Forest Ecology and Management 499 (2021) 119615

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Mapping the stock and spatial distribution of aboveground woody biomass in the native vegetation of the Brazilian Cerrado biome

Barbara Zimbres [a,*], Pedro Rodríguez-Veiga [b,c], Julia Z. Shimbo [a], Polyanna da Conceição Bispo [b,d], Heiko Balzter [b,c], Mercedes Bustamante [e], Iris Roitman [e], Ricardo Haidar [f], Sabrina Miranda [g], Letícia Gomes [e], Fabrício Alvim Carvalho [h], Eddie Lenza [i], Leonardo Maracahipes-Santos [a], Ana Clara Abadia [i], Jamir Afonso do Prado Júnior [j], Evandro Luiz Mendonça Machado [k], Anne Priscila Dias Gonzaga [l], Marcela de Castro Nunes Santos Terra [m], José Marcio de Mello [m], José Roberto Soares Scolforo [m], José Roberto Rodrigues Pinto [n], Ane Alencar [a]

[a] Amazon Environmental Research Institute (IPAM), Brasília-DF 71503-505, Brazil
[b] NERC National Centre for Earth Observation (NCEO), Leicester LE1 7RH, UK
[c] Centre for Landscape and Climate Research, School of Geography, Geology and the Environment, University of Leicester, Leicester LE1 7RH, UK
[d] Department of Geography, School of Environment, Education and Development, University of Manchester, Oxford Road, Manchester M13 9PL, UK
[e] Department of Ecology, University of Brasília (UnB) and Brazilian Research Network on Global Climate Change—Rede Clima, Brasília-DF 70910-900, Brazil
[f] Federal University of Tocantins (UFT), Palmas-TO, 77001-090, Brazil
[g] Goiás State University (UEG), Palmeiras de Goiás-GO 76190-000, Brazil
[h] Department of Botany, Federal University of Juiz de Fora (UFJF), Juiz de Fora-MG 36037-000, Brazil
[i] Mato Grosso State University (UNEMAT), Nova Xavantina-MT 78690-000, Brazil
[j] Federal University of Uberlândia (UFU), Uberlândia-MG 38402-020, Brazil
[k] Department of Forest Engineering, Federal University of Vales do Jequitinhonha e Mucuri (UFVJM), Diamantina-MG 39100-000, Brazil
[l] Department of Geography, Federal University of Vales do Jequitinhonha e Mucuri (UFVJM), Diamantina-MG 39100-000, Brazil
[m] Department of Forest Sciences, Federal University of Lavras (UFLA), Lavras-MG 37200-000, Brazil
[n] Department of Forest Engineering, University of Brasília (UnB), Brasília-DF 70910-900, Brazil

ARTICLE INFO

Keywords:
ALOS-2 PALSAR-2
Savanna
Landsat
LiDAR
Machine learning
Synthetic Aperture Radar (SAR)

ABSTRACT

The Brazilian Cerrado biome consists of a highly heterogeneous tropical savanna, and is one of the world's biodiversity hotspots. High rates of deforestation, however, place it as the second-largest source of carbon emissions in Brazil. Due to its heterogeneity, biomass and carbon stocks in the Cerrado vegetation are highly variable, and mapping and monitoring these stocks are not a trivial effort. To address this challenge, we built an aboveground woody biomass (AGWB) model for the Cerrado biome using 30-m resolution optical satellite imagery (Landsat-5 and Landsat-8), 25-m resolution SAR imagery (ALOS and ALOS-2), and a set of plot-based and LiDAR-derived AGWB estimates (n = 1858) from a wide network of researchers in Brazil. We implemented both a Classification and Regression Tree (CART) and a Random Forest (RF) algorithm to model AGWB over the native vegetation in the year 2019 (as classified by MapBiomas) in the Cerrado. The RF algorithms resulted in a slightly better result (R^2 = 53%; rel. RMSE = 57%) than the CART model (R^2 = 45%; rel. RMSE = 63%), but our map shows an underestimation of very high AGWB (negative bias over 200 t ha^{-1}) and a slight overestimation of low AGWB (positive bias), especially in the RF model (bias of 1.19 t ha^{-1} against 0.86 t ha^{-1} for the CART model). We believe we have contributed to knowledge on the woody biomass stocks in the biome, especially in the predominant savanna woodlands, which is where the highest current rates of conversion take place in the Cerrado.

* Corresponding author at: Amazon Environmental Research Institute (IPAM), SCN 211, bloco B, sala 201, 70836-520 Brasilia, Brazil.
E-mail address: barbara.zimbres@ipam.org.br (B. Zimbres).

https://doi.org/10.1016/j.foreco.2021.119615
Received 19 March 2021; Received in revised form 26 July 2021; Accepted 7 August 2021
Available online 20 August 2021
0378-1127/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Savannas are dynamic ecosystems, marked by the co-existence of trees and a continuous grassy-herbaceous layer, where growth patterns are determined by wet and dry seasons (Bourlière & Hadley, 1983). In general, savannas grow in tropical regions 8° to 20° from the Equator. Temperatures are warm to hot in all seasons, but significant rainfall occurs for only a few months each year - from about October to March in the Southern Hemisphere. Considering the length of the dry season, savannas can be divided into dry and wet savannas. In wet seasonal savannas, such as the Brazilian Cerrado, the dry season typically lasts 3–5 months, with approximately 80% of the annual rainfall falling during the wet season (Castro et al., 1994). Soil factors are also important determinants of Cerrado physiognomies. The landscape units are intensively weathered and well-drained terrains resulting in soils that are dominantly acidic Oxisols. In the Cerrado, soil, water availability, weather, and local and regional topographic variations create a significant variety of inner ecosystems and subtypes of savanna, such as grasslands, shrublands, woodlands, and dry forests (Ribeiro & Walter, 2008).

The Brazilian Cerrado is the second largest biome in South America, comprising approximately 2 million Km^2 (or 23% of Brazil's territory). This highly heterogeneous tropical savanna is one of the world's biodiversity hotspots (Myers et al., 2000). The Cerrado covers a wide latitudinal and environmental gradient and is composed of a mosaic of plant formations from open grasslands (7%) to dense forests (32%), while most of its area (61%) is typically woodland savanna (Ribeiro & Walter, 2008; Sano et al., 2010). Due to its high structural heterogeneity, biomass and carbon stocks in the Cerrado are highly variable, so mapping and monitoring these stocks is not a trivial effort. Moreover, most biomass mapping initiatives in Brazil or in the tropics are calibrated with the focus on tropical rainforests (de Almeida et al., 2019; Avitabile et al., 2016; Baccini et al., 2012; Saatchi et al., 2011; dos Santos et al., 2019), with the few exceptions of more localized efforts (e.g. Scolforo et al., 2016; Silveira et al., 2019, in the state of Minas Gerais). To date, there is no wide-ranging mapping effort covering the entire biome using a remote sensing approach well-calibrated with a range of local field inventory data along the Cerrado.

