Int J Appl Earth Obs Geoinformation 78 (2019) 175–188

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

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

# Object-based random forest modelling of aboveground forest biomass outperforms a pixel-based approach in a heterogeneous and mountain tropical environment

Eduarda M.O. Silveira[a,*], Sérgio Henrique G. Silva[b], Fausto W. Acerbi-Junior[a], Mônica C. Carvalho[a], Luis Marcelo T. Carvalho[a], Jose Roberto S. Scolforo[a], Michael A. Wulder[c]

[a] Forest Science Department (DCF), Federal University of Lavras, Lavras, Brazil
[b] Soil Science Department (DCS), Federal University of Lavras, Lavras, Brazil
[c] Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, Victoria, Canada

ARTICLE INFO

*Keywords:*
Landsat
Random forests
Spatial distribution
OBIA
Atlantic forest
AGB

ABSTRACT

The Brazilian Atlantic Forest is a highly heterogeneous biome of global ecological significance with high levels of terrestrial carbon stocks and aboveground biomass (AGB). Accurate maps of AGB are required for monitoring, reporting, and modelling of forest resources and carbon stocks. Previous research has linked plot-level AGB with environmental and remotely sensed data using pixel-based approaches. However, few studies focused on investigating possible improvements via object-based image analysis (OBIA) including terrain related data to predict AGB in topographically variable and mountainous regions, such as Atlantic forest in Minas Gerais, Brazil. OBIA is expected to reduce known uncertainties related to the positional discrepancy between the image and field data and forest heterogeneity, while terrain derivatives are strong predictors in forest ecosystems driving forest biomass variability. In this research, we compare an object-based approach to a pixel-based method for modeling, mapping and quantifying AGB in the Rio Doce basin, within the Brazilian Atlantic Forest biome. We trained a random forest (RF) machine learning algorithm using environmental, terrain, and Landsat Thematic Mapper (TM) remotely sensed imagery. We aimed to: (i) increase the precision of the AGB estimates; (ii) identify optimal variables that fit the best model, with the lowest root mean square error (RMSE, Mg/ha); (iii) produce an accurate map of the AGB for the study area, and subsequently (iv) describing the AGB spatial distribution as a function of the selected variables. The RF object-based model notably improved the AGB prediction by reducing the mean absolute error (MAE) from 28.64 to 20.95%, and RMSE from 33.43 to 20.08 Mg/ha, and increasing the $R^2$ (from 0.57 to 0.86) by using a combination of selected remote sensing, environmental, and terrain variables. Object-based modelling is a promising alternative to common pixel-based approaches to reduce AGB variability in topographically diverse and heterogeneous environments. Investigation of mapped outcomes revealed a decreasing AGB from west towards the east region of the Rio Doce Basin. Over the entire study area, we map a total of 195,799,533 Mg of AGB, ranging from 25.52 to 238 Mg/ha, following seasonal precipitation patterns and anthropogenic disturbance effects. This study provided reliable AGB estimates for the Rio Doce basin, one of the most important watercourses of the globally important Brazilian Atlantic Forest. In conclusion, we highlight that OBIA is a better solution to map forest AGB than the pixel-based traditional method, increasing the precision of AGB estimates in a heterogeneous and mountain tropical environment.

## 1. Introduction

The Atlantic Forest, a biome of global ecological significance, is one of the largest rainforests of the Americas, originally covering around 150 million ha, with highly heterogeneous environmental conditions (Ribeiro et al., 2009). The wide latitudinal, longitudinal, and elevation ranges present result in high levels of biodiversity, where flora and fauna encompass 1–8% of the world's total species (Silva and Tabarelli,

* Corresponding author.
*E-mail addresses:* dudalavras@hotmail.com, eduarda.silveira@posgrad.ufla.br (E.M.O. Silveira), sergio.silva@dcs.ufla.br (S.H.G. Silva), fausto@dcf.ufla.br (F.W. Acerbi-Junior), monicacanaan@gmail.com (M.C. Carvalho), passarinho@dcf.ufla.br (L.M.T. Carvalho), jscolforo@dcf.ufla.br (J.R.S. Scolforo), mike.wulder@canada.ca (M.A. Wulder).

https://doi.org/10.1016/j.jag.2019.02.004
Received 18 December 2018; Received in revised form 4 February 2019; Accepted 11 February 2019
0303-2434/ © 2019 Elsevier B.V. All rights reserved.

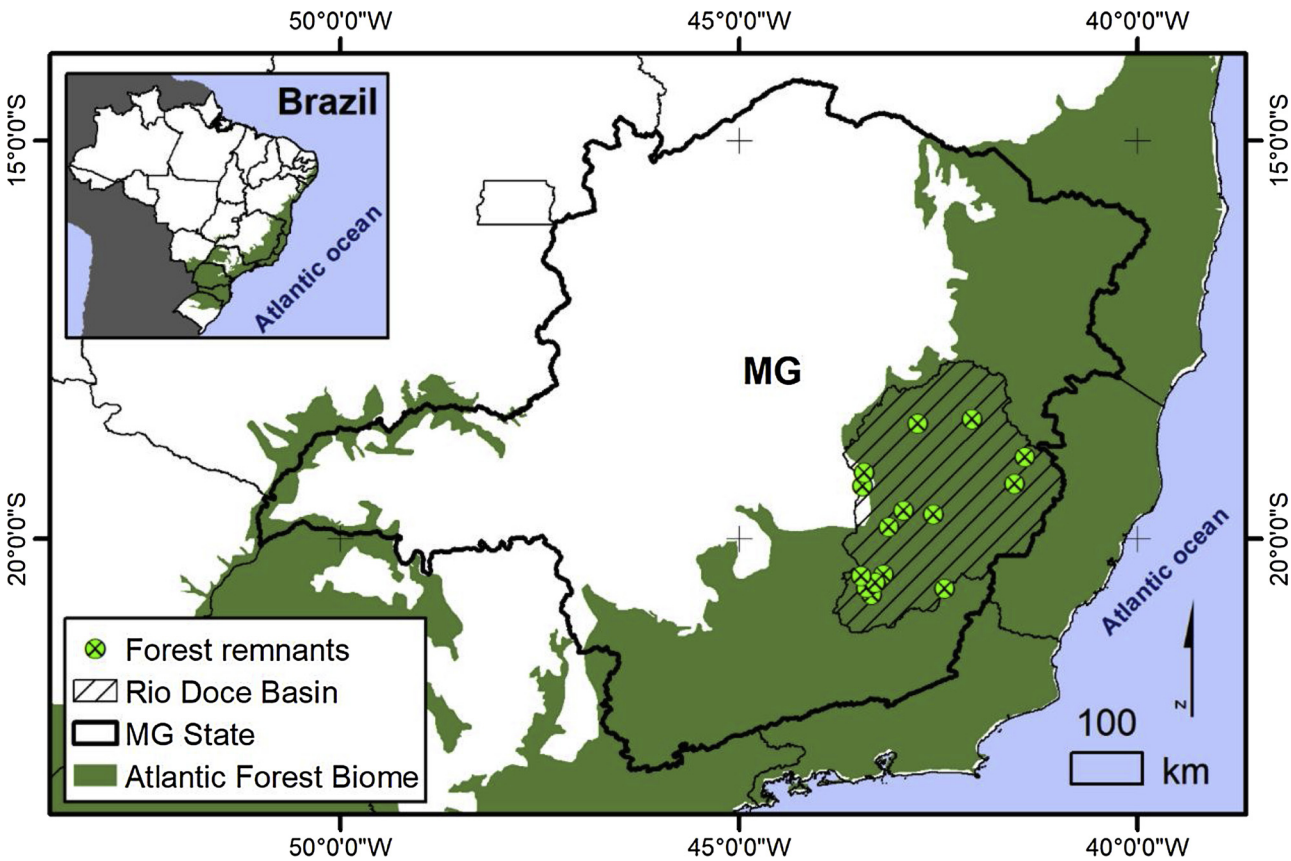


**Fig. 1.** Study area and forest remnants locations in the state of Minas Gerais (MG), Brazil.

2000) and high levels of terrestrial carbon stocks and biomass due to high tree mass (David et al., 2017; Scolforo et al., 2016). During the last 500 hundred years, the Atlantic Forest has been subject to land use change, with forests being replaced by croplands, cattle ranching and by Eucalyptus forest for cellulose and paper production (Colombo and Joly, 2010). When the area is converted to other land uses through burning and clearing, greenhouse gases are emitted (Houghton et al., 2009). Forest biomass stocks are a proxy of the quantity of these gases, hence, there is critical need for estimating the magnitude of Atlantic forest aboveground biomass (AGB) under scenarios of land use change for monitoring natural forests remnants and supporting climate change modeling studies (Houghton, 2005).

Spatially explicit estimates of AGB for the Brazilian Atlantic forest are rarely available, owing to large areas of complex forest, where extensive inventories are typically not available as such activity is time-consuming and expensive (Gardner et al., 2008). Sampling efforts with limited or poor spatial distribution of plots have resulted in significant geographical data gaps (Groves et al., 2002), making it particularly difficult to use this information to quantify and delineate AGB distribution accurately, thus resulting in uncertainties in forest ecosystem carbon cycle models (Xu et al., 2015).

Over the past two decades, a large number of studies have shown that AGB forest biomass can be retrieved using various predictor variables derived from remotely sensed images (Baccini et al., 2008; Galidaki et al., 2017; Hu et al., 2016; Su et al., 2016; Timothy et al., 2016) and by combining remote sensing with environmental data (Fayad et al., 2016). Up to now, Landsat images are the most commonly used medium spatial resolution data in AGB studies (Andersen et al., 2012; Powell et al., 2010; Zhu and Liu, 2015). Assisted by free and open access to the data (Wulder et al., 2012), the spatial resolution (30 m) of Landsat allows for capture of forest management activities, as well as robust calibration that enables consistent and predictable application of algorithms (Cohen and Goward, 2004). However, the estimates are mainly made following pixel-based approaches, which present some disadvantages when compared with object-based image analysis (OBIA). For example, using an object unit of analysis can reduce the uncertainty of positional discrepancy between the image and field data as there is a higher probability that a field plot is located in an object rather than in a single pixel; the local noise and heterogeneity can be effectively reduced when using a “pure” object; (Addink et al., 2007; Zhang et al., 2018). Additional object-based spatial features can be extracted for each object (Wulder and Seemann, 2003), offering potential to improve AGB estimation by including the spatial attributes in the model (Zhang et al., 2018). Studies applying object-based methods have reported notable improvements in AGB estimation accuracies (Kajisa et al., 2009; Mareya et al., 2018; Zhang et al., 2018).

Terrain data derived from digital elevation models and its derivative maps, such as slope, topographic wetness index (Beven and Kirkby, 1979) and elevation derivatives (e.g., landscape positon, curvature), are also reliable predictors in forest ecosystems which relate soil type, soil water content, soil nutrients and light availability that directly influence vegetation (Luizão et al., 2004; Matasci et al., 2018). For example, Kopecký and Čížková (2010) found a strong correspondence of topographic wetness index to plant distribution, suggesting this variable can be useful for vegetation studies. It is expected that AGB stock of the arboreal vegetation is influenced by elevation and its derivative measurements since such variables are correlated to climatic conditions, mainly in hilly and mountainous regions, such as the Atlantic forest in Minas Gerais (MG) state, Brazil. However, application examples using terrain data to estimate forest AGB in Brazil are rare, requiring more tests.

