha represented with red colour. The areas with open canopy cover were seen to have intermediate values of AGB represented by dark green colour, and the areas with higher AGB values were represented with a light green colour. A better representation of AGB for the entire study area can be observed in the case of RFR-based modelling as compared with that of MLR-based modelling. The regions of intermediate AGB and open areas were distinctly identifiable in case of the RFR-based AGB model and less distinctly identifiable for MLR-based AGB model.

The SWOT analysis of this study was evaluated as shown in Table 3. The main strengths and opportunities of this work were the availability of the quadpol C-band data sets, which enabled the proper characterization of the targets in the study area and helped to obtain acceptable results within the given time duration. The positive cooperation of the forest personnel of Dehradun Forest Division was helpful during field inventory work and while traversing through the forests, whereas the main weaknesses and threats encountered during this study were the uncertainty due to instrumental errors such as RTK and laser range finder. These were reduced by taking RTK readings towards the centre of each sampling unit and measuring tree height with the laser range finder 2–3 times for proper validation. Uncertainty was also incorporated due to the allometric equations used.

Conclusions

This study showed that the integrated use of PolSAR and PolInSAR parameters retrieved from an image pair of fully polarimetric Radarsat-2 C-band data used as input variables for RFR, and MLR models gives optimal prediction for AGB estimation. Out of these two, the RFR gives the best result for the prediction of AGB. The advantage of RFR over MLR is the flexibility of input variables. The number of input variables can be large in case of RFR, but MLR often fails for the same. RFR is also not vulnerable to over-fitting and noise as compared with MLR. RFR also allows to estimate model error without bias. The volume scattering, HV and HV + VH coherences and sigma-nought HV backscatter coefficient components were identified to be the most sensitive and important parameters for the prediction of AGB.

Acknowledgements

The authors would like to express their sincere gratitude to the whole research team of ESA for providing SNAP 6.0 and PolSARPro 4.2 tools for polarimetric processing of the SAR data. The Technology

Development Programme (TDP) of Indian Space Research Organisation (ISRO) supported this research. The Radarsat-2 data were purchased from the MDA corporation under the ISRO TDP. The spaceborne C-band Radarsat-2 PolInSAR data were procured under a TDP (code: Z1Z12PR71601-03) on Polarimetric SAR Calibration of PolSAR Data. The authors are thankful to the Indian Institute of Remote Sensing (IIRS), ISRO for providing all the support to carry out this research work. The authors are grateful to Ms. Shefali Agarwal, Group Head, Geospatial Technology and Outreach Program Group (GT&OP Group), IIRS, Dehradun for her invaluable contribution and guidance in permitting field data collection and other necessary facilities in this research. The authors would like to express their gratitude to Mr. M.E. Bhanu Prakash (M.Tech.), Mr. Shahid Shuja Shafai (M.Sc.) and Mr. Sriharsha Yegireddi (M.Sc.) for their spontaneous contributions to field data collection. The authors would like to express their sincere thanks to the Uttarakhand Forest Department, Uttarakhand, India, for granting permission for field data collection in Malhan Forest Range, Dehradun Forest Division, Uttarakhand India.

Author's contributions

R.M. contributed to data processing, methodology, formal analysis, validation and manuscript preparation. **S.K.** contributed to conceptualization, supervision, methodology, satellite data selection, field data collection, software, validation, writing – reviewing and editing and project administration. **H.A.** contributed to supervision, visualization, investigation, reviewing and editing; **A.K.** supervised the study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The Radarsat-2 data used in this study were purchased under the Technology Development Programme (TDP) of Indian Space Research Organisation (ISRO) on Polarimetric SAR Calibration of PolSAR Data.