The Cerrado has one of the lowest proportions of protected carbon stocks in Brazil since only around 6.5% of its territory are legally protected areas (Françoso et al., 2015). This biome currently faces high rates of land conversion, and the main driver of this process is the expanding agricultural and cattle ranching frontier (Alencar et al., 2020; Spera et al., 2016; Strassburg et al., 2017). Between 1985 and 2019, agricultural areas increased 225%, while pastures increased 13% across the biome, leaving only 53% of its original vegetation (MapBiomas Collection 5.0). Due to the high rates of vegetation loss, the conversion of the Cerrado represents the second greatest source of carbon emission in Brazil, after the Amazon deforestation (SEEG Brasil, 2020). At the same time, according to the Brazilian law that regulates the native vegetation conversion in private lands, named Forest Code, landowners must conserve the native vegetation in 20–35% of their property in the Cerrado, whereas those in the Amazon biome should conserve 80% (Soares-Filho et al., 2014). This represents 326 thousand Km^2 of legal native vegetation clearing in the Cerrado, releasing 3.2 $GtCO_2$ emissions into the atmosphere (Russo et al., 2018). In addition, there are 25.6 thousand Km^2 of undesignated public areas in the Cerrado, with no clear land tenure and being easily targeted by irregular deforestation and land grabbers. Also, most of deforestation (98%) in the Cerrado has evidence of irregularity or illegality (Azevedo et al., 2021). There is great urgency, therefore, to develop a fast and reliable method to monitor large-scale aboveground biomass while minimizing uncertainties and improving resolution.

Aboveground woody biomass (AGWB) is traditionally estimated by direct plot-based forest inventories. These methods depend on field-measured tree attributes, such as diameter at breast height, tree height, and other variables such as species and/or wood density, which are then used to estimate biomass by means of allometric equations (e.g. Alvarez et al., 2012; Baskerville, 1972; Chave et al., 2005; Chave et al., 2014; Feldpausch et al., 2011; Oliveira Filho & Scolforo, 2008; Rezende et al., 2006; Roitman et al., 2018). However, the need to quantify and monitor AGWB over large extents and at short time scales cannot be met with these field methods. The spatial and temporal constraints of these approaches, together with the costs of large-extent and repetitive field campaigns, hinder the capturing of stock change over short periods of time, making these methods less appropriate for detecting disturbance and recovery, which is crucial information for reliable carbon accounting (Houghton, 2005). Nevertheless, field measurements are critical to calibrate and validate remote sensing estimates of biomass and vegetation structure.

The estimation of vegetation structure and biomass stocks over large extents depends on the calibration of models relating plot-based inventory data and remotely-derived products. Information on field-based reference aboveground biomass is, in this way, related to the spectral signatures at the pixel scale, and a predictive model aimed at extrapolating this relationship is built (Avitabile et al., 2016; Baccini et al., 2012). Openly available optical satellite data, such as from the Landsat series or from Sentinel-2, despite facing issues with cloud cover, provide multispectral data of the entire Earth's surface at high repeat-pass rates, thus providing the spatial and temporal scales necessary for monitoring vegetation structure at moderate resolutions (10–30 m) (e.g. Avitabile et al., 2012; Baccini et al., 2008; Chen et al., 2018; Houghton et al., 2011; Leitão et al., 2018; Schwieder et al., 2016). A large body of literature has been focused on deriving and testing indices from satellite imagery that reflect differences in the vegetation structure and aboveground woody biomass at the pixel-scale (Powell et al., 2010; Schultz et al., 2016). Active sensors, such as Synthetic Aperture RADAR (SAR), are a different and complementary set of products that capture information on the structure of vegetation stands (Bispo et al., 2014; Bispo et al., 2019; Joshi et al., 2017; Le Toan et al., 1992; Ouchi, 2013). Therefore, the integration of a multi-sensor approach using multispectral optical data and SAR data with longer wavelengths (L or P bands) is still the best option to continuously monitor vegetation structure (Omar et al., 2017; Chang and Shoshany, 2016; Sinha et al., 2016).

In this study, we compiled field-based vegetation data from a network of organizations across the Cerrado biome, and investigated the relationships between plot-based field measurements and satellite imagery to train an aboveground woody biomass (AGWB) model for the biome. We explored the potential of a set of remote-sensed (optical and radar) indices to predict AGWB in the Brazilian Cerrado, and thus generate a benchmark map of AGWB from field plot data for the three predominant vegetation types in the biome for the year 2019 (savannas, forests, and grasslands), while also providing a spatially-explicit account of model uncertainty.

2. Materials and methods

2.1. Study area

The Cerrado biome is located in central Brazil. This biome presents wide latitudinal variation (22.4°), and elevation ranges from sea level to 1800 m altitude. The climate is tropical, classified as Köppen's Aw class, due to its strong seasonality: a dry winter from May to September and a rainy summer from October to April (Alvares et al., 2013). The annual precipitation varies between 600 mm and 2000 mm, being drier in the Northeast bordering the semi-arid Caatinga and wetter in the West bordering the Amazon (Assad, 1994). The average annual temperature is approximately 22–23 °C (Coutinho, 2002). Although the absolute maximum temperature does not vary much over the year but can exceed 40 °C, the absolute minimum temperature varies widely, occasionally dropping below freezing in the winter months (June and July) in some

southern parts of the biome (Coutinho, 2002).

The Cerrado vegetation can be classified into three predominant formations, from grasslands to savanna woodlands and forests (Ribeiro & Walter, 2008). The degree of woodiness is the basic criterion to distinguish them. According to Ribeiro and Walter (2008), grasslands are characterized by the dominance of a herbaceous-shrub stratum, with sparsely distributed trees; the savanna woodlands have variable tree/shrub/grass strata, with canopy cover ranging from 50% to 70%; and forest formations are denser, with relatively larger and taller trees, absence of a grass stratum, with canopy cover ranging from 50% to 95%. Total plant biomass varies according to the type of formation, with an average of 24.75 tC ha^{-1} in grasslands, 39.9 tC ha^{-1} in savannas, and 80.6 tC ha^{-1} in forests (MCTIC, 2020).

2.2. Response variables and predictors

2.2.1. Calibration data

We obtained data from field-based inventories of the woody vegetation from a network of collaborators, initially based on the compilation by Roitman et al. (2018), and from the literature (Ottmar et al., 2001 for grasslands). These field plots are widely distributed in the biome, despite presenting clear gaps in the northern and southern regions (Fig. 1). Some plots were sampled outside the official limits of the Cerrado (IBGE, 2019), but consisted of Cerrado formations. We still considered those located at a maximum distance of 100 Km from the boundaries of the biome since transition zones between biomes are not abrupt, but a continuum with vegetation-structure gradients.

Our field data compilation resulted in a total of 1373 plots, out of which 598 (43.5%) were forest plots, 762 (55.5%) were in savanna, and 13 (<1%) were grassland plots. In addition to the field plot data, we used airborne LiDAR footprint data available in the Rio Vermelho Watershed, state of Goiás, produced by the Sustainable Landscapes project led by the Brazilian Enterprise for Agricultural Research Corporation (EMBRAPA), which was calibrated to predict AGWB in a previous study (da Bispo et al., 2020), with highly satisfactory results ($R^2 = 0.93$, and RMSE = 6.74 t ha^{-1}). The AGWB footprint was divided into 25 × 25-m pixels,