Thus, we explored the question of whether OBIA together with remote sensing, environmental and terrain data, can influence the predictive quality of forest AGB estimation in hilly and mountainous

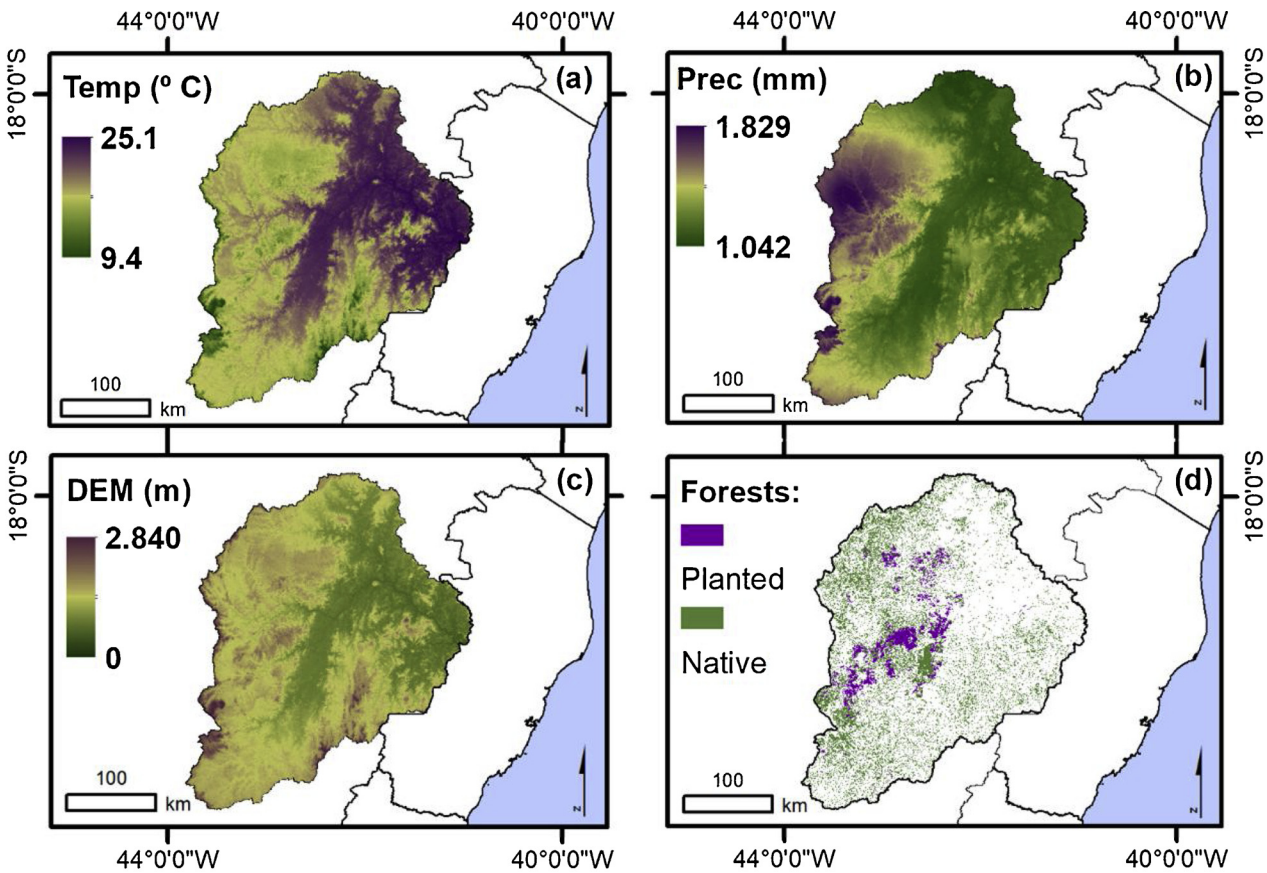


**Fig. 2.** (a) Mean Annual Temperature (b) and Precipitation (c), Digital Elevation Model (DEM) (d), and distribution of forest types within the Rio Doce Basin, state of Minas Gerais, Brazil. Temp – Mean Annual Temperature; Prec - Mean Annual Precipitation.

regions by reducing variability to improve modeling results. Hence, this study aims focusses on: (i) increasing the precision of the AGB estimates; (ii) finding the optimal variables among environmental, terrain, and remote sensing data to predict AGB; (iii) producing an accurate map of Atlantic forest AGB in Rio Doce basin, in Minas Gerais state, Brazil; and subsequently (iv) describing the AGB spatial distribution as a function of the selected variables. As an outcome, we aim to offer insights and options to improve the AGB estimates in high heterogeneity tropical forests, such as Rio Doce basin, helping guide the selection of the appropriate analysis unit that reduces variability and support for selecting variables that best capture AGB variability.

## 2. Study area and field measurements

The Rio Doce Basin has a total area of approximately 84,000 km$^2$, of which 14% lies in the state of Espírito Santo (ES), whilst the remaining 86% are found in Minas Gerais state (MG), Brazil (Guevara et al., 2018). The study area covers a portion of the Rio Doce Basin in MG territory (Fig. 1), where climate is tropical humid with mean annual temperature between 9.4 °C and 25.1 °C (Fig. 2a) and average annual precipitation close to 1500 mm (Fig. 2b). The regional relief is classified as hilly and mountainous (Fig. 2c) (Hijmans et al., 2005). Approximately 98% of the basin area belongs to the Atlantic Forest biome, composed of semi-deciduous forest (19.90%), grassland (0.80%), rupestrian fields - "campo rupestre" (0.17%), shrub savannas (0.01%), pinus (0.02%) and eucalyptus (3.64%), totalizing 1.417,970 ha of native vegetation and 259,537 ha of planted vegetation (Carvalho et al., 2006) (Fig. 2d).

We obtained the field measurements from 188 plots (10 × 100 m) from 15 forest remnants, which are part of the project: "Forest Inventory of Minas Gerais", conducted by Federal University of Lavras (UFLA) from 2006 to 2008 (Fig. 3). This project generated abundant information regarding the quality of forest remnants and information related to the determination of AGB and the ongoing monitoring of forests development throughout permanent plots.

During the field surveys, all trees were tagged and identified. The diameter at breast height (DBH) and the total height of the trees with a minimum DBH of 5 cm were measured. The trees used to determine the AGB were all from destructive sampling campaigns. Scaling was conducted on 674 trees, distributed over 456 species. The destructive sampling was conducted by first measuring the DBH, total height and stem height prior to felling the tree. The tree was then georeferenced, cut down and had its height measured with a measuring tape. The tree was then scaled using Huber's method. Wood discs were sampled in each tree to determine dry weight. We used the amount of AGB per plot converted to hectares (Mg/ha) to model the spatial distribution of AGB across the Rio Doce basin, with additional details provided in Scolforo et al. (2015). The plots descriptive statistics are presented in Table 1.

## 3. Material and methods

A summary of our methods to derive the forest AGB maps is provided in Fig. 4. We described the data and individual steps taken in detail below.

### 3.1. Data collection and pre-processing

We downloaded nineteen (19) bioclimatic variables with 1 km$^2$ spatial resolution (Table 2) from WorldClim (Hijmans et al., 2005). The bioclimatic variables represent annual trends (e.g., mean annual temperature, annual precipitation) seasonality (e.g., annual range in temperature and precipitation) and extreme environmental factors. Studies in tropical forests have shown the importance of climatic variables in

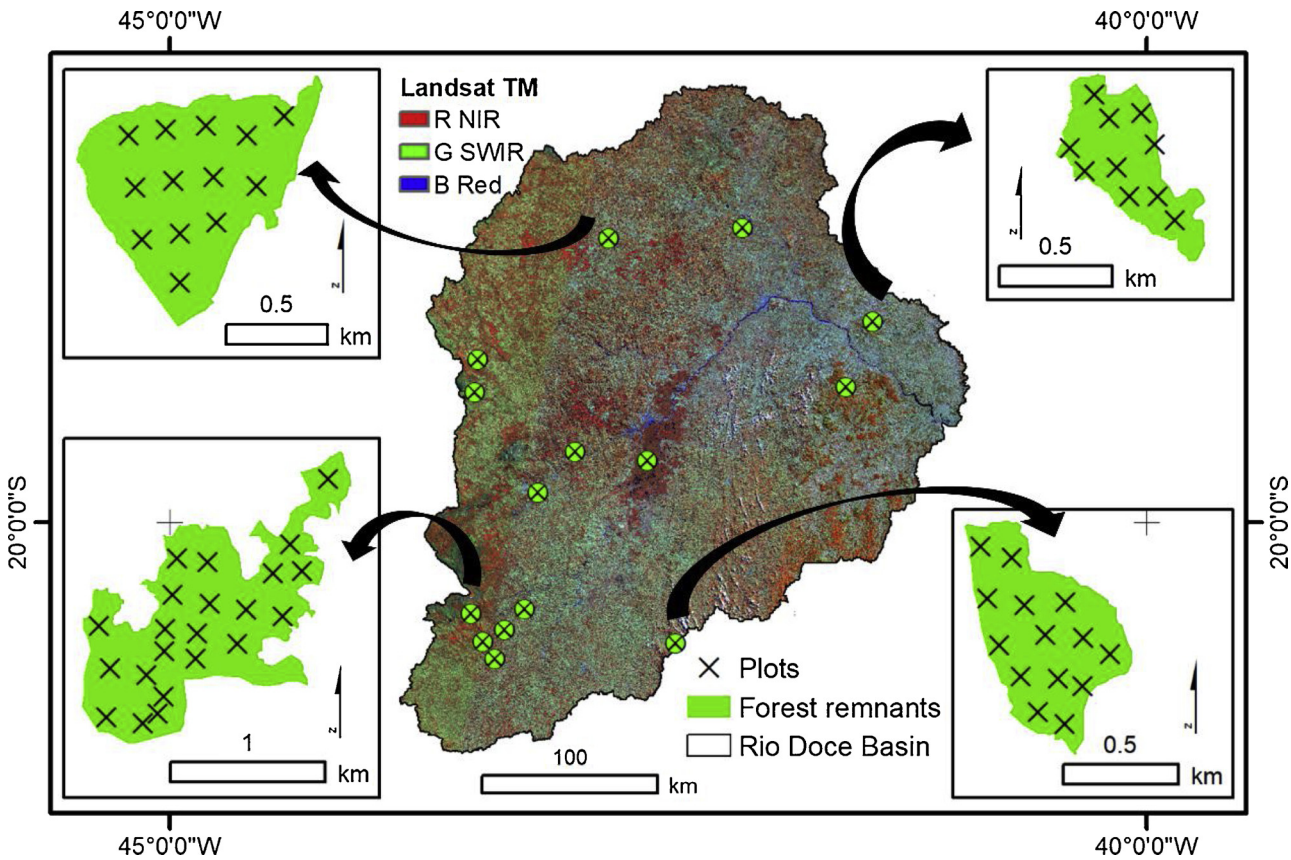


**Fig. 3.** Spatial distribution of forest remnants within the Rio Doce Basin, detailing the spatial distribution of plots found in four forest remnants. Image background is a Landsat TM false colour composite with NIR, SWIR and Red bands in red, green and blue colours, respectively (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

**Table 1**
Descriptive statistics of the measured AGB (Mg/ha).

| Minimum | Mean | Maximum | Standard Deviation |
|---|---|---|---|
| 11.05 | 93.72 | 292.96 | 59.06 |

determining biomass (Feng et al., 2011; Vieilledent et al., 2016; Kumar and Mutanga, 2017).

To support the terrain analysis, a digital elevation model (DEM) of 30 m of spatial resolution developed from the Shuttle Radar Topography Mission (SRTM) was obtained. The data was pre-processed to generate a depressionless DEM prior to the calculation of terrain variables. SRTM and ASTER-derived DEMs are post-processed elevation datasets that are frequently used for a wide range of applications due to their high level of detail and wide coverage (Nikolakopoulos et al., 2006). We chose SRTM supported by previous studies that have shown desirable qualities and a superior accuracy when compared to ASTER-derived DEMS (Jarvis et al., 2004). Recent studies confirmed that surface processes such as hydrological, environmental, and pedological modeling can be appropriately studied when using SRTM-DEM by generating topographical indices (Daleles et al., 2008; Liu et al., 2012; Collard et al., 2014; Guitet et al., 2015; Samuel-Rosa et al., 2015).

We obtained eight (8) Landsat-5 TM images to represent the study area (Table 3). We downloaded the images from the free and open archive of the United States Geological Survey for Earth Observation and Science (USGS), Earth Resources Observation and Science (EROS) Center, prepared as geometrically corrected and to surface reflectance, Collection 1, Level 2 (Young et al., 2017). For the year 2007 (the same year of the inventory data collection in the study area), we selected one image date in which the scene was entirely without clouds. The images were mosaicked and resampled to Albers Equal Area Conic projection due to area and shape preserving characteristics of this projection (Duro et al., 2012).

### 3.2. Deriving variables

We derived 14 terrain variables from the DEM (Table 4) using the software SAGA GIS (Conrad et al., 2015). This software applies several algorithms to the DEM, producing different terrain variables to portray features, such as slopes, curvature, vertical distance of each place on the landscape above water channels, shading, places proximity to water channels, shading, and accumulation zones. The variables chosen in this study have been widely used for predictions of soil properties (Forkuor et al., 2017; Malone et al., 2017), but less so for tropical forest AGB predictions. Thus, by adopting them, it is possible to evaluate if these terrain variables effectively contribute to the modelling and predicting of AGB in tropical forests.

The terrain variables selected was based on previous works on tropical environments (Asner et al., 2009; Ferry et al., 2010; Xu et al., 2015) and due to their individual capacity for representing terrain-related data that can be associated with vegetation types, accumulation of water, slope gradient, elevation, erosion, shading, etc. It is expected that such information may help predict vegetation patterns and, hence, AGB throughout the study area. For example, the MRVBF map identifies valley bottoms that are important hydrologic and geomorphic features. Thus, distinguishing between hillslopes and valley bottoms is necessary because of the substantial difference in hydrological processes in the two landforms (Gallant and Dowling, 2003), which, in turn, influence vegetation patterns.