ORCID

Ritwika Mukhopadhyay  <http://orcid.org/0000-0002-6052-3685>

Shashi Kumar  <http://orcid.org/0000-0002-2442-7143>

Hossein Aghababaei  <http://orcid.org/0000-0003-3417-2591>

Anurag Kulshrestha  <http://orcid.org/0000-0002-4640-8998>

References

- Adams EE. 2012. Forest Cover: World Forest Area Still on the Decline. *Eco-Economy Indicators*. [accessed 2019 Sept 14]. <http://www.earth-policy.org/indicators/C56>.
- Agrawal N. 2015. PolInSAR based scattering information retrieval for forest aboveground biomass estimation. Enschede: University of Twente Faculty of Geo-Information and Earth Observation (ITC).
- Agrawal N, Kumar S, Tolpekin VA. 2019. Polinsar based scattering information retrieval for forest above-ground biomass estimation. *Int Arch Photogramm Remote Sens Spatial Inf Sci*. XLII-2/W13: 1913–1920.
- Attema EPW, Ulaby FT. 1978. Vegetation modeled as a water cloud. *Radio Sci*. 13(2):357–364.
- Behera MD, Tripathi P, Mishra B, Kumar S, Chitale VS, Behera SK. 2016. Above-ground biomass and carbon estimates of shorea robusta and tectona grandis forests using QuadPOL ALOS PALSAR Data. *Adv Space Res*. 57(2):552–561.
- Bharadwaj PS, Shashi Kumar SPS, Kushwaha W, Bijker 2015. Polarimetric scattering model for estimation of above ground biomass of multilayer vegetation using ALOS-PALSAR quad-pol data. *Phys Chem Earth Parts A/B/C*. 83–84:187–195.
- Brandeis TJ, Delaney M, Parresol BR, Royer L. 2006. Development of equations for predicting Puerto Rican subtropical dry forest biomass and volume. *For Ecol Manage*. 233(1):133–142.

- Brigot G, Simard M, Colin-Koeniguer E, Boulch A. 2019. Retrieval of forest vertical structure from PolInSAR data by machine learning using LIDAR-derived features. *Remote Sens.* 11(4):381.
- Brown S. 1997. Methods for estimating biomass density from existing data. *Advances in remote sensing.* Rome: FAO FORESTRY PAPER.
- Carreiras J, Melo J, Vasconcelos M. 2013. Estimating the above-ground biomass in miombo savanna woodlands (Mozambique, East Africa) using L-band synthetic aperture radar data. *Remote Sens.* 5(4): 1524–1548.
- Champion SHG, Seth SK. 1968. A revised survey of the forest types of India. Delhi: Manager of Publications.
- Chandola S. 2014. Polarimetric SAR interferometry for forest aboveground biomass estimation. Enschede: ITC Faculty of Geo-Information Science and Earth Observation of the University of Twente.
- Chi Y, Guo Z, Shi H-H, Shen C-C, Liu Y-Z. 2016. The plantation biomass on Beichangshan Island based on forest health. *Zhongguo Huanjing Kexue/China Environ Sci.* 36(8):2522–2535. [accessed 2019 Sept 14]. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84986326000&partnerID=40&md5=dc78aab7304ee0202171b8b196aaba3d>.
- Cloude SR. 2010. Polarisation: applications in remote sensing. Oxford: Oxford University Press.
- Dhanda P. 2013. OPTIMISING PARAMETERS OBTAINED FROM MULTIPLE SENSORS FOR BIOMASS ESTIMATION AT ICESAT FOOTPRINT LEVEL USING DIFFERENT REGRESSION ALGORITHMS [M. Tech. Thesis]. Visakhapatnam: Andhra University.
- Duncanson L, Neuschwander A, Hancock S, Thomas N, Fatoyinbo T, Simard M, Silva CA, Armston J, Luthcke SB, Hofton M, et al. 2020. Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. *Remote Sens Environ.* 242:111779.
- Esteban J, McRoberts R, Fernández-Landa A, Tomé J, Naesset E. 2019. Estimating forest volume and biomass and their changes using random forests and remotely sensed data. *Remote Sens.* 11(16):1944.
- FAO. 2018. The state of world's forests 2018 - forest pathways to sustainable development. Rome: FAO.
- Forest Survey of India. 1996. Volume equations for Forests of India, Nepal and Bhutan. Dehra Dun: Forest Survey of India, Ministry of Environment and Forests.
- Freeman A, Durden SL. 1998. A three-component scattering model for polarimetric SAR data. *IEEE Trans Geosci Remote Sens.* 36(3):963–973.
- Garestier F, Dubois-Fernandez PC, Guyon D, Le Toan T. 2009. Forest biophysical parameter estimation using L- and P-band polarimetric SAR data. *IEEE Trans Geosci Remote Sens.* 47(10):3379–3388.
- Ghosh SM, Behera MD. 2018. Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest. *Appl Geogr.* 96:29–40.
- Hogg EH, Brandt JP, Michaelian M. 2008. Impacts of a regional drought on the productivity, dieback, and biomass of western Canadian aspen forests. *Can J for Res.* 38(6):1373–1384.
- Husch B, Kershaw JA, Beers TW. 2003. Forest mensuration. 4th ed. Hoboken (NJ): John Wiley & Sons Inc.
- Husch B, Miller CI, Beers TW. 1972. Forest mensuration. 2nd ed. New York (NY): The Ronald Press Company.
- Joshi SK, Kumar S. 2017a. Performance of PolSAR backscatter and PolInSAR coherence for scattering characterization of forest vegetation using single pass X-band spaceborne synthetic aperture radar data. *J Appl Remote Sens.* 11(2):026022. <http://dx.doi.org/10.1117/1.JRS.11.026022>.
- Joshi SK, Kumar S. 2017b. Spaceborne PolInSAR tomography for vertical profile retrieval of forest vegetation. *J Appl Remote Sens.* 11 (1):016001.
- Kershaw JA, Ducey MJ, Beers TW, Husch B. 2016. Forest mensuration. 5th ed. Chichester: John Wiley & Sons, Ltd.
- Khanna LS, Chaturvedi AN. 1994. Forest mensuration and biometry. Dehradun: International Book Distributors.
- Khati UG. 2014. POLINSAR based scattering information and physical property retrieval of vegetation. Dehradun: Indian Institute of Remote Sensing, ISRO. [accessed 2019 Sept 14]. http://www.iirs.gov.in/iirs/sites/default/files/StudentThesis/Unmesh_Khati_mtech_2012-14.pdf.
- Khati UG, Lavalle M, Shiroma GHX, Meyer V, Chapman B. 2020. Assessment of forest biomass estimation from dry and wet SAR acquisitions collected during the 2019 UAVSAR AM-PM campaign in southeastern United States. *Remote Sens.* 12 (20):3397.
- Krieger G, Papathanassiou KP, Cloude SR. 2005. Spaceborne polarimetric SAR interferometry: performance analysis and mission concepts. *EURASIP J Adv Signal Process.* 2005(20):3272–3292.
- Kumar S. 2009. Retrieval of Forest Parameters from Envisat ASAR Data for Biomass Inventory in Dudhwa National Park, U.P., India. Enschede: ITC Faculty of Geo-Information Science and Earth Observation of the University of Twente.