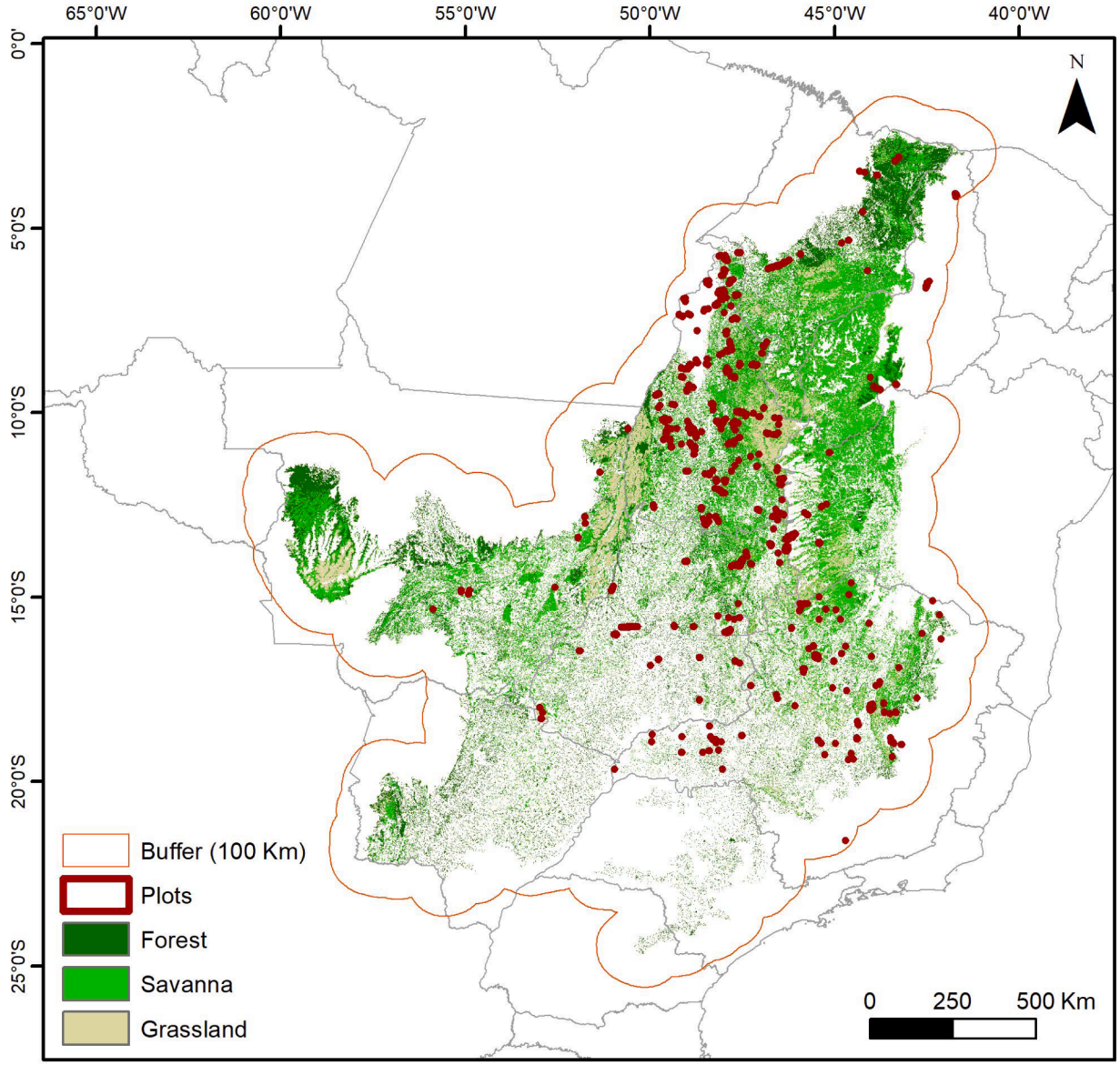

Fig. 1. Cerrado biome and the location of all compiled plots (and an airborne LiDAR footprint) with information on aboveground woody biomass. The three predominant vegetation types considered are shown according to the stable classified map by the MapBiomas initiative (Collection 5.0), over which our models will be predicted. Some plots were located outside the official Cerrado boundaries, but since they comprise sampling efforts of the Cerrado vegetation in transition regions, we considered a 100-Km buffer for inclusion in our study.

and each pixel was included as an additional sample in our dataset. This produced additional 485 samples extracted from the LiDAR footprint, which resulted in a total dataset to train the models composed of 1858 samples across the Cerrado.

The dates of sampling varied between 1995 and 2020 (Table S1). Plot size also varied. Most savanna plots had 20 × 50-m dimensions; grasslands had 10 × 10-m plots, while forest plots were the ones with the highest size variability, most of them (75% of the forest plots) being 20 × 20-m in size. Some forest plots were contiguous and were treated as a single cluster, especially in riparian forests (gallery forests), which are commonly narrow strips of forest. For instance, four 10 × 10-m plots forming a contiguous line were considered a single cluster of 10 × 40 m. Original plot sizes are presented according to location in Table S1. A few plots in riparian forests from the Tocantins state dataset had to be excluded from our study. They were composed of single 10x10-m plots presenting very high AGWB estimates due to the presence of very large trees within the small plots. These plots could not be used, since the extrapolation of these small plots to an AGWB density approach (t ha^{-1}) rendered unrealistic large values of biomass.

In order to account for errors in the geolocation of the plots, as well as account for differences in scale between the sampled plots and the imagery pixels, we built buffers around each plot/cluster of plots, so that spectral signatures were extracted from the pixels that intersected the buffers, and consisted of the mean values of these pixels. In the case of 20 × 50 m plots, we built a 25 m radius buffer around the centroid of the plot; for 10 × 10 m plots/clusters, we built a 15 m radius buffer; for 20 × 20 m plots, we built a 20 m radius buffer; and in the case of the few 100 × 100 m plots present in the dataset, we built a 50 m radius buffer.

We estimated the AGWB in each sampling plot from field-measured diameter at breast height (DBH) or at base height (Db, which is 30 cm from the base of the stem), and height (H), when available. Specific allometric equations were used for each of the three predominant formations (Table 1). For savannas (phytophysiognomies: *cerrado ralo, típico, denso,* and *rupestre*), we used the equation proposed by Roitman et al. (2018); and for grasslands, either all trees found were measured (as in the case of *campo limpo* phytophysiognomy, where woody individuals are rare), or trees with Db over 2 cm were measured and the equation by Abdala et al. (1998) was used (as in *campo sujo*, with a higher prevalence of trees). In the case of forests, several equations from Scolforo et al. (2008) were applied, depending on the vegetation physiognomy: a specific equation for deciduous forests, one for semi-deciduous forests and gallery (riparian) forests, and two for forested savannas (*cerradão*), depending on the presence of larger trees (DBH > 25 cm), following an analysis of model behavior we conducted (data not shown). All the above allometric equations are regional models built for specific Cerrado vegetation types and calibrated with data from plots located in the biome. However, because we had no information on tree height for a few plots (n = 5 in forest, and n = 48 in savanna) in the state of Minas Gerais, we also used the pantropical equation by Chave et al., (2014), which is based on tree diameter, specific wood density (ρ) and an environmental stress term (E). The specific wood density was obtained from a global database (Zanne et al., 2009), and the average value of wood density from the South America region for each species was considered (or for each genus or each family, when more specific information was not available). The E measure of environmental stress is used to compensate for the absence of tree height data by using the known effect of macroscale climatic patterns on tree diameter-height relationship, once this relationship reflects the effect of drought tolerance and tolerance to temperature variability (Chave et al., 2014).

The inclusion criterion for measurement of the woody vegetation was variable, depending on the prevalence of larger individuals and according to the vegetation type: Db ≥ 5 cm or Db ≥ 3 cm in the case of savannas; DBH ≥ 5 cm or DBH ≥ 10 cm for forests; and all trees or Db ≥ 2 cm in grasslands (Table S1).