Using the six pre-processed optical bands of Landsat-5 TM, we computed a selection of spectral indices (11) (Table 5), shown in

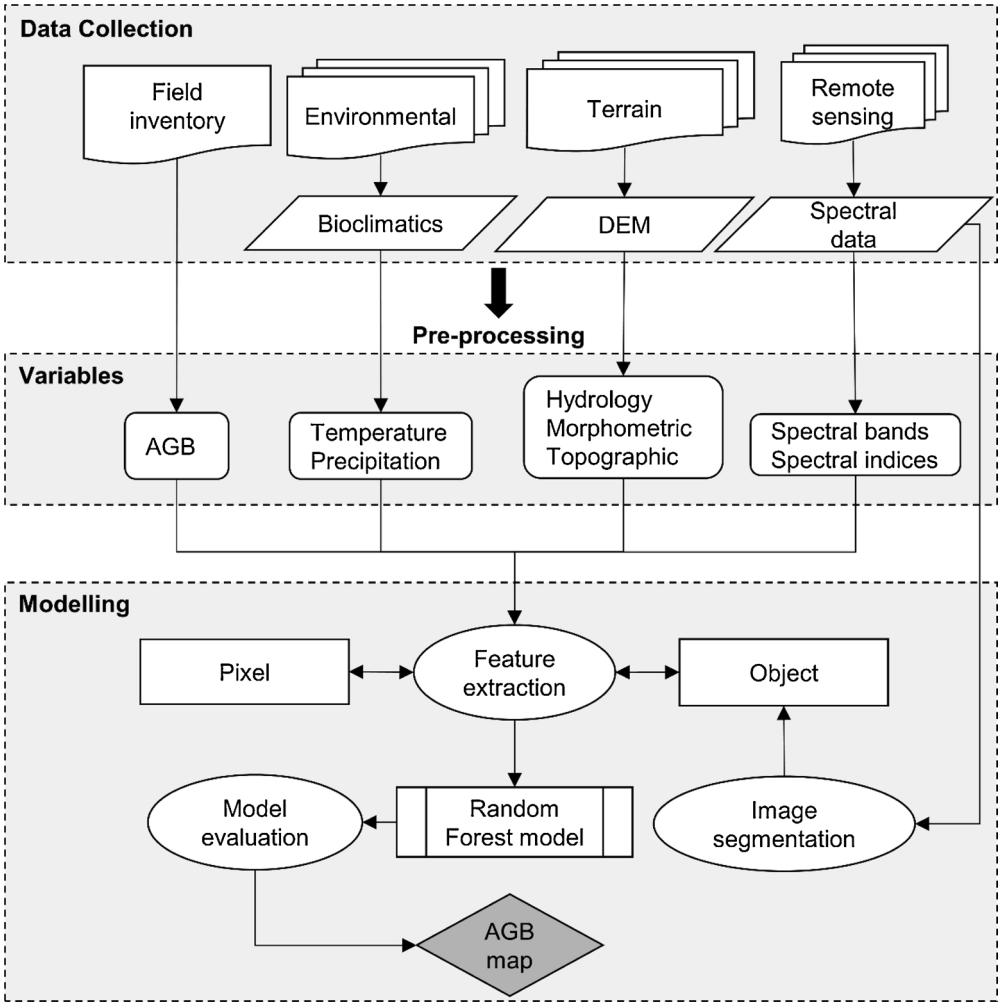


**Fig. 4.** Aboveground biomass (AGB) mapping procedure.

**Table 2**
Environmental variables acquired from WorldClim.

| Variables | Description |
|---|---|
| BIO 1 | Annual Mean Temperature (ºC) |
| BIO 2 | Mean Diurnal Range (Mean of monthly (ºC) |
| BIO 3 | Isothermality (BIO2/BIO7) (* 100) (ºC) |
| BIO 4 | Temperature Seasonality (standard deviation *100) (ºC) |
| BIO 5 | Max Temperature of Warmest Month (ºC) |
| BIO 6 | Min Temperature of Coldest Month (ºC) |
| BIO 7 | Temperature Annual Range (BIO5-BIO6) (ºC) |
| BIO 8 | Mean Temperature of Wettest Quarter (ºC) |
| BIO 9 | Mean Temperature of Driest Quarter (ºC) |
| BIO 10 | Mean Temperature of Warmest Quarter (ºC) |
| BIO 11 | Mean Temperature of Coldest Quarter (ºC) |
| BIO 12 | Annual Precipitation (mm) |
| BIO 13 | Precipitation of Wettest Month (mm) |
| BIO 14 | Precipitation of Driest Month (mm) |
| BIO 15 | Precipitation Seasonality (Coefficient of Variation) (mm) |
| BIO 16 | Precipitation of Wettest Quarter (mm) |
| BIO 17 | Precipitation of Driest Quarter (mm) |
| BIO 18 | Precipitation of Warmest Quarter (mm) |
| BIO 19 | Precipitation of Coldest Quarter (mm) |

**Table 3**
Landsat scenes and acquisition dates used to calculate spectral indices.

| Landsat Path/Row | Acquisition date |
|---|---|
| 216/73 | 27 August 2007 |
| 216/74 | 27 August 2007 |
| 217/72 | 01 July 2007 |
| 217/73 | 14 May 2007 |
| 217/74 | 14 May 2007 |
| 217/75 | 02 August 2007 |
| 218/73 | 25 August 2007 |
| 218/74 | 25 August 2007 |

previous research to be sensitive to vegetation characteristics (Galidaki et al., 2017; Timothy et al., 2016). Spectral indices have been used to estimate AGB, but there was no universal index applied to all vegetation conditions (Gómez et al., 2014). Identifying the spectral indices that are most strongly correlated with a specific biophysical parameter in a given study area is difficult, especially in tropical environments due to their complex forest stand structure (Lu et al., 2004).

Although vegetation indices such as NDVI, EVI, SAVI and MSAVI have been proposed in previous research to estimate biomass (Lu et al., 2016), some researchers in tropical forests found that spectral indices including near-infrared (NIR) wavelength presented weaker relationships with biomass than those including shortwave infrared (SWIR) wavelength (Lu, 2005). Thus, we also included in the analysis the NBR, NBR2 and NDMI spectral indices to evaluate their contribution to predict AGB in tropical environments. The Tasselled Cap components have been widely used to characterize vegetation conditions (Crist and Cicone, 1984). These indices can measure the presence and density of green vegetation, overall reflectance, soil moisture content, and vegetation density (structure) (Cohen and Goward, 2004).

**Table 4**
Terrain variables derived from digital elevation model (DEM) used to model AGB.

| Name | Description |
|---|---|
| Digital Elevation Model (DEM) | Represents the elevation (m) of each pixel of the study area |
| Aspect | Slope orientation in relation to North |
| Analytical Hillshade | Represents the angle between sunlight and the terrain surface |
| Channel Network Base Level (CNBL) | Distance to a channel network base level |
| Convergence Index | Relief representation taking into account convergent and divergent areas |
| Cross-sectional Curvature | Terrain plan curvature |
| Longitudinal Curvature | Terrain profile curvature |
| LS Factor | Slope-length factor |
| Multiresolution Index of Ridge Top Flatness (MRRTF) | Identifies flat and high elevation places |
| Multiresolution Index of Valley Bottom Flatness (MRVBF) | Identifies flat and low elevation places |
| Relative Slope Position | Position of each cell in relation to the top and bottom areas |
| Topographic Position Index | Comparison of each pixel to its neighbour values |
| Topographic Wetness Index | Shows the feasibility of a site to accumulate water |
| Valley Depth (VD) | Vertical distance in relation to the valley |
| Vertical Distance to Channel Network (VDCN) | Vertical distance of each pixel to the channel network |

### 3.3. Development of AGB models

#### 3.3.1. Feature extraction

We used the AGB (in Mg/ha) from sample plots as dependent variable and the entire group of environmental, terrain, and remote sensing data as independent variables. We tested two units of analysis: pixel- and object-based. The data were resampled to 100 m × 100 m (1 ha), so all the data had the same cell size.

In the pixel-based process, we conducted the feature extraction considering co-located pixel at the plot location. In the object-based process, we applied an automatic procedure using the eCognition software for object-based image analysis. We used a multiresolution segmentation algorithm (Baatz and Schäpe, 2000) set with three parameters: shape, compactness, and scale. These parameters control the shape, size, and spectral variation of segmented image objects. The shape parameter was set to 0.1, compactness was set to 0.5, and scale parameter was set to 30. Following Duro et al. (2012), we used a trial and error approach to determine the appropriate scale parameter (SP) based on a visual assessment of segmentation suitability, resulting in an average object size of 15,000 $m^2$. We conducted the feature extraction considering the average values of the environmental, terrain, and remote sensing variables inside the objects to link with the AGB (Mg/ha) from sample plots.

#### 3.3.2. Random forests algorithm

In the Random Forests algorithm (Breiman, 2001), a large number of trees are grown with the root node containing a different bootstrap sample of the data with the same number of cases as the original data. At each node, splitting is performed using a randomly selected subset of the predictor variables. Modelling using the Random Forests algorithm is less sensitive to noise in the training data and tends to result in more accurate models (Baccini et al., 2008). We used randomForest Package (Liaw and Wiener, 2002) available in the R software (R Core Team, 2014).

Approximately 70% of the data (132 plots) were randomly selected and used to fit the model. The remaining 30% data (56 plots) were used to test the model performance. The random sampling was defined to preserve the overall class distribution of the data. The number of decision trees (Ntree) was set to 10,000. In order to improve the AGB model parsimony and predictive performance of AGB, we used the OOB (out of bag) error as minimization criterion. We identified the smallest number of variables with the lowest mean square error (MSE) employing a backward feature elimination approach (Díaz-Uriarte and Alvarez de Andrés, 2006). The least important variables returned by Random Forests was successively eliminated. We first take into account the measures of importance (Random Forests variables importance) to obtain an initial variable ranking and then proceeded with an iterative backward elimination of the least important variables. In each iteration, we eliminated the least important variables and a new Random Forests

**Table 5**
Summary of spectral bands and indices used to model aboveground biomass (AGB) in this study. Coefficients for the three basic Tasselled Cap indices (brightness, greenness and wetness) are shown as B, G, and W, respectively.

| RS data | Name | Formula | Reference |
|---|---|---|---|
| EVI | Enhanced Vegetation Index | $G\frac{NIR-Red}{NIR+C1*Red-C2*Blue+L}$ | (Huete et al., 2002) |
| SAVI | Soil Adjusted Vegetation Index | $\frac{NIR-Red}{NIR+Red}(1+l)$ | (Huete, 1988) |
| MSAVI | Modified Soil Adjusted Vegetation Index | $\frac{2NIR+1-\sqrt{(2NIR+1)^2-8(NIR-Red)}}{2}$ | (Rondeaux et al., 1996) |
| NBR | Normalized Burn Ratio | $\frac{NIR-SWIR2}{NIR+SWIR2}$ | (Miller and Thode, 2007) |
| NBR2 | Normalized Burn Ratio 2 | $\frac{SWIR1-SWIR2}{SWIR1+SWIR2}$ | (Miller and Thode, 2007) |
| NDMI | Normalized Difference Moisture Index | $\frac{NIR-SWIR1}{NIR+SWIR1}$ | (Wilson and Sader, 2002) |
| NDVI | Normalized Difference Vegetation Index | $\frac{NIR-Red}{NIR+Red}$ | (Tucker, 1979) |
| TCB | Tasselled Cap Brightness | $B1*Blue+B2*Green+B3*Red+B4*NIR+B5*SWIR1+B6*SWIR2$ | (Crist and Cicone, 1984) |
| TCG | Tasselled Cap Greenness | $G1*Blue+G2*Green+G3*Red+G4*NIR+G5*SWIR1+G6*SWIR2$ | (Crist and Cicone, 1984) |
| TCW | Tasselled Cap Wetness | $W1*Blue+W2*Green+W3*Red+W4*NIR+W5*SWIR1+W6*SWIR2$ | (Crist and Cicone, 1984) |
| TCA | Tasselled Cap Angle | $\tan^{-1}\frac{TCB}{TCG}$ | (Gómez et al., 2014) |

was built by training with the remaining features for the assessment of MSE. Besides obtaining the best overall predictive accuracy, variables selection allowed us to simplify the modelling process and identify both the minimum number of features that offer the best predictive accuracy (Ismail and Mutanga, 2010) and the most important variables for AGB prediction.

To compare the performance between pixel- and object-based approaches, we computed the adjusted coefficient of determination ($R^2$, in %), which relates the part of the observed variability that is explained by the model, and adjusts for the number of terms in a model; the Root-Mean-Square Error (RMSE, in Mg/ha; Equation 1), which measures the average difference between values predicted by the model and observations; the Mean Absolute Error (MAE, in %; Equation 2), which indicates an average over or underestimation bias by the model (Vieilledent et al., 2016); and by visual analysis of residuals graphs.

$$\mathrm{RMSE}\left(\frac{\mathrm{Mg}}{\mathrm{ha}}\right)=\sqrt{\frac{\sum_{i=1}^{n}(E_i-M_i)^2}{n}} \quad (1)$$

$$\mathrm{MAE}\ (\%)=\frac{1}{n}\sum_{i=1}^{n}\frac{M_i-E_i}{M_i}100 \quad (2)$$

Where $M_i$ and $E_i$ are measured and estimated AGB values (Mg/ha), respectively; and n corresponds to the number of plots.