- Kumar S, Garg RD, Govil H, Kushwaha SPS. 2019. PolSAR-decomposition-based extended water cloud modeling for forest aboveground biomass estimation. *Remote Sens.* 11(19):2287.
- Kumar S, Garg RD, Kushwaha SPS, Jayawardhana WGNN, Agarwal S. 2017. Bistatic PolInSAR inversion modelling for plant height retrieval in a tropical forest. *Proc Natl Acad Sci, India, Sect A Phys Sci.* 87(4):817–826.
- Kumar S, Govil H, Srivastava PK, Thakur PK, Kushwaha SPS. 2020. Spaceborne multifrequency PolInSAR-based inversion modelling for forest height retrieval. *Remote Sens.* 12 (24):1–27.
- Kumar S, Khati UG, Chandola S, Agrawal S, Kushwaha SPS. 2017. Polarimetric SAR interferometry based modeling for tree height and aboveground biomass retrieval in a tropical deciduous forest. *Adv Space Res.* 60(3):571–586.
- Kumar S, Pandey U, Kushwaha SP, Chatterjee RS, Bijker W. 2012. Aboveground biomass estimation of tropical forest from envisat advanced synthetic aperture radar data using modeling approach. *J Appl Remote Sens.* 6(1):063588.
- Kumar S, Sara R, Singh J, Agrawal S, Kushwaha SPS. 2018. Spaceborne PolInSAR and ground-based TLS data modeling for characterization of forest structural and biophysical parameters. *Remote Sens Appl: Soc Environ.* 11:241–253.
- Lavalle M, Solimini D, Pottier E, Desnos YL. 2009. Forest Parameters Inversion Using Polarimetric and Interferometric SAR Data. *International Geoscience and Remote Sensing Symposium (IGARSS)*; July 12–17; Cape Town, South Africa. Vol. 4.
- Lee JS, Pottier E. 2009. *Polarimetric radar imaging - from basics to applications.* New York (NY): Taylor & Francis Group.
- Mangla R. 2015. Machine learning based regression model for forest aboveground biomass estimation using RISAT-1 PolSAR and TLS Lidar Data. *Andhra University.*
- Mangla R, Kumar S, Nandy S. 2016. Random forest regression modelling for forest aboveground biomass estimation using RISAT-1 PolSAR and terrestrial LiDAR Data. *Proc. SPIE 9879, Lidar Remote Sensing for Environmental Monitoring XV, 98790Q.*
- ME BP, Kumar S. 2021. PolInSAR decorrelation-based decomposition modelling of spaceborne multifrequency SAR data. *Int J Remote Sens.* 42(4):1398–1419.
- Mette T, Papathanassiou K, Hajnsek I. 2004. Biomass estimation from polarimetric SAR interferometry over heterogeneous forest terrain. *IGARSS'04. 2004 IEEE International Geoscience and Remote Sensing Symposium.* Vol. 1, p. 511–514.
- Moreira A, Prats-Iraola P, Younis M, Krieger G, Hajnsek I, Papathanassiou KP. 2013. A tutorial on synthetic aperture radar. *IEEE Geosci Remote Sens Mag.* 1(1):6–43.
- Neumann M. 2009. Remote Sensing of Vegetation Using Multi-Baseline Polarimetric SAR Interferometry: Theoretical Modeling and Physical Parameter Retrieval. *Université de Rennes I, France.* [accessed 2019 Sept 14]. <http://www.opensthesis.org/documents/Remote-sensing-vegetation-using-multi-600827.html>.
- Neumann M, Ferro-Famil L, Reigber A. 2010. Estimation of forest structure, ground, and canopy layer characteristics from multibaseline polarimetric interferometric SAR data. *IEEE Trans Geosci Remote Sens.* 48(3):1086–1104.
- O'Brien RM. 2007. A caution regarding rules of thumb for variance inflation factors. *Qual Quant.* 41(5): 673–690.
- Papathanassiou KP, Cloude SR. 2001. Single-baseline polarimetric SAR interferometry. *IEEE Trans Geosci Remote Sens.* 39(11):2352–2363.
- Rencz AN, Ryerson RA, editors. 1999. *Remote sensing of earth sciences-manual of remote sensing.* 3rd ed. New York (NY): American Society for Photogrammetry and Remote Sensing.
- Richards JA. 2009. *Remote sensing with imaging radar.* Berlin, Heidelberg: Springer.
- Santoro M. 2003. Estimation of biophysical parameters in Boreal Forests from ERS and JERS SAR Interferometry [Ph.D. dissertation]. Department of Radio and Space Science, Chalmers University of Technology.
- Santoro M, Eriksson L, Askne J, Schmillius C. 2006. Assessment of stand-wise stem volume retrieval in boreal forest from JERS-1 L-Band SAR backscatter. *Int J Remote Sens.* 27(16):3425–3454.
- Shafai SS, Kumar S. 2020. PolInSAR coherence and entropy-based hybrid decomposition model. *Earth Space Sci.* 7(10):1–17.
- Shataee S, Kalbi S, Fallah A, Pelz D. 2012. Forest attribute imputation using machine-learning methods and ASTER data: comparison of k -NN, SVR and random forest regression algorithms. *Int J Remote Sens.* 33(19):6254–6280.
- Stein A, Meer FVD, Gorte B. 1999. *Spatial statistics for remote sensing.* Dordrecht: Kluwer Academic Publishers.