A photographic representation of each vegetation type included in our dataset is presented in Fig. S1.

Table 1
Allometric equations used for different vegetation types and their source. Individual standard error, as well as plot-based standard error for each equation, are also presented.

Reference/ terminology	Phytophysiognomy	Vegetation type	Equation (AGWB in tonnes)
Scolforo et al. 2008 (1)	Forested savanna*	Forest	Ln(AGWB) = −11.3710317049 + 2.433521972 * Ln(D) + 0.8433902218 * Ln(H)
Scolforo et al. 2008 (2)	Forested savanna**	Forest	Ln(AGWB) = −12.2999911901 + 2.6961223975 * Ln(D) + 0.8094354054 * Ln(H)
Scolforo et al. 2008 (3)	Deciduous forest	Forest	Ln(AGWB) = −10.5940 + 1.6027*Ln(D) + 1.5879* Ln(H)
Scolforo et al. 2008 (4)	Semi-deciduous forest	Forest	Ln(AGWB) = −10.439791707 + 2.1182873001 * Ln(D) + 0.8339834928 * Ln(H)
Scolforo et al. 2008 (4)	Gallery forest	Forest	Ln(AGWB) = −10.439791707 + 2.1182873001 * Ln(D) + 0.8339834928 * Ln(H)
Chave et al. 2014	General***	General***	Ln(AGWB) = −1.803–0.976*E + 0.976 * Ln(ρ) + 2.673*Ln(D) − 0.0299*(Ln(D))^2
Roitman et al. 2018	Woodland savanna	Savanna	AGWB = ((409.0469739 * $V^{0.97545}$) * 1.17) / 1,000,000
Abdala et al. 1998	Grassland	Grassland	Ln(AGWB) = 0.9967*Ln (V) + 2.587

*Equation applied to lower biomass forested savannas with the absence of large individuals (DBH > 25 cm).
**Equation applied to higher biomass forested savannas with the presence of large individuals (DBH > 25 cm).
***Equation from Chave et al. (2014) is a pantropical equation built to cover different vegetation types. It does not consider tree height (H) but includes the terms: E (environmental stress), and ρ (wood specific gravity).

2.2.2. Remotely sensed variables

The first step before creating image composites was to define the period of analysis. Since the dataset was composed of samples obtained over a wide period of time from 1995 (for the grassland plots) to 2020, we defined two periods that matched the concentration of plots over time: one time period from 2007 to 2009 and another from 2016 to 2018 (Fig. S2). All sampling plots except for the 1995 grassland plots were less than five years apart from the two defined periods, which is enough time for significant changes in woody biomass in the Cerrado to occur (Miranda et al., 2017). In the case of the grassland plots, AGWB estimates were matched with the 2007–2009 period, because ALOS radar data is only available from 2007 and because these areas have very low woody biomass and are known to be undisturbed since 1995. Therefore, we assumed these are old-growth grasslands and are not changing greatly over time.

The optical satellite data used in our study data consisted of the Landsat-5 Thematic Mapper (TM) and the Landsat-8 Operational Land Imager (OLI) surface reflectance collection (Collection 1 Tier 1) with 30 m spatial resolution, freely available on the Google Earth Engine (GEE) platform. The images were atmospherically corrected using the Surface Reflectance Code for Landsat (LaSRC), which includes a cloud, shadow, water, and snow mask produced using the C Function of Mask (CFMASK) algorithm (Foga et al., 2017). A median composite was built for both seasons (rainy and dry) in each of the periods defined, since spectral differences between seasons may help in the prediction of AGWB (Arantes et al., 2016). From these composites, the pixel-based values of the raw bands as well as of several spectral indices were extracted. The indices were selected based on their previous use in the literature to

predict AGWB and other related vegetation structural attributes (Table 2).

The 25-m resolution L-band SAR images of backscattering coefficient acquired by the Advanced Land Observing Satellite (ALOS) and Advanced Land Observing Satellite-2 (ALOS-2) were also obtained from GEE. The scenes were pre-processed by the Japan Aerospace Exploration Agency (JAXA) and combined into seasonal mosaics of each defined period, which consist of 10 × 10° ortho-rectified tiles corrected for topographic effects (Shimada et al., 2014).

We calibrated these annual mosaics to γ0 according to the equation:

$$\gamma 0 = 10 * \log_{10}(DN)^2 + CF,$$

where γ0 is the terrain-corrected backscatter in decibels (dB), DN is the digital number in unsigned 16 bit, and CF is a calibration constant of −83.0 dB.

We applied a multi-temporal filter to reduce noise in the composite mosaics, which consisted of a 5 × 5 mean moving window (Quegan & Yu, 2001). We used the ALOS mosaic as baseline to normalize the ALOS-2 mosaic in order to increase temporal consistency. We also corrected a geolocation error (approx. 80 m) on the ALOS and ALOS-2 mosaics, by a SAR-to-SAR co-registration using a Sentinel-1 composite built with scenes from 2016 to 2018 as geolocation reference.

The variables selected as potential predictors of AGWB in the Cerrado included the SAR backscatter coefficients (HH, HV), and two SAR indices: Radar Forest Degradation Index (RDFI) and Cross-Polarised Ratio (CpR) (Mitchard et al., 2012; Shimada et al., 2014).

2.3. Modeling

Forest structure and AGWB often exhibit complex, nonlinear variations, autocorrelation, and variable interactions across temporal and spatial scales (Saatchi et al., 2011). Nonparametric methods, such as machine learning approaches, are increasingly and successfully being used to model these parameters (Baccini et al., 2008; Mutanga & Adam, 2000). Classification And Regression Trees (CART) is a machine learning, nonparametric approach, robust to premise violations, able to model both continuous and categorical variables with highly different distributions. Random Forests (RF) uses a bootstrap technique to construct multiple decision trees, with either categorical (classification tree) or continuous (regression tree) response variables (Breiman, 2001).

We implemented both a regression CART and a RF algorithm using the remote-sensing variables as predictors. Variable importance was quantitatively assessed based on a Jackknife analysis (Rodríguez-Veiga et al., 2019), but all predictor variables could be included in the modeling procedures since decision trees are robust to multicollinearity in the predictors.

We used a k-fold approach to train and cross-validate the models. All the samples were used both for training and validation. The first step in the k-fold approach was to divide the plot dataset into three evenly distributed stratified groups of samples (low, medium, and high levels of AGWB, with roughly equal numbers of observations). Then a value of k (from 1 to 10) was randomly attributed to each sample in these three groups. k-1 fold samples were selected to train the model and produce an AGWB map, while the remaining fold was used for validation. This process was run k times, where each k fold was used once for validation. The final k AGWB maps are then combined to generate a final average AGWB map. We adapted this k-fold framework from the procedures designed in previous studies (da Bispo et al., 2020; Pedro Rodríguez-Veiga et al., 2020), and the overview of the method can be seen in Fig. 2.

Lastly, we mapped the resulting predictions over a native vegetation mask, derived from the MapBiomas Collection 5.0 (mapbiomas.org) so that predictions are only made over areas within the range of sampled spectral attributes. This mask comprised all pixels classified by the MapBiomas initiative as either forest, savanna, or grassland in the year 2019, the year most recently mapped from Collection 5.0.

Table 2
Remote sensing-derived indices from optical satellites Landsat 5 TM and Landsat 8 OLI, which have been used to capture structural attributes of natural vegetation both in the rainy and the dry seasons, and which were tested in the current study as predictors of aboveground biomass in the Brazilian Cerrado.