### 3.4. Deriving the aboveground biomass map

We generated two maps of AGB for Rio Doce basin. In both maps, we used a forest mask (Carvalho et al., 2006) considering the semi-deciduous forest class to predict the AGB only in forest remains covered by this land cover class.

In the pixel-based approach, we created continuous cells with dimensions of 1 ha (100 × 100 m) for the entirety of the study area. In each cell containing the values of the selected variables, we applied the Random Forests regression model to predict the AGB. In the object-based approach, in each object containing the average values of the selected variables, we applied the Random Forests regression model to predict the AGB.

## 4. Results

### 4.1. Model performance

According to the results, the Random Forests model using the object-based unit analysis performed better compared to the pixel-based unit. The Random Forests object-based model significantly improved the AGB prediction by both reducing the MAE and RMSE (RMSE decreased from 33.43 to 20.08 Mg/ha and MAE reduced from 28.64 to 20.95%) and increasing the $R^2$ (from 0.57 to 0.86). The increasing on the performance of Random Forests object-based model is also highlighted in the scatter plot graphics, presented randomly distributed residuals with no trends (Fig. 5).

In the object-based RF model, the analyses between calculated AGB in sample plots and predictors (environmental, terrain, and vegetation indices) showed that 6 variables (Fig. 6a), out of the entire set of 44 variables, were the minimum number of features that offered the best predictive accuracy. The selected variables with their respective Increment in Mean Square Error (% IncMSE) were: Valley Depth (VD-34.81%), Annual Precipitation (BIO12-30.66%), Annual Mean Temperature (BIO1-23.77%), Tasselled Cap Wetness (TCW-21.80%), Multiresolution Index of Valley Bottom Flatness (MRVBF-19.91%) and Digital Elevation Model (DEM-18.33%).

In the pixel-based RF model, the features selected to best predict the AGB were 4: environmental (2) and terrain-related (2) (Fig. 6b). The selected variables with their respective % IncMSE were: Channel Network Base Level (CNBL, 32.79%), Minimum Temperature of Coldest Month (BIO6, 29.27%), Valley Depth (VD, 28.93%) and Mean Temperature of Coldest Quarter (BIO11, 28.81%).

The spatial distribution of the selected variables from the objected-based RF model and pixel-based RF model within the study area is presented in Figs. 7 and 8, respectively.

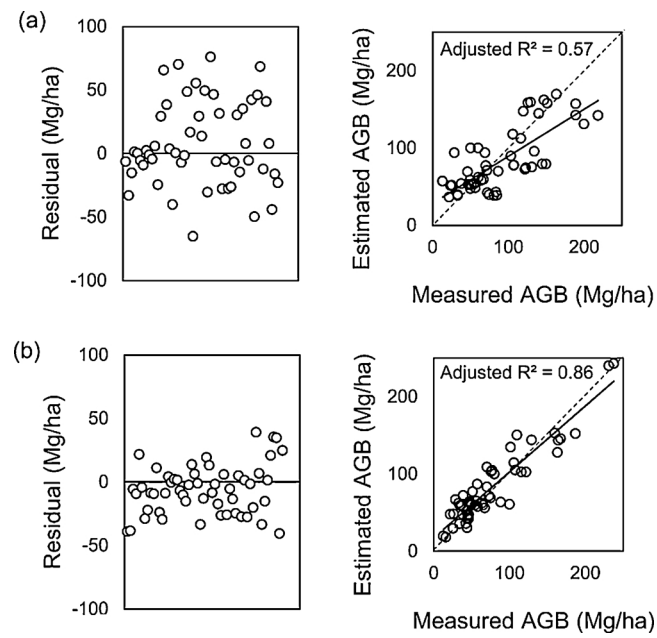


**Fig. 5.** Residuals analysis and scatter plots of measured value versus estimated value by (a): pixel-based Random Forests (RF) model and (b) object-based RF model. A 1:1 line (black, dashed) is provided for reference.

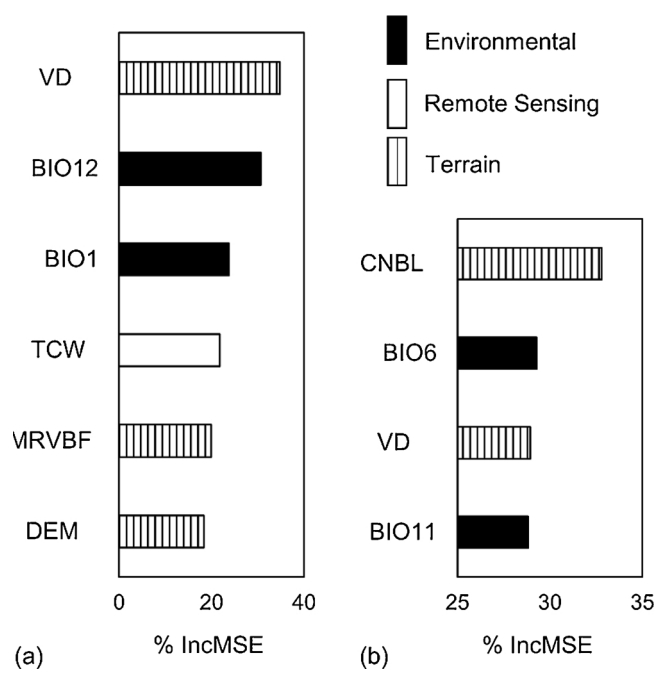


**Fig. 6.** The importance of the variables for the aboveground biomass (AGB) estimation in this study measured using: (a) Random Forests object-based approach and (b) Random Forests pixel-based approach. TCW - Tasselled Cap Wetness; MRVBF – Multiresolution Index of Valley Bottom Flatness; DEM – Digital Elevation Model; CNBL – Channel Network Base Level; VD – Valley Depth; BIO1 - Annual Mean Temperature; BIO6 - Minimum Temperature of Coldest Month; BIO11 - Mean Temperature of Coldest Quarter; BIO12 - Annual Precipitation.

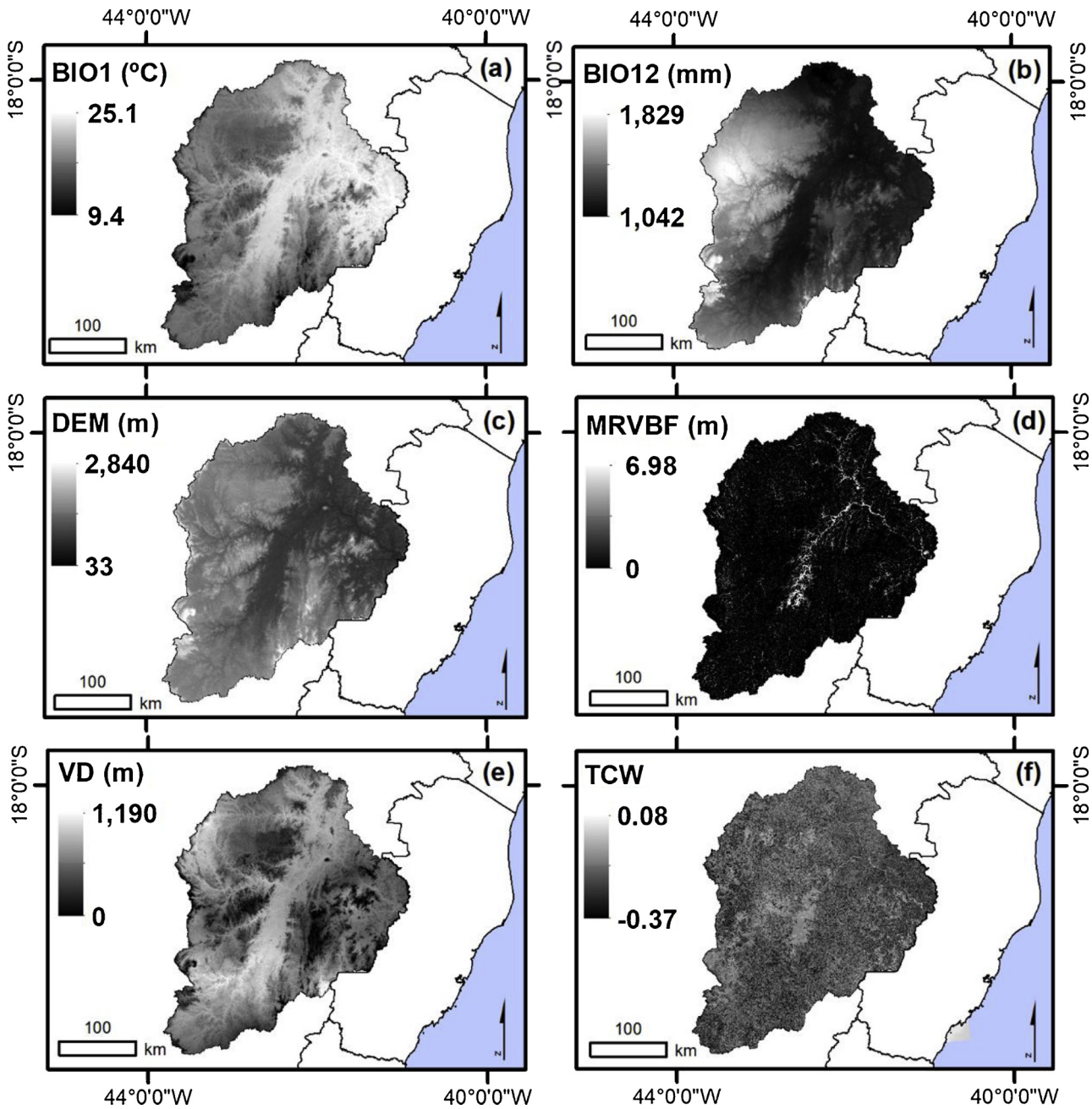


**Fig. 7.** Selected features that fit the best model using the object-based approach: (a) BIO1 - Annual Mean Temperature, (b) BIO12 - Annual Precipitation, (c) DEM – Digital Elevation Model, (d) MRVBF - Multiresolution Index of Valley Bottom Flatness, (e) VD – Valley Depth, (f) TCW – Tasselled Cap Wetness.

### *4.2. AGB maps in Rio Doce Basin*

Fig. 9 shows the maps of forest AGB spatial distribution from the Rio Doce Basin, divided into five (5) water resources planning and management units (DO1, DO2, DO3, DO4 and DO5), which were obtained by both object-based and pixel-based estimation methods.

Both the object-based and pixel-based methods presented similar spatial patterns relating the distribution of forest biomass, revealing a decrease in AGB from the west (DO1, DO2, DO3) towards the east (DO4 and DO5) regions. This is primarily attributable to the variations in the precipitation regime, as indicated in Fig. 7b (spatial distribution of BIO12, annual precipitation). The western portion of the study area, representing by DO1, DO2 and DO3, holds the forest remnants with the highest mean AGB, ranging from 96.61 to 100.68 Mg/ha (object-based model) and 86.28 to 98.26 Mg/ha (pixel-based model). In the eastern portion of the Rio Doce basin (DO4 and DO5) lies the regions with lower AGB. These areas have experienced anthropogenic disturbances, such as exploitation of vegetation (Carvalho et al., 2006), leading to few forest remnants in this region (see Fig. 2d). The total estimated AGB is 195,799,533 Mg, ranging from 25.52 Mg/ha to 238.26 Mg/ha (object-based method) and 178,967,656 Mg, ranging from 20.04 Mg/ha to 167.72 Mg/ha (pixel-based method) (Table 6).