- Stelmaszczuk-Górska MA, Urbazaev M, Schmillius C, Thiel C. 2018. Estimation of above-ground biomass over boreal forests in Siberia using updated in situ, ALOS-2 PALSAR-2, and RADARSAT-2 Data. *Remote Sens.* 10(10):1550.
- Tomar KS. 2015. Hybrid Polarimetric Decomposition for Aboveground Biomass Estimation Using Semi-Empirical Modelling. University of Twente.
- Tomar KS, Kumar S, Tolpekin VA. 2019. Evaluation of hybrid polarimetric decomposition techniques for forest biomass estimation. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 12(10):3712–3718.
- Trisasongko BH, Paull DJ, Griffin AL, Jia X, Panuju DR. 2019. On the relationship between the circumference of rubber trees and L-band waves. *Int J Remote Sens.* 40(16):6395–6417.
- van Leeuwen HJC. 2014. Multifrequency and multitemporal analysis of scatterometer radar data with respect to agricultural crops using the cloud model; 1991 Jun 3–6; Espoo, Finland. *IEEE.* Vol. 4, p. 1893–1897.
- Van Zyl JJ, Aari M, Kim Y. 2011. Model-based decomposition of polarimetric SAR covariance matrices constrained for nonnegative eigenvalues. *IEEE Trans Geosci Remote Sens.* 49(9):3452–3459.
- Vapnik VN. 2000. *The nature of statistical learning theory.* New York (NY): Springer.
- Yamaguchi Y, Moriyama T, Ishido M, Yamada H. 2005. Four-component scattering model for polarimetric SAR image decomposition. *IEEE Trans Geosci Remote Sens.* 43(8):1699–1706.
- Yamaguchi Y, Yajima Y, Yamada H. 2006. A four-component decomposition of POLSAR images based on the coherency matrix. *IEEE Geosci Remote Sens Lett.* 3(3):292–296.
- Zhang L, Zou B, Cai H, Zhang Y. 2008. Article in *IEEE geoscience and remote sensing letters.* *IEEE Geosci Remote Sens Lett.* 5(4):603–607.