Index	Formula	Reference
Normalized Difference Vegetation Index (NDVI)	(NIR-Red)/(NIR + Red)	Rouse et al., 1973
Enhanced Vegetation Index (EVI2)	2.4*(NIR-RED)/(NIR + Red + 1.0)	Jiang et al., 2008
Normalized Difference Water Index (NDWI)	(Green-NIR)/(Green + NIR)	Gao, 1996
Soil-adjusted Vegetation Index (SAVI)	1.5*(NIR-Red)/(NIR + Red + 0.5)	Huete, 1988
Simple Ratio (SR)	NIR/Green	Birth & McVey, 1968
Generalized Difference Vegetation Index (GDVI1)	(SR-1)/(SR + 1)	Wu, 2014
Normalized Difference Moisture Index (NDMI)	(NIR-SWIR-1)/(NIR + SWIR-1)	Cibula et al., 1992
Green Chlorophyll Vegetation Index (GCVI)	NIR/Green − 1	Gitelson et al., 2003

2.4. Uncertainty analysis

We estimated the overall error associated with the final AGWB map (ε_{AGWB}) in an error propagation approach, according to Saatchi et al. (2011) and explained in more detail in Rodríguez-Veiga et al. (2016), as follows:

$$\varepsilon_{AGWB} = (\varepsilon_{measurement}{}^2 + \varepsilon_{allometry}{}^2 + \varepsilon_{sampling}{}^2 + \varepsilon_{prediction}{}^2)^{1/2}$$

The measurement error ($\varepsilon_{measurement}$) of the tree level parameters (DBH and height measurements) averaged at plot scale is assumed as 10% according to Mitchard et al., (2011). The allometric error ($\varepsilon_{allometry}$) varied according to vegetation type (Scolforo et al., 2008), and consisted of the plot-based uncertainty (CV_{plot}) proposed by Chave et al. (2014), calculated as:

$$CVplot = CV*(\sqrt{\Sigma i[AGWB(i)]2\tilde{A}\cdot\Sigma iAGWB(i)}$$

which basically distributed the individual error estimates (CV) for each tree individual in the plot according to their contribution to total plot AGWB. This plot-based uncertainty is larger than simply averaging the individual CV of all individuals, since larger trees contribute more to the overall plot AGWB. The only cases for which we did not carry out this CV_{plot} calculation were the Minas Gerais plots for which the Chave et al. (2014) equation was applied. In these cases, we did not have access to the full raw data, and since the Chave paper already states that the uncertainty for a 1-ha plot using their equation ranges close to 10%, we considered this value as the CV_{plot} estimate for these plots. With this plot-based calculation of uncertainty, we defined an average plot-based allometric error for each vegetation type, namely 18.48% for forests, 15.48% for savannas, and 18.75% for grasslands.

The sampling error ($\varepsilon_{sampling}$), which consists of the variability of field measured AGWB across sampling plots was conservatively estimated at 45.6%, based on the sampling error of using 0.04-ha plots (our smallest plots) with a 0.09-ha (30 × 30-m) pixel based on data extrapolated from Réjou-Méchain et al. (2014). Finally, the standard deviation (SD) of the estimates of AGWB at the pixel level was used as a prediction error ($\varepsilon_{prediction}$). This error also accounts for the representativeness by the sampled plots of the true distribution of AGWB across the biome (Zhang & Lu, 2012).

2.5. AGWB comparisons with reference maps

It is worthwhile to observe how the different AGWB reference data

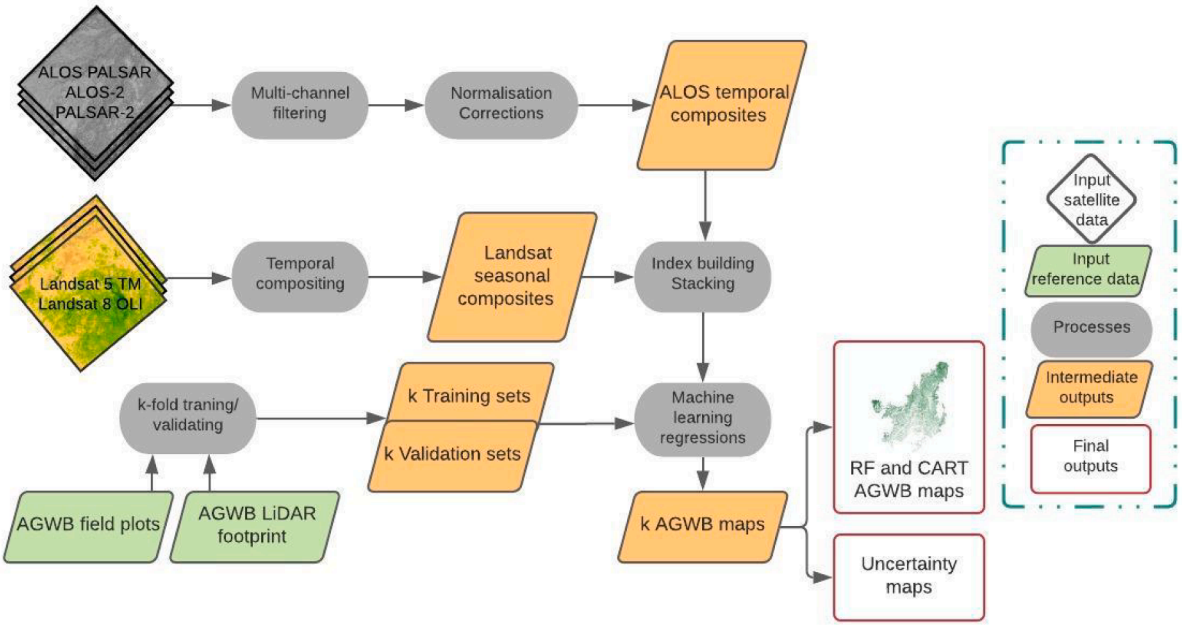

Fig. 2. Overview of the method used to produce the aboveground woody biomass (AGWB) maps for the Cerrado biome using a Random Forest algorithm (RF) and a Classification and Regression Tree algorithm (CART), adapted from the methods developed and used in Rodriguez-Veiga et al. (2020) and Bispo et al. (2020),

products available for the Cerrado compare to the maps produced in this study, even though these reference maps were estimated using different methods at different resolutions, for different time periods, and were focused on tropical forests. In order to make the maps comparable in terms of land cover proportion, which changes greatly over time, we masked all reference maps based on the Native Vegetation mask in 2019, according to MapBiomas Collection 5. We based our comparisons with six AGWB maps, five available in the literature, and one which consists of the map from the Fourth National Inventory (FNC) released by the Science, Technology, Innovation and Communication Ministry (MCTIC, 2020) at a resolution of 250 m. The FNC map is an effort to quantify the "original" carbon stocks of pristine vegetation cover (i.e. without human disturbance), according to the weighted-average estimates compiled for different native vegetation types, and mapped over a categorical "original" native vegetation class map from IBGE (2012). So in the greatest part, and with the exception of typical savanna (vegetation class 'Sa' according to IBGE), the FNC map is categorical in structure. The five maps from the literature are efforts based on remote sensing approaches, and include (a) the Saatchi et al. (2011) map at a resolution of 1 Km for the years 2000–2001; (b) the Baccini et al. (2012) map at 500-m resolution for the years 2007–2008; (c) the Avitabile et al. (2016) map, a 1-Km resolution fusion map between the two previous maps; (d) the Santoro et al. (2020), at 100-m resolution for 2010 (available at the PANGAEA platform, https://doi.org/10.1594/PANGAEA.894711); and (e) a modified Baccini map, at 30-m resolution for the year 2000 published online in 2019 (available at https://data.globalforestwatch.org/).