## 5. Discussion

### *5.1. Object-based versus pixel-based method to model AGB*

Research aiming to predict AGB often utilize pixel level analysis instead of object-based approaches. OBIA analyses instead of pixels can reduce the “salt-and-pepper” effect in heterogeneous landscapes and

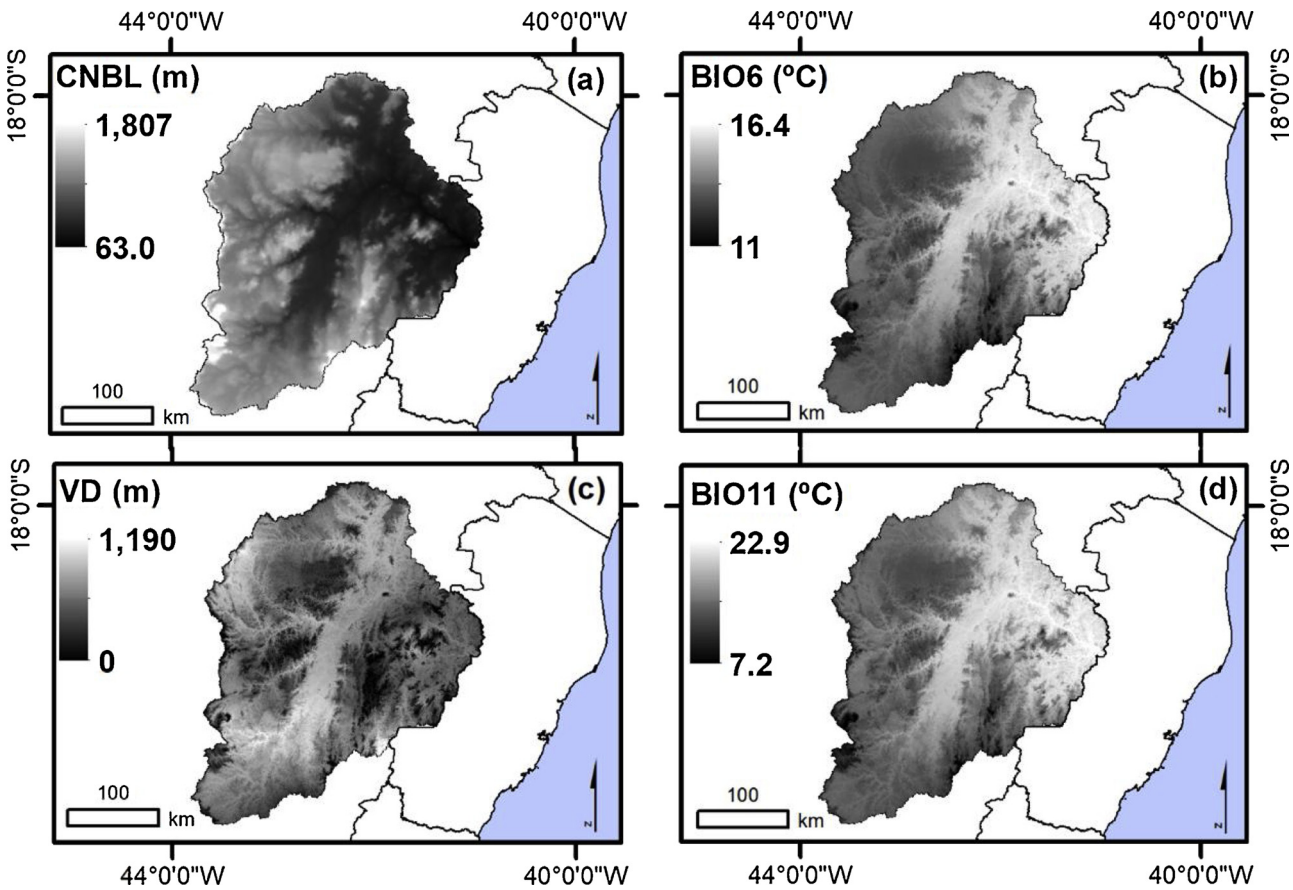


**Fig. 8.** Selected features that fit the best model using the pixel-based approach: (a) CNBL - Channel Network Base Level, (b) BIO6 - Minimum Temperature of Coldest Month, (c) VD - Valley Depth, (d) BIO11 - Mean Temperature of Coldest Quarter.

enhance the analysis accuracy in these environments (Dronova, 2015). OBIA has been widely used in image classification (Blaschke, 2010; Silveira et al., 2017; Wu et al., 2015) and change detection analysis (Chen et al., 2012; Silveira et al., 2018a,b,c), however, its applications in AGB modelling are sparse (Timothy et al., 2016). Also, there is a higher probability that a field plot is located in an object rather than in a single pixel, which can in turn be capitalized upon using an object-based analysis (Addink et al., 2007).

Here, we compared the accuracy of object-based and pixel-based methods in the estimation of Atlantic forest AGB using environmental, remote sensing, and terrain data. The accuracy achieved with the object-based method using the mean of 6 variables BIO1 - Annual Mean Temperature, BIO12 - Annual Precipitation, DEM – Digital Elevation Model, MRVBF - Multiresolution Index of Valley Bottom Flatness, VD – Valley Depth and, TCW – Tasselled Cap Wetness – was compared to that achieved with the pixel-based method using the singular values of CNBL - Channel Network Base Level, BIO6 - Minimum Temperature of the Coldest Month, VD - Valley Depth, and BIO11 - Mean Temperature of the Coldest Quarter. Our results demonstrated that, using the object-based approach, the AGB was notably improved, with $R^2$ reaching 0.86, RMSE equal to 20.08 Mg/ha and 20.95% of MAE. Thus, as we hypothesized, OBIA was a better solution to map forest AGB than the pixel-based traditional method.

Studies have indicated that several textural attributes extracted from image objects showed a significant relationship with forest biomass (Kajisa et al., 2009). Although in our study we have only used the spectral information of data, and did not exploit the spatial context from OBIA, the mean values of the selected variables inside the objects were enough to reduce the heterogeneity within objects and improve the AGB relationship with the independent variables.

Although the object-based approach provides good results in biomass modeling, the method has limitations related to image segmentation. The selection of the appropriate scale parameter (SP) to generate objects for linking to the AGB plot measurement is a difficult task. Finding an appropriate SP is a challenge, due to erroneous set impacts the modeling results. Different levels of segmentation result in different accuracy values for estimation biomass (Addink et al., 2007). Small SP generates more homogenous objects than higher values. On the contrary, applying higher values can produce heterogeneous objects, which might fail to represent the true structure of the plots (Zhang et al., 2018). Starting with individual pixels and increasing the object size, the predictions improve until an optimum value is reached. After that point, increasing object size results in worse predictions (Addink et al., 2007).

### 5.2. Selection of suitable variables for AGB modeling

Selection of suitable variables is important for developing biomass estimation models. Here, we combined environmental, remote sensing, and terrain data for AGB modeling. To allow for a better interpretation of the relationship among the spatial distribution of selected variables and the improved AGB map (object-based one), we divided the independent and predictive dependent variables into 2 classes: low and high, and then, combined each one, generating four (4) classes (Table 7): HH (Both high values), HL (High predictive AGB and low independent variable value), LH (low predictive AGB and High independent variable value), and LL (both low values). We used Jenks natural breaks (de Smith et al., 2018) interval to partition the data. Natural breaks classes are based on natural groupings inherent in the data. The Jenks method determines class breaks by minimizing within

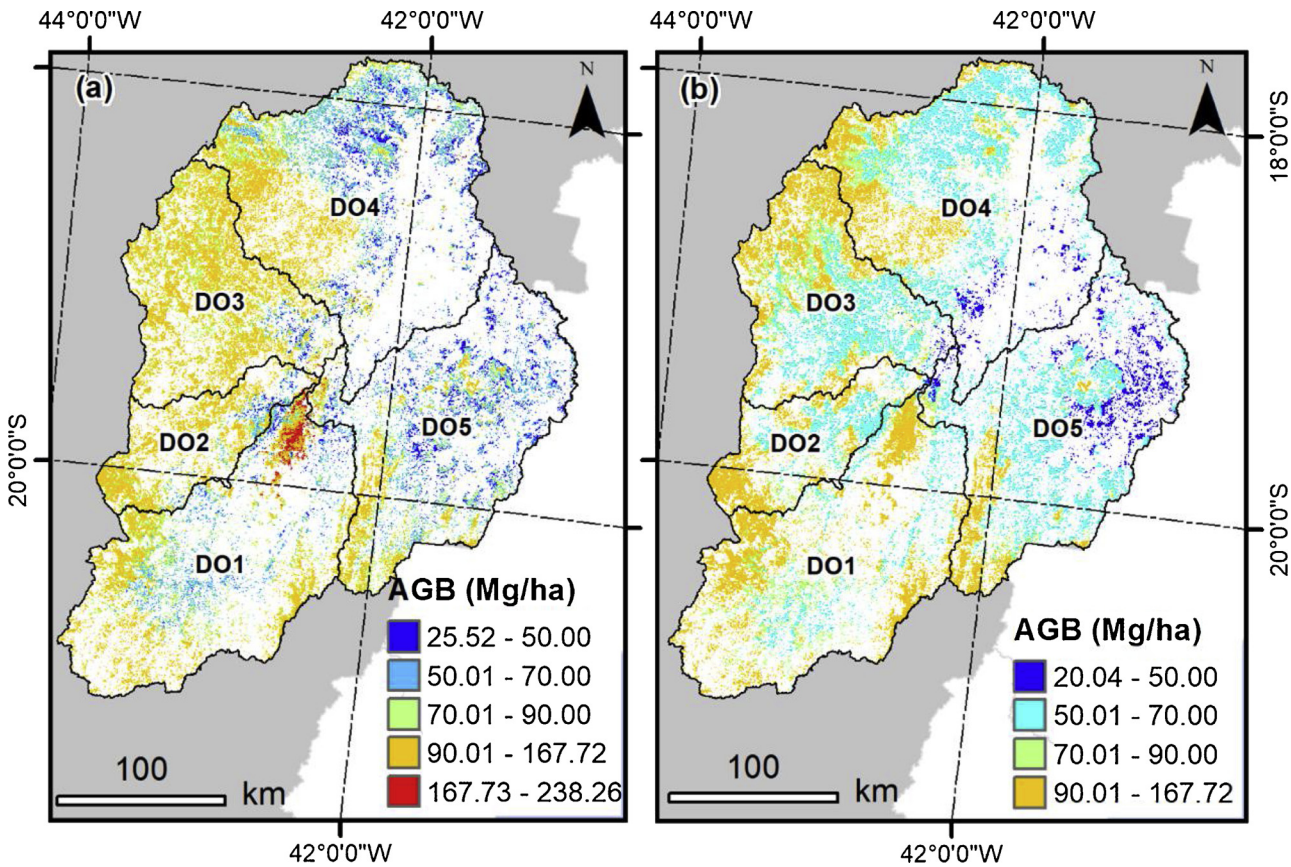


**Fig. 9.** Forest aboveground biomass (AGB) estimation maps modelled using: (a) the object-based approach and (b) the pixel-based approach.

**Table 6**
Mean and total aboveground biomass for each water resources planning and management unit (DO1, DO2, DO3, DO4, and DO5) present in the Rio Doce basin in Minas Gerais state, Brazil.

| PLAN | Object-based | | Pixel-based | |
|---|---|---|---|---|
| | Mean (Mg/ha) | Total (Mg) | Mean (Mg/ha) | Total (Mg) |
| DO1 | 96.61 | 44,434,813 | 98.25 | 45,184,132 |
| DO2 | 100.68 | 21,492,684 | 88.28 | 18,824,760 |
| DO3 | 100.11 | 50,016,351 | 86.28 | 43,125,089 |
| DO4 | 82.04 | 48,416,658 | 75.42 | 44,474,970 |
| DO5 | 73.86 | 31,439,025 | 64.28 | 27,358,704 |
| **Total (Mg)** | **89.45** | **195,799,533** | **81.78** | **178,967,656** |

**Table 7**
Natural breaks interval to separate variables from low to high values.

| Variables[1] | Interval | | **Unit** |
|---|---|---|---|
| | Low - L | High - H | |
| AGB | 25.5 – 87.0 | 87.1 – 238.3 | Mg/ha |
| BIO1 | 13.4 – 21.3 | 21.4 – 25.1 | ºC |
| BIO12 | 1,052.0 – 1354.0 | 1354.1 – 1792.0 | mm |
| DEM | 96.0 – 658.0 | 658.1 – 2078.0 | m |
| MRVBF | 0.0 – 2.2 | 2.3 – 6.9 | |
| VD | 0.0 – 525.0 | 525.1 -1,190.0 | m |
| TCW | −0.36 - -0.11 | −0.10 - 0.07 | |

[1] AGB – above ground biomass; BIO1 - Annual Mean Temperature; BIO12 - Annual Precipitation; DEM – Digital Elevation Model; MRVBF - Multiresolution Index of Valley Bottom Flatness; VD – Valley Depth; TCW – Tasselled Cap Wetness.

class variance and maximizing differences between classes (Pflugmacher et al., 2012).

Environmental data (climate and soil properties) are expected to have an influence on the AGB variability (Ferry et al.,2010; Malhi et al., 2006). As shown in previous studies, even in the case of strong environmental contrasts monitored at fine scale, environmental effects only explain a small fraction of variations in AGB and interact largely with more important structural effects (Guitet et al., 2015). This is mainly explained by the limitations: (1) few sets of plots exist for which high-quality data are available describing climate, soil properties, and plot data, (2) even where these data exist, as is the case of this study (188 field plots), rarely do plot networks cover broad gradients that can disentangle covariation among climate, soil, and forest structure such as stem density, DBH, height and basal area (Baraloto et al., 2011).

David et al. (2017) analysed the climate variables to develop a general rule, applicable to large areas, for classifying tropical forests into classes of aboveground forest carbon stocks. Their results suggested that climate variables helped to discriminate carbon classes and their effect differed between biomes and different sets of variables. The simultaneous use of mean annual temperature and mean annual precipitation enhanced the classification accuracy by up to 9.2% in the Atlantic Forest biome across Minas Gerais state, Brazil, covering a portion of Rio Doce Basin, the same study area assessed here. Forests with the highest stocking levels are negatively correlated with mean annual temperature and positively correlated with mean annual precipitation, meaning that, in warmer and drier regions, the forests have lower carbon stocks.