3. Results

Our field plot dataset was biased towards forests, presenting a higher proportion of forest plots than the available proportion of forest in the biome. The forest–savanna–grassland proportion observed in our dataset was 53%/46%/0.06%, while the proportion of these three vegetation types in the biome (according to the MapBiomas 2018 land cover map) was 31%–52%–16%. The range of AGWB obtained from the compiled plots is well distributed over the known range of AGWB in forests and savannas, and less so in grasslands (Fig. S3). It is difficult to obtain tree AGWB data for grasslands in the Cerrado, and this formation is widely understudied in comparison to other more wooded vegetation types.

3.1. AGWB modeling

The RF and CART algorithms resulted in similar models. The most important variables indicated as important for the prediction of AGWB included the green band in the dry season, the SWIR-2 band in the dry season, and NDMI in the rainy season; but their relative contribution differed between algorithms (Fig. 3 and Table S2). Radar-derived variables had lower contributions than optical-derived variables towards the models' predictions in this study.

The RF algorithm resulted in a slightly better result, according to the variation explained by the model and Root Mean Square Error (R^2 = 53%; RMSE = 36.55 t ha^{-1} [57%]), than the CART model (R^2 = 45%; RMSE = 40.35 t ha^{-1} [63%]). Our map shows an underestimation of very high AGWB (negative bias over 200 t ha^{-1}) and a slight overestimation of low AGWB (positive bias) (Fig. 4). That means that our model predictions tend quite strongly towards the overall mean, and therefore reducing our effective prediction range (Fig. 4), with -the RF model presenting slightly higher bias than the CART model (Tables 3 and 4).

The largest RMSE occurred for the lowest biomass range (Table 3), which might be related to a lack of training samples of very low biomass vegetation types, namely grasslands and low biomass savannas. This is corroborated by the highest error estimates observed for savannas (Table 4). Moreover, the greatest bias (positive) was obtained for grasslands, indicating the aforementioned overestimation of AGWB for this vegetation type. For the highest biomass range (AGWB > 100 t/ha), bias was negative, also corroborating the models' tendency towards the mean.

3.2. AGWB and uncertainty maps

According to the RF model, the predicted AGWB varied from 0 to 204 t ha^{-1} per pixel, while the CART model indicated a predicted AGWB range from 0 to 255 t ha^{-1}. The pixel-scale uncertainty (i.e. SD) estimated by the error propagation approach reached values up to 135 t ha^{-1} (255%) in the CART model, and up to 103 t ha^{-1} (54%) in the RF model. According to both resulting maps, the Cerrado harbors a total of 5.6 Gt (billion tonnes) of AGWB (with a 95% confidence interval ranging from 3 to 9 Gt), nearly equally distributed between savannas and forests

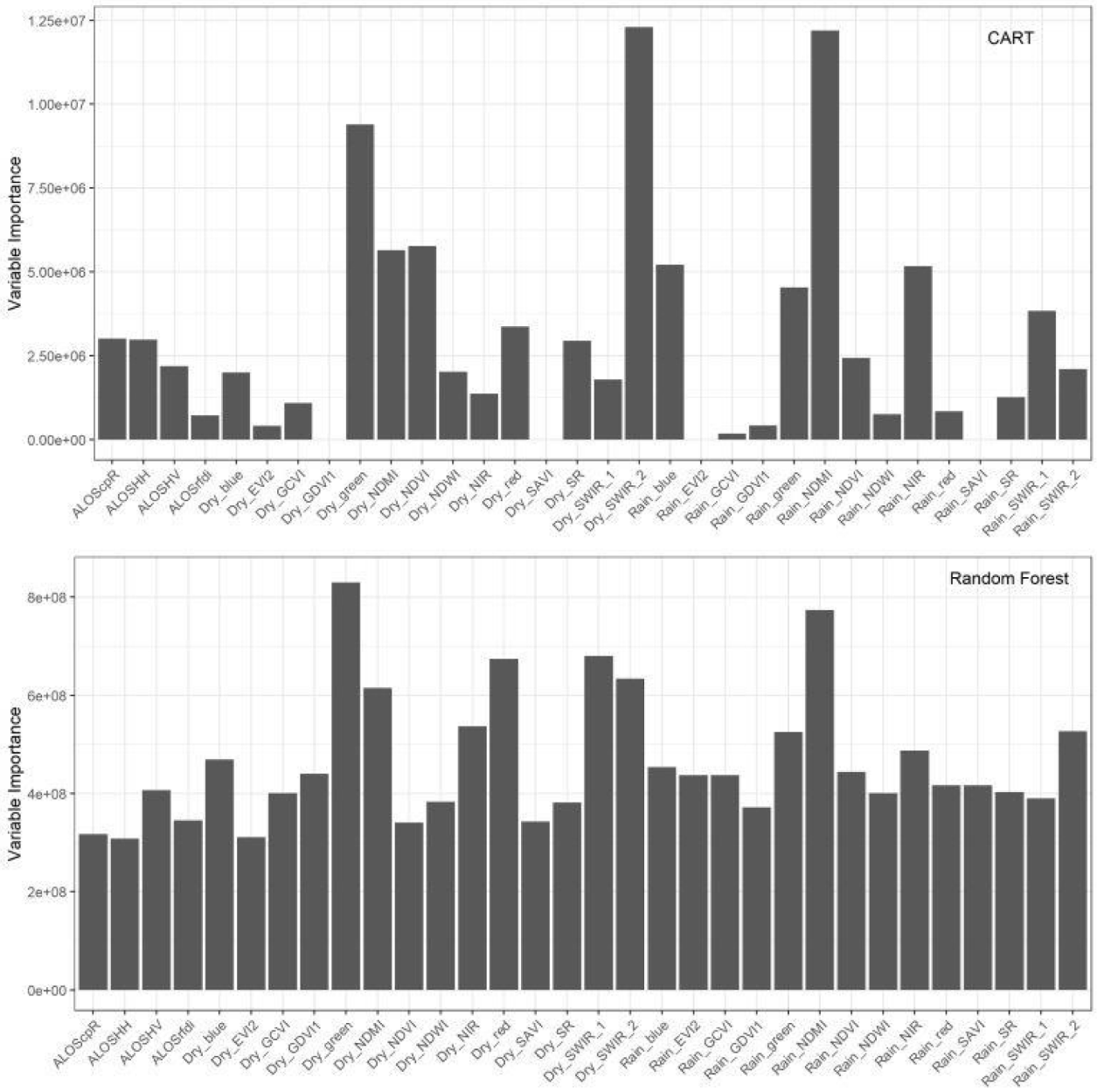

Fig. 3. Relative importance of predictors derived from Landsat (L5 and L8) and ALOS/PALSAR (ALOS and ALOS-2) for the estimation of aboveground biomass in the Cerrado biome, according to both algorithms used (CART and Random Forest). Variable importance was assessed based on a Jackknife analysis, but all predictors were included in the models. Radar-derived predictors (prefix ALOS) include backscatter coefficients (HH, HV), the Radar Forest Degradation Index (RDFI) and the Cross-Polarised Ratio (CpR). Optical predictors include raw bands (blue, green, red, NIR, SWIR-1, SWIR-2) and indices, which can be consulted in Table 2. The prefix Rain or Dry indicate the season of the optical composites used.