Our study demonstrates the same behaviour, with an inverse relation between AGB and BIO1 and a positive relation with BIO12. The west region of Rio Doce basin is characterized by high AGB values and low temperatures (HL = 48.70% of the total area). On the contrary, the east region present low AGB values and high temperatures

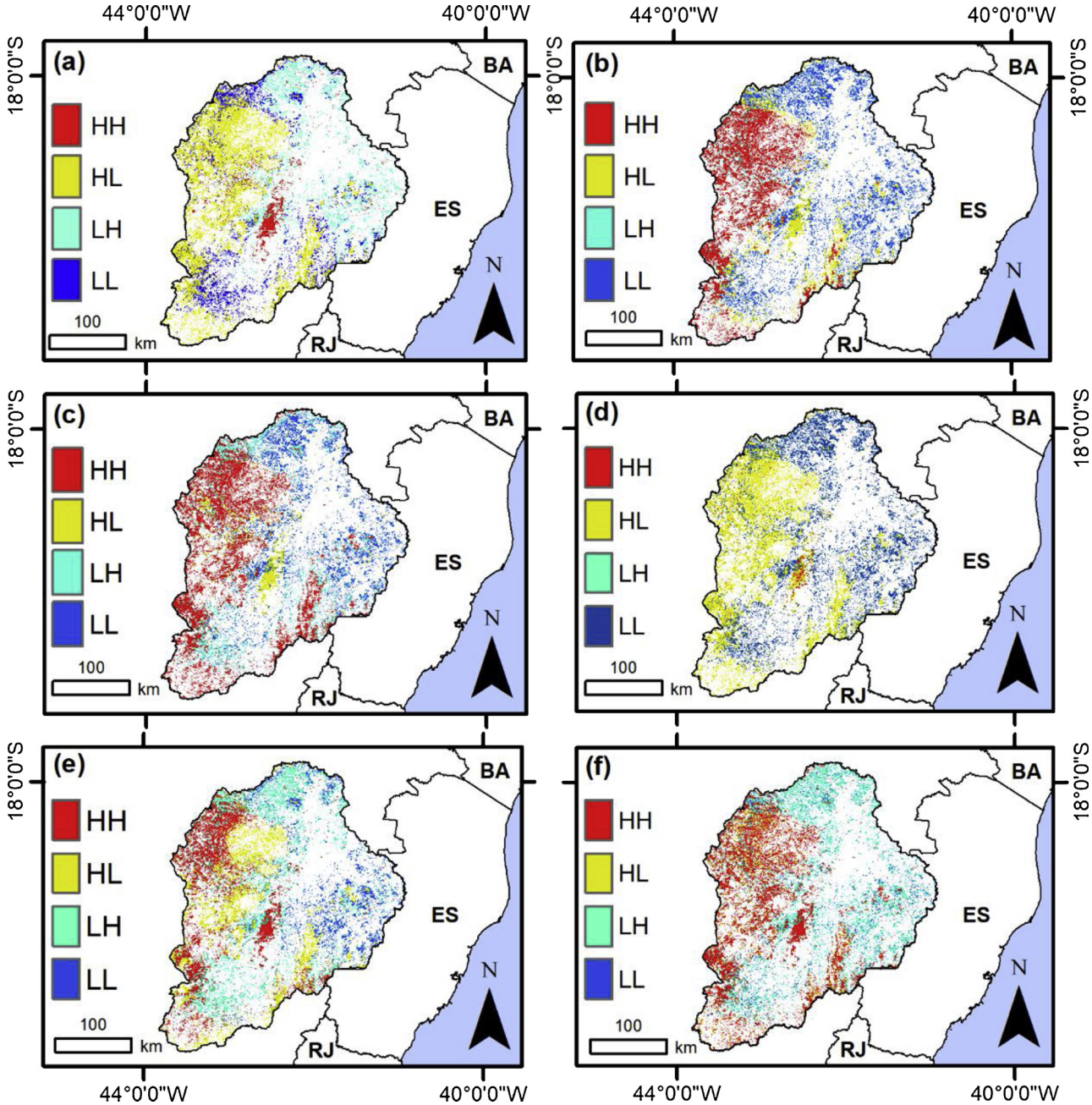


**Fig. 10.** Spatial distribution of combined values of aboveground biomass (AGB) with (a) BIO1 - Annual Mean Temperature, (b) BIO12 - Annual Precipitation, (c) DEM – Digital Elevation Model, (d) MRVBF - Multiresolution Index of Valley Bottom Flatness, (e) VD – Valley Depth, (f) TCW – Tasselled Cap Wetness.

(LH = 25.55% of the total area) (Fig. 10a). About precipitation, a positive relation with AGB was found (Fig. 10b). The west region is characterized by high values (HH = 36.94% of the total area) and the east by low values (LL = 36.94% of the total area). These results corroborate with studies such Li et al. (2015) in central Asian states and, Raich et al. (2006) working in intact tropical evergreen forests globally, in which precipitation behaves inversely in relation to temperature.

Terrain-related data are good predictors in forest ecosystems, which usually relates to other variables that directly influence plant growth, such as soil type (Jafari et al., 2014; Pelegrino et al., 2016), soil water content (Ågren et al., 2014; Silva et al., 2014), soil fertility (Hengl et al., 2015; Luizão et al., 2004), and soil texture (Chagas et al., 2016; Forkuor et al., 2017). AGB in forest ecosystems vary with elevation (Rajput et al., 2015; Shah et al., 2014). In most of these studies, carbon stocks decreased with elevation, while carbon stocks at a given elevation varied strongly between and within continents. This pattern is likely to be driven by differences in climate (Ensslin et al., 2015). Some studies did not find any link between AGB and elevation (Culmsee et al., 2010; Dossa et al., 2013). They found a fairly constant AGB along an altitudinal gradient. Some researchers have suggested that a hump-shaped relationship may be due to drier climates at lower elevations (Dossa et al., 2013). Here, elevation-DEM presented significant positive relation with AGB (Fig. 10c), with HH = 46.62% and LL = 26.98% of the study area, totalling 73.60% of forest remnants that are direct influenced by elevation, in terms of AGB.

Our study also selected MRVBF and VD data as potential predictors of AGB. We found two main patterns (Fig. 10d): (1) high AGB and low MRVBF in the west region (HL = 52.88%), probably due to the high elevation and slope in this region and; (2) a positive correlation between AGB and MRVBF in the east region of Rio Doce basin

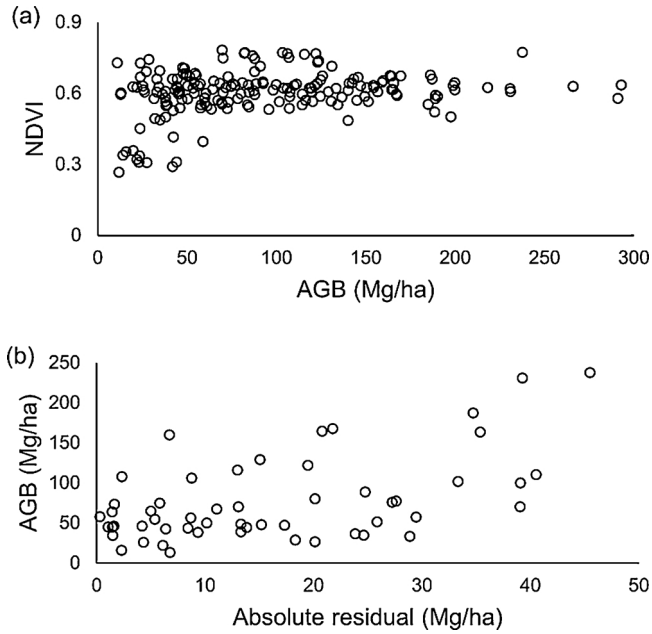


**Fig. 11.** Scatterplots of measured plots in Rio Doce Basin: (a) Normalized Difference Vegetation Index (NDVI) versus aboveground biomass (AGB) in Mg/ha; (b) Measured AGB (Mg/ha) versus absolute residuals (Mg/ha).

(LL = 45.35%), indicating that the closer the place from the valley bottom area, the higher the AGB, probably due to the presence of water in greater amounts in such places.

Regarding the VD feature selected, we found four patterns (Fig. 10e): (1) high AGB and high values of VD variable in the west part of the Rio Doce basin and in a specific region in the central area, totalling 29.09%; (2) high AGB and low VD (HL = 25.45%, mainly in the west region; (3) low values of AGB and high VD (LH = 27.26%) in the east part and; (4) both low values (LL = 18.20%) in the east region of the study area. A negative relation means that closer to drainage network, the lower the AGB. This situation can be explained due to the human-caused disturbances in flat areas near to the water courses, where forest areas were converted to pastures or croplands. On the contrary, in steep slope areas (lower VD), the forest remnants are preserved.

It is important to note that almost all studies that evaluate topographic variables as determinants of forest biomass and forest structure have been carried out in preserved forests (Jaramillo et al., 2003; Laumonier et al., 2010), that is, in forests that are mostly unaffected by human factors. The addition of anthropogenic factors complicates the relationships between environmental variables and biomass and makes prediction of biomass levels over space more challenging (Salinas-Melgoza et al., 2018), but this could be achieved in this work.

Optical remote sensing variables, such as spectral bands, vegetation indices, texture derivatives, Tasselled Cap transformations and principal component analysis can be used as potential predictors for AGB modelling for both object and pixel levels. However, some challenges remain, such as (1) the saturation problem for forest sites with high biomass density and the fact that (2) spectral variables are influenced by external factors such as atmosphere, soil moisture, and vegetation phenology (Lu et al., 2016). The roles of spectral bands in AGB modelling are dependent on the complexity of forest structures (Lu, 2005). Some indices saturate for forest sites with high biomass, which decreased their potential for AGB modelling (Lu et al., 2016). Fig. 11a illustrate the relationship between NDVI and measured AGB from our 188 plots. It is possible to observe that NDVI saturated above 0.6, associated with a biomass of approximately 100 Mg/ha, beyond which further increases in biomass did not result in corresponding increases. The model estimates are more accurate in areas with lower average AGB (Fig. 11b) and lower standard deviation of AGB due to a saturation between 100 and 200 Mg/ha using Landsat data (as shown by Duncanson et al., 2010). Others are sensitive to terrain illumination effects in rugged surfaces or mountains like the study area (Galvão et al., 2016). For example, Matsushita et al. (2007) when studying high-density plantations over mountains in Japan, reported that the EVI is more sensitive to topographic effects than the NDVI.

We found a positive relation between AGB and TCW (Fig. 10f) corroborating with studies such Vicharnakorn et al. (2014) in tropical forest lands of Thailand. The west region in the map is characterized by both high values of AGB and TCW (HH = 41.87%). However, we found areas with low AGB and high TCW values (LH = 32.29%) in the east region of the study area.

## 6. Conclusions

We compared an object-based approach versus pixel-based method for modelling and mapping aboveground biomass (AGB) in a Brazilian Atlantic Forest basin by training Random Forests (RF) machine learning algorithm using environmental, terrain, and remote sensing data. Results suggested that object-based method is promising due to the better performance in AGB prediction.

Object-based modelling is a viable alternative to the traditional pixel-based method to reduce AGB variability in hilly, mountainous and heterogeneous environmental that are influenced by climate and topography conditions, captured by a combination of selected environmental, terrain related and remote sensing variables. The spatial distribution of AGB revealed a decreasing trend from the west towards the east region of the Rio Doce Basin, with a total of 195,799,533 Mg, ranging from 25.52 to 238 Mg/ha, following the precipitation pattern and anthropogenic disturbance effects.

This study provided accurate AGB estimates for the Rio Doce basin, one of the most important watercourses of Brazilian Atlantic Forest. In conclusion, we highlight that OBIA is a preferred solution to map forest AGB over common pixel-based methods, increasing the precision of AGB estimates in a heterogeneous and mountain tropical environment.

## Acknowledgements

The authors would like to thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) and FAPEMIG for financing part of this study.

## References

Addink, E.A., Jong, S.M., Pebesma, E.J., 2007. The importance of scale in object-based mapping of vegetation parameters with hyperspectral imagery. Photogramm. Eng. Remote Sens. 73, 905–912. https://doi.org/10.14358/PERS.73.8.905.

Ågren, A.M., Lidberg, W., Strömgren, M., Ogilvie, J., Arp, P.A., 2014. Evaluating digital terrain indices for soil wetness mapping–a Swedish case study. Hydrol. Earth Syst. Sci. 18, 3623–3634. https://doi.org/10.5194/hess-18-3623-2014.

Andersen, H.E., Strunk, J., Temesgen, H., Atwood, D., Winterberger, K., 2012. Using multilevel remote sensing and ground data to estimate forest biomass resources in remote regions: a case study in the boreal forests of interior Alaska. Can. J. Remote. Sens. 37, 596–611. https://doi.org/10.5589/m12-003.