(approximately 47% in forests, 46% in savannas, and only 7% in grasslands).

The highest biomass estimates were obtained at the northernmost and western portions of the biome, in the transition zones of the Cerrado with the Amazon (Fig. 5). These were also the regions with the highest pixel-based error estimates, due to the highest AGWB range (Fig. 6). Overall, the error map accompanied well the range of AGWB predictions, which is to be expected, since errors are presented in absolute terms.

Large values of biomass were also observed in the southernmost parts of the Cerrado, near the transition zones with the Atlantic Forest; however, this region had a gap in our plot data distribution, so that estimates in these regions might be less reliable.

3.3. AGWB reference maps

Maps from this study and the map from Saatchi (S11) presented a more similar distribution of AGWB overall (Figs. 7 and 8). The adaptation of Baccini's original map (B12) to the current 30-m resolution map from Global Forest Watch (B19) produced an AGWB distribution tending towards the minimum value range but spreading the distribution over a larger AGWB value range as well. Avitabile's map (A16) also showed a more narrow range of biomass estimates over the biome, tending towards lower values. Santoro's map (S20) overestimates the proportion of AGWB in the lower ranges, but less so than A16 (Figs. 7 and 8). The FNC map rendered low estimates overall in the biome (Fig. 8). Also, since it is a categorical map in structure, its histogram revealed multiple peaks probably related to the mean values mapped over the most

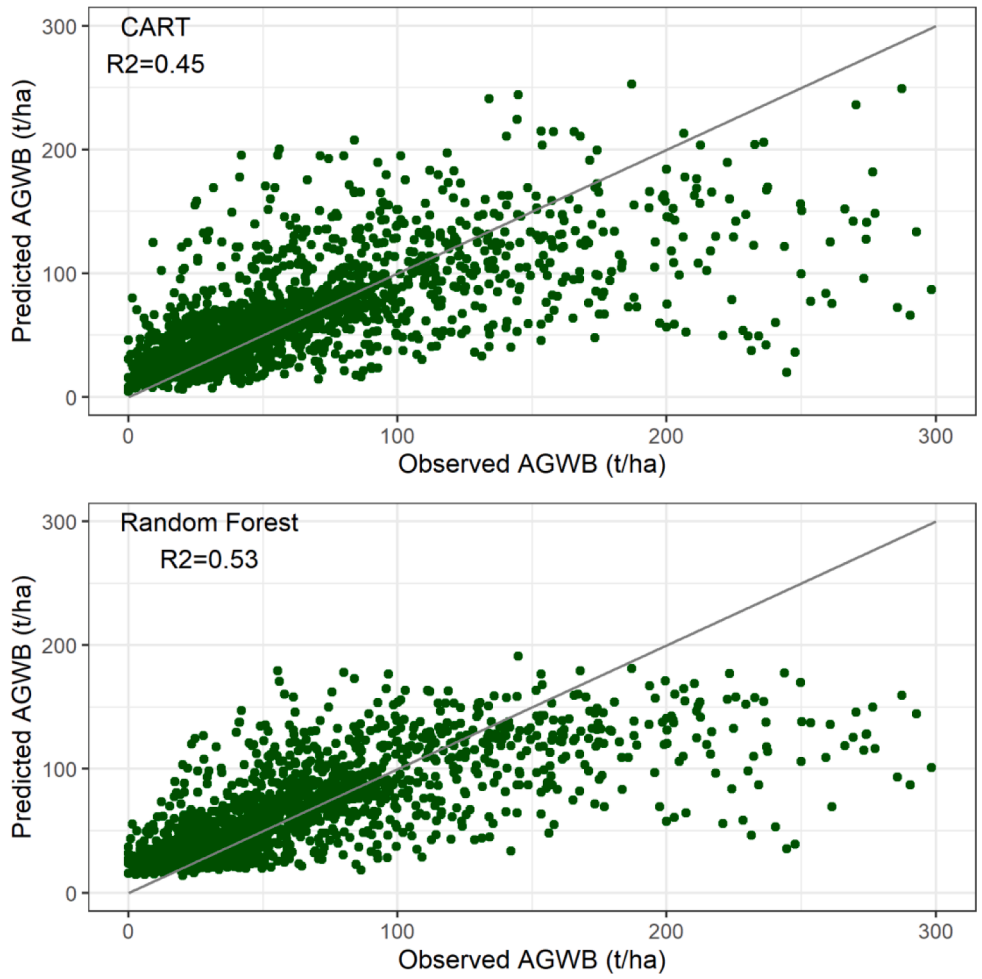

Fig. 4. Observed versus predicted aboveground woody (AGWB) according to the Classification And Regression Trees (CART) and Random Forest (RF) algorithms. The black solid line in the plots corresponds to the y = x line.

Table 3
Bias and RMSE (root-mean-square error) according to calibration data aboveground woody biomass (AGWB) range for each of the modeled algorithms (Random Forest – RF and Classification and Regression Tree – CART), as well as the mean overall RMSE and bias.

AGWB range (t ha^{-1})	Mean bias RF (t ha^{-1})	Mean bias CART (t ha^{-1})	RMSE RF (t ha^{-1} [%])	RMSE CART (t ha^{-1} [%])
0–50	13.24	12.51	22.15 [82%]	25.41 [95%]
50–100	7.05	6.51	28.60 [40%]	33.38 [47%]–
100–200	−25.8	−25.00	44.83 [32%]	51.67 [37%]
> 200	−114.48	−111.59	123.73 [52%]	126.64 [54%]
Total	1.19	0.86	36.55 [57%]	40.35 [63%]

predominant vegetation classes and is not comparable to the other histograms (but is still presented separately in Fig. S4).

As is to be expected, the greatest variations between reference maps and the ones produced in this study are located in regions with the greatest predicted estimates of AGWB (Fig. 8). The pairwise comparison between the modeled map (represented by the RF map) with the reference maps highlighted that the greatest differences between maps are located in forest areas in the transition zones with other biomes, especially in the northern and western portions of the Cerrado (Fig. S5). Different reference maps rendered slightly different patterns, but it is interesting to note that the FNC map showed the lowest differences with our modeled maps in the northern region (Fig. S5).

The overall AGWB stock in the Cerrado biome, modeled in this study, is higher than most of the reference stocks, except for B19 and S20 (Fig. 9). These references however highly underestimate the proportion of AGWB in grasslands, possibly allocating values close to null in this vegetation type, as suggested by the histograms (Fig. 7) and by the mean estimates per vegetation type (Table 5). Most reference maps show similar relative proportions of forest, savanna, and grassland, with the exception of a stark difference observed in B12, which presented the highest stocks in savannas compared to forests (Fig. 9). Our maps produced higher values of stock than the estimates of the FNC, but due to our high level of uncertainty, it is not possible to point out significant differences.

Table 4
Bias and RMSE (root-mean-square error) according to vegetation type for each of the modeled algorithms (Random Forest – RF and Classification and Regression Tree – CART), as well as the mean overall RMSE and bias.