Asner, G.P., Hughes, R.F., Varga, T.A., Knapp, D.E., Kennedy-bowdoin, T., 2009. Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. Ecosystems 12, 261–278. https://doi.org/10.1007/s10021-008-9221-5.

Baatz, M., Schäpe, A., 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. J. Photogramm. Remote Sens. 58, 12–23.

Baccini, A., Laporte, N., Goetz, S.J., Sun, M., Dong, H., 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. Environ. Res. Lett. 3. https://doi.org/10.1088/1748-9326/3/4/045011.

Baraloto, C., Rabaud, S., Molto, Q., Blanc, L., Fortunel, C., Hérault, B., Dávila, N., Mesones, I., Rios, M., Valderrama, E., Fine, P.V.A., 2011. Disentangling stand and environmental correlates of aboveground biomass in Amazonian forests. Glob. Change Biol. 17, 2677–2688. https://doi.org/10.1111/j.1365-2486.2011.02432.x.

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. Hydrol. Sci. Bull. 24, 43–69. https://doi.org/10.1080/02626667909491834.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS J.

Photogramm. Remote. Sens. 65, 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.
Carvalho, L.M.T., Scolforo, J.R., Oliveira, A.D., Mello, J.M., Oliveira, L.Tde, Acerbi-Junior, F.W., Cavalcanti, H.C., Filho, R.V., 2006. Procedimentos de mapeamento, in: Mapeamento e Inventário Da Flora e Dos Reflorestamentos de Minas Gerais. pp. 37–57.
Chagas, C.S., Carvalho Junior, W., Bhering, S.B., Calderano Filho, B., 2016. Spatial prediction of soil surface texture in a semiarid region using random forest and multiple linear regressions. Catena 139, 232–240. https://doi.org/10.1016/j.catena.2016.01.001.
Chen, G., Hay, G.J., Carvalho, L.M.T., Wulder, M.A., 2012. Object-based change detection. Int. J. Remote Sens. https://doi.org/10.1080/01431161.2011.648285.
Cohen, W.B., Goward, S.N., 2004. Landsat's role in ecological applications of remote sensing. BioScience 54, 535–545. https://doi.org/10.1641/0006-3568(2004)054[0535:LRIEAO]2.0.CO;2.
Collard, F., Kempen, B., Heuvelink, G.B.M., Saby, N.P.A., Richer de Forges, A.C., Lehmann, S., Nehlig, P., Arrouays, D., 2014. Refining a reconnaissance soil map by calibrating regression models with data from the same map (Normandy, France). Geoderma 1, 21–30. https://doi.org/10.1016/j.geodrs.2014.07.001.
Colombo, A., Joly, C., 2010. Brazilian Atlantic Forest lato sensu: the most ancient Brazilian forest, and a biodiversity hotspot, is highly threatened by climate change. Braz. J. Biol. 70, 697–708. https://doi.org/10.1590/S1519-69842010000400002.
Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1.4. Geosci. Model. Dev. 8, 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015.
Crist, E.P., Cicone, R.C., 1984. A phisically-based transformation of thematic mapper data - the tasseled cap. IEEE Trans. Geosci. Remote. Sens. 22, 256–263. https://doi.org/10.1109/TGRS.1984.350619.
Culmsee, H., Leuschner, C., Moser, G., Pitopang, R., 2010. Forest aboveground biomass along an elevational transect in Sulawesi, Indonesia, and the role of Fagaceae in tropical montane rain forests. J. Biogeogr. 37, 960–974. https://doi.org/10.1111/j.1365-2699.2009.02269.x.
Daleles, C., Donato, A., Adriana, L., Vianei, J., Hodnett, M.G., Tomasella, J., Waterloo, M.J., 2008. HAND, a new terrain descriptor using SRTM-DEM: mapping terra- firme rainforest environments in Amazonia. Remote Sens. Environ. 112, 3469–3481. https://doi.org/10.1016/j.rse.2008.03.018.
David, H.C., de Araújo, E.J.G., Morais, V.A., Scolforo, J.R.S., Marques, J.M., Péllico Netto, S., MacFarlane, D.W., 2017. Carbon stock classification for tropical forests in Brazil: understanding the effect of stand and climate variables. For. Ecol. Manage. 404, 241–250. https://doi.org/10.1016/j.foreco.2017.08.044.
De Smith, M.J., Goodchild, M.F., Longley, P.A., 2018. Geospatial Analysis - A Comprehensive Guide to Principles, Techniques and Software Tools, 3rd edition. 602p.
Díaz-Uriarte, R., Alvarez de Andrés, S., 2006. Gene selection and classification of microarray data using random forest. BMC Bioinformatics 7, 1–13. https://doi.org/10.1186/1471-2105-7-3.
Dossa, G.G.O., Paudel, E., Fujinuma, J., Yu, H., Chutipong, W., Zhang, Y., Paz, S., Harrison, R.D., 2013. Factors determining forest diversity and biomass on a Tropical Volcano, Mt. Rinjani, Lombok, Indonesia. PLoS One 8, 1–11. https://doi.org/10.1371/journal.pone.0067720.
Dronova, I., 2015. Object-based image analysis in wetland research: a review. Remote Sens. 7, 6380–6413. https://doi.org/10.3390/rs70506380.
Duncanson, L.I., Niemann, K.O., Wulder, M.A., 2010. Integration of GLAS and Landsat TM data for aboveground biomass estimation. Can. J. Remote Sens. 36, 129–141. https://doi.org/10.5589/m10-037.
Duro, D.C., Franklin, S.E., Dube, M.G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens. Environ. 118, 259–272. https://doi.org/10.1016/j.rse.2011.11.020.
Ensslin, A., Rutten, G., Pommer, U., Zimmermann, R., Hemp, A., Fischer, M., 2015. Effects of elevation and land use on the biomass of trees, shrubs and herbs at Mount Kilimanjaro. Ecosphere 6 (3), 1–15. https://doi.org/10.1890/ES14-00492.1.
Fayad, I., Baghdadi, N., Guitet, S., Bailly, J.S., Hérault, B., Gond, V., El Hajj, M., Tong Minh, D.H., 2016. Aboveground biomass mapping in French Guiana by combining remote sensing, forest inventories and environmental data. Int. J. Appl. Earth Observ. Geoinf. 52, 502–514. https://doi.org/10.1016/j.jag.2016.07.015.
Feng, Y., Wu, J., Zhang, J., Zhang, X., Song, C., 2011. Identifying the relative contributions of climate and grazing to both direction and magnitude of Alpine grassland productivity dynamics from 1993 to 2011 on the Northern Tibetan plateau. Remote Sens. 9, 1–15. https://doi.org/10.3390/rs9020136.
Ferry, B., Bontemps, Jean-Daniel, François, M., Blanc, L., Freycon, V., 2010. Higher treefall rates on slopes and waterlogged soils result in lower stand biomass and productivity in a tropical rain forest. J. Ecol. 98, 106–116. https://doi.org/10.1111/j.1365-2745.2009.01604.x.
Forkuor, G., Hounkpatin, O.K.L., Welp, G., Thiel, M., 2017. High resolution mapping of soil properties using remote wensing variables in south-western Burkina Faso: A comparison of machine learning and multiple linear regression models. PLoS One 12, e0170478. https://doi.org/10.1371/journal.pone.0170478.
Galidaki, G., Zianis, D., Gitas, I., Radoglou, K., Karathanassi, V., Tsakiri–Strati, M., Woodhouse, I., Mallinis, G., 2017. Vegetation biomass estimation with remote sensing: focus on forest and other wooded land over the Mediterranean ecosystem. Int. J. Remote Sens. 38, 1940–1966. https://doi.org/10.1080/01431161.2016.1266113.
Gallant, J.C., Dowling, T.I., 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. Water Resour. Res. 39, 1–14. https://doi.org/10.1029/2002WR001426.
Galvão, L.S., Breunig, F.M., Teles, T.S., Gaida, W., Balbinot, R., 2016. Investigation of terrain illumination effects on vegetation indices and VI-derived phenological metrics in subtropical deciduous forests. GIsci. Remote Sens. 53, 360–381. https://doi.org/10.1080/15481603.2015.1134140.
Gardner, T.A., Barlow, J., Araujo, I.S., Ávila-Pires, T.C., Bonaldo, A.B., Costa, J.E., Esposito, M.C., Ferreira, L.V., Hawes, J., Hernandez, M.I.M., Hoogmoed, M.S., Leite, R.N., Lo-Man-Hung, N.F., Malcolm, J.R., Martins, M.B., Mestre, L.A.M., Miranda-Santos, R., Overal, W.L., Parry, L., Peters, S.L., Ribeiro, M.A., Da Silva, M.N.F., Silva Motta, C., Peres, C.A., 2008. The cost-effectiveness of biodiversity surveys in tropical forests. Ecol. Lett. 11, 139–150. https://doi.org/10.1111/j.1461-0248.2007.01133.x.
Gómez, C., White, J.C., Wulder, M.A., Alejandro, P., 2014. Historical forest biomass dynamics modelled with Landsat spectral trajectories. Isprs J. Photogramm. Remote. Sens. 93, 14–28. https://doi.org/10.1016/j.isprsjprs.2014.03.008.
Groves, C.R., Jensen, D.B., Valutis, L.L., Redford, K.H., Shaffer, M.L., Scott, J.M., Baumgartner, J.V., Higgins, J.V., Beck, M.W., Anderson, M.G., 2002. Planning for biodiversity conservation: putting conservation science into practice. BioScience 52, 499. https://doi.org/10.1641/0006-3568(2002)052.
Guevara, Y.Z.C., de Souza, J.J.L.L., Veloso, G.V., Veloso, R.W., Rocha, P.A., Abrahão, W.A.P., Fernandes Filho, E.I., 2018. Reference values of soil quality for the Rio Doce Basin. Revista Brasileira de Ciência do Solo 42, 1–16. https://doi.org/10.1590/18069657rbcs20170231.
Guitet, S., Hérault, B., Molto, Q., Brunaux, O., Couteron, P., 2015. Spatial structure of above-ground biomass limits accuracy of carbon mapping in rainforest but large scale forest inventories can help to overcome. PLoS One 10, 1–22. https://doi.org/10.1371/journal.pone.0138456.
Hengl, T., Heuvelink, G.B.M., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Shepherd, K.D., Sila, A., MacMillan, R.A., Mendes de Jesus, J., Tamene, L., Tondoh, J.E., 2015. Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. PLoS One 10, e0125814. https://doi.org/10.1371/journal.pone.0125814.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978. https://doi.org/10.1002/joc.1276.
Houghton, R.A., 2005. Aboveground forest biomass and the global carbon balance. Glob. Change Biol. 11, 945–958. https://doi.org/10.1111/j.1365-2486.2005.00955.x.
Houghton, R.A., Hall, F.G., Scott, J., 2009. Importance of biomass in the global carbon cycle. J. Geophys. Res. 114, 1–13. https://doi.org/10.1029/2009JG000935.
Hu, T., Su, Y., Xue, B., Liu, J., Zhao, X., Fang, J., Guo, Q., 2016. Mapping global forest aboveground biomass with spaceborne LiDAR, optical imagery, and forest inventory data. Remote Sens. 8. https://doi.org/10.3390/rs8070565.
Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.
Huete, A., Didan, K., Miura, H., Rodriguez, E.P., Gao, X., Ferreira, L.F., 2002. Overview of the radiometric and biopyhsical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213.
Ismail, R., Mutanga, O., 2010. A comparison of regression tree ensembles: predicting Sirex noctilio induced water stress in Pinus patula forests of KwaZulu-Natal, South Africa. Int. J. Appl. Earth Obs. Geoinf. 12, 45–51. https://doi.org/10.1016/j.jag.2009.09.004.
Jafari, A., Khademi, H., Finke, P.A., Van de Wauw, J., Ayoubi, S., 2014. Spatial prediction of soil great groups by boosted regression trees using a limited point dataset in an arid region, southeastern Iran. Geoderma 232–234, 148–163. https://doi.org/10.1016/j.geoderma.2014.04.029.
Jaramillo, J., Kauffman, J.B., Renterı, L., Cummings, D.L., Ellingson, L.J., 2003. Biomass, carbon, and nitrogen pools in mexican tropical dry forest landscapes. Ecosystems 6, 609–629. https://doi.org/10.1007/s10021-002-0195-4.
Jarvis, A., Rubiano, J., Nelson, A., Farrow, A., Mulligan, M., 2004. Practical Use of SRTM Data in the Tropics – Comparisons With Digital Elevation Models Generated From Cartographic Data. Working Document No. 198. 32p. CIAT, Cali, Columbia.
Kajisa, T., Murakami, T., Mizoue, N., Top, N., Yoshida, S., 2009. Object-based forest biomass estimation using Landsat ETM+ in Kampong Thom Province, Cambodia. J. For. Res. 14, 203–211. https://doi.org/10.1007/s10310-009-0125-9.
Kopecký, M., Čížková, Š., 2010. Using topographic wetness index in vegetation ecology: does the algorithm matter? Appl. Veg. Sci. 13, 450–459. https://doi.org/10.1111/j.1654-109X.2010.01083.x.
Kumar, L., Mutanga, O., 2017. Remote sensing of above-ground biomass. Remote Sens. 9, 1–8. https://doi.org/10.3390/rs9090935.
Laumonier, Y., Edin, A., Kanninen, M., Munandar, A.W., 2010. Landscape-scale variation in the structure and biomass of the hill dipterocarp forest of Sumatra: implications for carbon stock assessments. For. Ecol. Manage. 259, 505–513. https://doi.org/10.1016/j.foreco.2009.11.007.
Li, C., Zhang, C., Luo, G., Chen, X., Maisupova, B., Madaminov, A.A., Han, Q., Djenbaev, B.M., 2015. Carbon stock and its responses to climate change in Central Asia. Glob. Change Biol. 21, 1951–1967. https://doi.org/10.1111/gcb.12846.
Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2, 18–22. https://doi.org/10.1023/A:1010933404324.
Liu, F., Geng, X., Zhu, A.X., Fraser, W., Waddell, A., 2012. Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS. Geoderma 171–172, 44–52. https://doi.org/10.1016/j.geoderma.2011.05.007.
Lu, D., 2005. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. Int. J. Remote Sens. 26, 2509–2525. https://doi.org/10.1080/01431160500142145.
Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., Moran, E., 2016. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. Int. J. Digit.

Earth 9, 63–105. https://doi.org/10.1080/17538947.2014.990526.
Luizão, R.C.C., Luizão, F.J., Paiva, R.Q., Monteiro, T.F., Souza, L.S., Kruijt, B., 2004. Variation of carbon and nitrogen cycling processes along a topographic gradient in a central Amazonian forest. Glob. Change Biol. 10, 592–600. https://doi.org/10.1111/j.1529-8817.2003.00757.x.
Malhi, Y., Wood, D., Baker, T.R., Wright, J., Phillips, O.L., Cochrane, T., Meir, P., Chave, J., Almeida, S., Arroyo, L., Higuchi, Ni., Killeen, T.J., Laurance, S.G., Laurance, W.F., Lewis, S.L., Monteagudo, A., Neill, D.A., Vargas, P.N., Pitman, N.C.A., Quesada, C.A., Salomão, R., Silva, J.N.M., Lezama, A.T., Terborgh, J., Martínez, R.V., Vinceti, B., 2006. The regional variation of aboveground live biomass in old-growth Amazonian forests. Glob. Change Biol. 12, 1107–1138. https://doi.org/10.1111/j.1365-2486.2006.01120.x.
Malone, B.P., Styc, Q., Minasny, B., McBratney, A.B., 2017. Digital soil mapping of soil carbon at the farm scale: a spatial downscaling approach in consideration of measured and uncertain data. Geoderma 290, 91–99. https://doi.org/10.1016/j.geoderma.2016.12.008.
Mareya, H.T., Tagwireyi, P., Ndaimani, H., Gara, T.W., Gwenzi, D., 2018. Estimating tree crown area and aboveground biomass in Miombo Woodlands from high-resolution RGB-Only imagery. Ieee J. Sel. Top. Appl. Earth Observ. Remote Sens. 11, 868–875. https://doi.org/10.1109/JSTARS.2018.2799386.
Matasci, G., Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Zald, H.S.J., 2018. Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots. Remote Sens. Environ. 209, 90–106. https://doi.org/10.1016/j.rse.2017.12.020.
Matsushita, B., Yang, W., Chen, J., Onda, Y., Qiu, G., 2007. Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to topographic effects: a case study in high-density Cypress forest. Sensors 7, 2636–2651. https://doi.org/10.3390/s7112636.
Miller, J.D., Thode, A.E., 2007. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta normalized Burn Ratio (dNBR). Remote Sens. Environ. 109, 66–80. https://doi.org/10.1016/j.rse.2006.12.006.
Lu, D., Mausel, P., Brondızio, E., Moran, E., 2004. Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. For. Ecol. Manage. 198, 149–167. https://doi.org/10.1016/j.foreco.2004.03.048.
Nikolakopoulos, K.G., Kamaratakis, E.K., Chrysoulakis, N., 2006. SRTM vs ASTER elevation products. Comparison for two regions in Crete, Greece. Int. J. Remote Sens. 27, 4819–4838. https://doi.org/10.1080/01431160600835853.
Pelegrino, M.H.P., Silva, S.H.G., Menezes, M.D., Silva, E., Owens, P.R., Curi, N., 2016. Mapping soils in two watersheds using legacy data and extrapolation for similar surrounding areas. Ciência e Agrotecnologia 40, 534–546. https://doi.org/10.1590/1413-70542016405011416.
Pflugmacher, D., Cohen, W.B., Kennedy, R.E., 2012. Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. Remote Sens. Environ. 122, 146–165. https://doi.org/10.1016/j.rse.2011.09.025.
Powell, S.L., Cohen, W.B., Healey, S.P., Kennedy, R.E., Moisen, G.G., Pierce, K.B., Ohmann, J.L., 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: a comparison of empirical modeling approaches. Remote Sens. Environ. 114, 1053–1068. https://doi.org/10.1016/j.rse.2009.12.018.
R Core Team, 2014. A Language and Environment for Statistical Computing.
Raich, J.W., Russel, A.E., Kitayama, K., Parton, W.J., Vitousek, P.M., 2006. Temperature influences carbon accumulation in moist tropical forests. Concepts Synth. 87, 76–87. https://doi.org/10.1890/05-0023.
Rajput, B.S., Bhardwaj, D.R., Pala, N.A., 2015. Carbon dioxide mitigation potential and carbon density of different land use systems along an altitudinal gradient in northwestern Himalayas. Agrofor. Syst. 89, 525–536. https://doi.org/10.1007/s10457-015-9788-8.
Ribeiro, M.C., Metzger, J.P., Martensen, A.C., Ponzoni, F.J., Hirota, M.M., 2009. The Brazilian Atlantic Forest: how much is left, and how is the remaining forest distributed? Implications for conservation. Biol. Conserv. 142, 1141–1153. https://doi.org/10.1016/j.biocon.2009.02.021.
Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Remote Sens. Environ. 55, 95–107. https://doi.org/10.1016/0034-4257(95)00186-7.
Salinas-Melgoza, M.A., Skutsch, M., Lovett, J.C., 2018. Predicting aboveground forest biomass with topographic variables in human-impacted tropical dry forest landscapes. Ecosphere 9, 1–20. https://doi.org/10.1002/ecs2.2063.
Samuel-Rosa, A., Heuvelink, G.B.M., Vasques, G.M., Anjos, L.H.C., 2015. Do more detailed environmental covariates deliver more accurate soil maps? Geoderma 243–244, 214–227. https://doi.org/10.1016/j.geoderma.2014.12.017.
Scolforo, H.F., Scolforo, J.R.S., Mello, C.R., Mello, J.M., Ferraz Filho, A.C., 2015. Spatial distribution of aboveground carbon stock of the arboreal vegetation in Brazilian Biomes of Savanna, Atlantic Forest and Semi-arid woodland. PLoS One 10, 1–20. https://doi.org/10.1371/journal.pone.0128781.
Scolforo, H.F., Scolforo, J.R.S., de Mello, J.M., de Mello, C.R., Morais, V.A., 2016. Spatial interpolators for improving the mapping of carbon stock of the arboreal vegetation in Brazilian biomes of Atlantic forest and Savanna. For. Ecol. Manage. 376, 24–35. https://doi.org/10.1016/j.foreco.2016.05.047.
Shah, S., Sharma, D.P., Pala, N.A., Tripathi, P., Kumar, M., 2014. Temporal variations in carbon stock of Pinus roxburghii Sargent forests of Himachal Pradesh, India. J. Mount. Sci. 11, 959–966. https://doi.org/10.1007/s11629-013-2725-2.
Silva, J.M.C., Tabarelli, M., 2000. Tree species improverishment and the future flora of the Atlantic Forest of northeast Brazil. Nature 404, 72–74.
Silva, B.M., Silva, S.H.G., Oliveira, G.C., Peters, P.H.C.R., Santos, W.J.R., Curi, N., 2014. Soil moisture assessed by digital mapping techniques and its field validation. Ciência e Agrotecnologia 38, 140–148. https://doi.org/10.1590/S1413-70542014000200005.
Silveira, E.M.O., Menezes, M.D., Acerbi Júnior, F.W., Terra, M.C.N.S., Mello, J.M., 2017. Assessment of geostatistical features for object-based image classification of contrasted landscape vegetation cover. J. Appl. Remote Sens. 11, 036004. https://doi.org/10.1117/1.JRS.11.036004.
Silveira, E.M.O., Bueno, I.T., Acerbi Junior, F.W., Mello, J.M., Scolforo, J.R.S., Wulder, M.A., 2018a. Using spatial features to reduce the impact of seasonality for detecting tropical forest changes from landsat time series. Remote Sens. 10. https://doi.org/10.3390/rs10060808.
Silveira, E.M.O., Mello, J.M., Acerbi Júnior, F.W., Carvalho, L.M.T., 2018b. Object-based land-cover change detection applied to Brazilian seasonal savannahs using geostatistical features. Int. J. Remote Sens. 39, 2597–2619. https://doi.org/10.1080/01431161.2018.1430397.
Silveira, E.M.O., Espírito-santo, F.D.B., Acerbi-júnior, F.W., Galvão, L.S., Withey, K.D., Blackburn, G.A., Mello, J.M., Shimabukuro, Y.E., Domingues, T., Scolforo, J.R.S., 2018c. Reducing the effects of vegetation phenology on change detection in tropical seasonal biomes. GIsci. Remote Sens. 00, 1–19. https://doi.org/10.1080/15481603.2018.1550245.
Su, Y., Guo, Q., Xue, B., Hu, T., Alvarez, O., Tao, S., Fang, J., 2016. Spatial distribution of forest aboveground biomass in China: estimation through combination of spaceborne lidar, optical imagery, and forest inventory data. Remote Sens. Environ. 173, 187–199. https://doi.org/10.1016/j.rse.2015.12.002.
Timothy, D., Onisimo, M., Cletah, S., Adelabu, S., Tsitsi, B., 2016. Remote sensing of aboveground forest biomass: a review. Trop. Ecol. 57, 125–132.
Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150. https://doi.org/10.1016/0034-4257(79)90013-0.
Vicharnakorn, P., Shrestha, R.P., Nagai, M., Salam, A.P., Kiratiprayoon, S., 2014. Carbon stock assessment using remote sensing and forest inventory data in Savannakhet, Lao PDR. Remote Sens. 6, 5452–5479. https://doi.org/10.3390/rs6065452.
Vieilledent, G., Gardi, O., Grinand, C., Burren, C., Andriamanjato, M., Camara, C., Gardner, C.J., Glass, L., Rasolohery, A., Rakoto Ratsimba, H., Gond, V., Rakotoarijaona, J.R., 2016. Bioclimatic envelope models predict a decrease in tropical forest carbon stocks with climate change in Madagascar. J. Ecol. 104, 703–715. https://doi.org/10.1111/1365-2745.12548.
Wilson, E.H., Sader, S.A., 2002. Detection of forest harvest type using multiple dates of Landsat TM imagery. Remote Sens. Environ. 80, 385–396. https://doi.org/10.1016/S0034-4257(01)00318-2.
Wu, X., Peng, J., Shan, J., Cui, W., 2015. Evaluation of semivariogram features for object-based image classification. Geo-spatial Inf. Sci. 18, 159–170. https://doi.org/10.1080/10095020.2015.1116206.
Wulder, M.A., Seemann, D., 2003. Forest inventory height update through the integration of lidar data with segmented Landsat imagery. Can. J. Remote. Sens. 29 (5), 536–543.
Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: how free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ. 122, 2–10. https://doi.org/10.1016/j.rse.2012.01.010.
Xu, Y., Franklin, S.B., Wang, Q., Shi, Z., Luo, Y., Lu, Z., Zhang, J., Qiao, X., Jiang, M., 2015. Topographic and biotic factors determine forest biomass spatial distribution in a subtropical mountain moist forest. For. Ecol. Manage. 357, 95–103. https://doi.org/10.1016/j.foreco.2015.08.010.
Young, N.E., Anderson, R.S., Chignell, S.M., Vorster, A.G., Lawrence, R., Evangelista, P.H., 2017. A survival guide to Landsat preprocessing. Ecology 98, 920–932. https://doi.org/10.1002/ecy.1730.
Zhang, C., Denka, S., Cooper, H., Mishra, D.R., 2018. Quantification of sawgrass marsh aboveground biomass in the coastal Everglades using object-based ensemble analysis and Landsat data. Remote Sens. Environ. 204, 366–379. https://doi.org/10.1016/j.rse.2017.10.018.
Zhu, X., Liu, D., 2015. Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. ISPRS J. Photogramm. Remote. Sens. 102, 222–231. https://doi.org/10.1016/j.isprsjprs.2014.08.014.