Vegetation type	Mean bias RF (t ha^{-1})	Mean bias CART (t ha^{-1})	RMSE RF (t ha^{-1} [%])	RMSE CART (t ha^{-1} [%])
Forest	−3.83	−3.68	46.36 [49%]	50.92 [54%]
Savanna	6.95	6.18	18.56 [66%]	21.41 [77%]
Grassland	26.9	18.33	27.80 [20%]	21.77 [16%]
Total	1.19	0.86	36.55 [57%]	40.35 [63%]

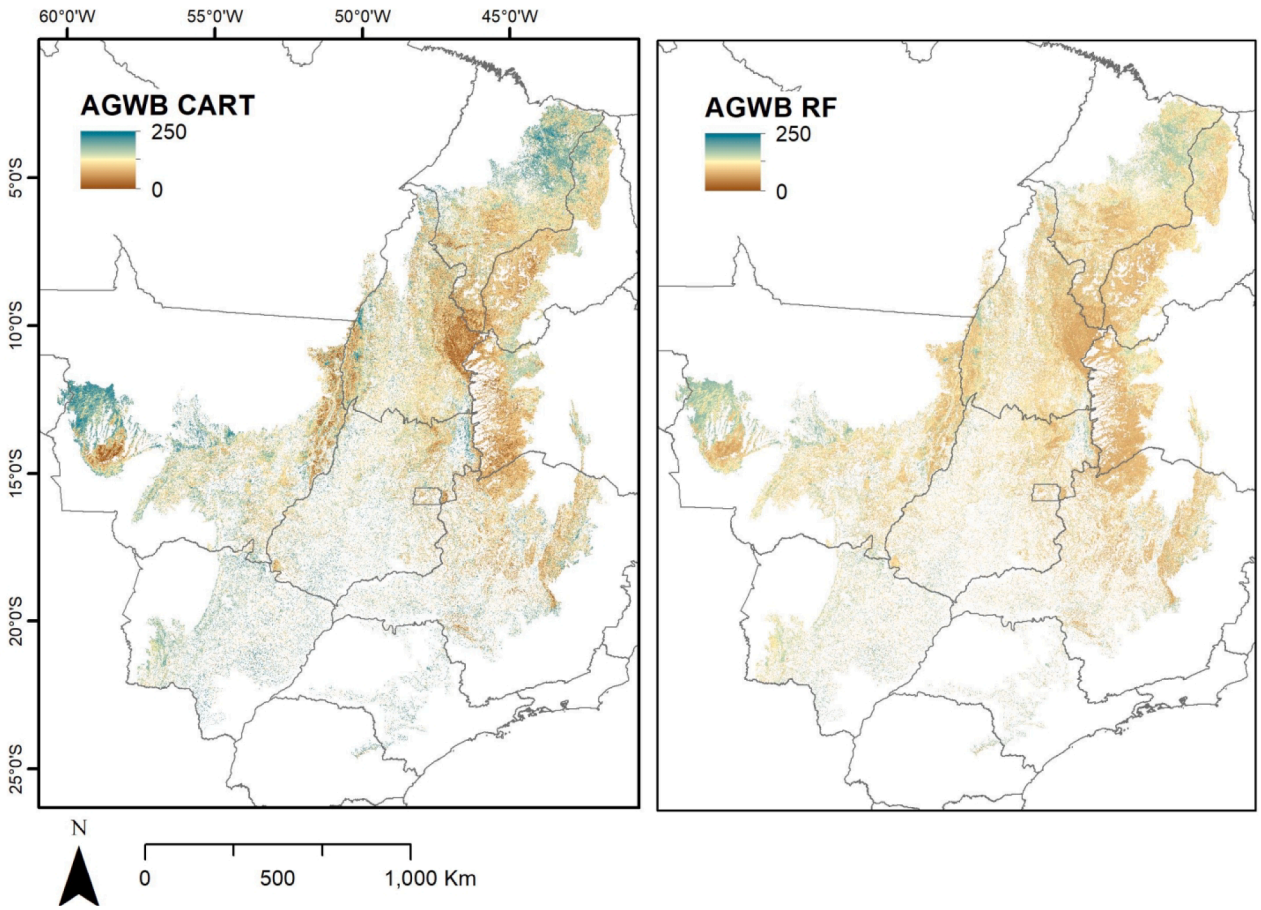

Fig. 5. Aboveground woody biomass (AGWB, in t.ha^{-1}) maps at 30-m resolution for the Cerrado biome, based on two machine learning algorithms tested (Classification And Regression Trees – CART, left panel, and Random Forest – RF, right panel). Predictions are mapped over the native vegetation pixels, classified by the MapBiomas Project Collection 5.0 (forest, savanna, and grasslands).

It is clear that the proportion of AGWB in grasslands modeled in this study is higher than most of the reference maps, except for B12 (Table 5). The general distribution of AGWB among vegetation types is more similar to the S11 reference map (Table 5), as corroborated by Fig. 9 and 10. The B19 and S20 products produced higher mean values in forests, and the FNC map produced average (and sd) values that are slightly lower but not different from our maps (Table 5).

4. Discussion

4.1. AGWB mapping improvements

A medium to high-resolution aboveground woody biomass (AGWB) map for the continental-scale Cerrado, was still lacking, after nearly a decade of mapping efforts focused on tropical rainforests in South America. Most of these tropical forest AGWB mapping initiatives that included the Cerrado presented few plots in the biome and concentrated in the transitions between the Amazon biome, overestimating biomass in central Cerrado. The reasons for this discrepancy rest foremost in the fact that the Amazon is the most sampled biome in Brazil due to it having over 80% of its original forest extent still preserved, storing the vast majority of biomass and carbon stocks in the country. On the other hand, in comparison, the Cerrado biome was still poorly sampled, and Brazilian savannas are not given enough attention by national and international forest inventory initiatives. The Brazilian Amazon, for instance, was the object of a massive tree biomass mapping initiative, conducted by the RADAM (RADAR in the Amazon) Project in the 1970s. Even if the focus was to map natural resources in the forest, the data produced by the project has been used for ecological understanding of the rainforest (e.g. Nogueira et al., 2008). More recently, the project Biomass Estimation in the Amazon (http://www.ccst.inpe.br/projetos/eba-estimativa-de-biomassa-na-amazonia/) produced a new biomass map for the Amazon based on data from new LIDAR transects, which formed the basis for the building of the FNC map for that biome. On the other hand, initiatives focused on the vegetation structure and plant communities of the Cerrado biome are fewer, but still produced a valuable set of information (e.g. Felfili et al., 2001; Morandi et al., 2018; the Federal District National Inventory available at https://www.florestal.gov.br/publicacoes/574-relatorio-inventario-florestal-nacional-df; among others), which can be used to improve on the availability of maps of aboveground biomass in the biome.

The Cerrado, after the Amazon, is the second biome in Brazil with the highest rates of conversion and degradation, which in turn makes it the second highest source of carbon emission in the country (SEEG Brasil, 2020). While the Amazon has lost about 20% of its original range, more than 50% of the native vegetation of the Cerrado has been lost (MapBiomas Collection 5), so that these conversion rates are even more alarming. The official initiative led by the MCTIC (Science, Technology, Innovation, and Communications Ministry) to quantify emissions due to land use change in the Cerrado is a progressing collaborative effort carried out by experts and researchers (MCTIC, 2020). Current estimates are based on a potential biomass map based on biomass estimates compiled from the literature for each vegetation physiognomy (IBGE, 2012). Even though this process generates a map that is categorical in structure (for all vegetation types except savanna woodlands), the resulting biomass map has a surprising amount of agreement to the maps

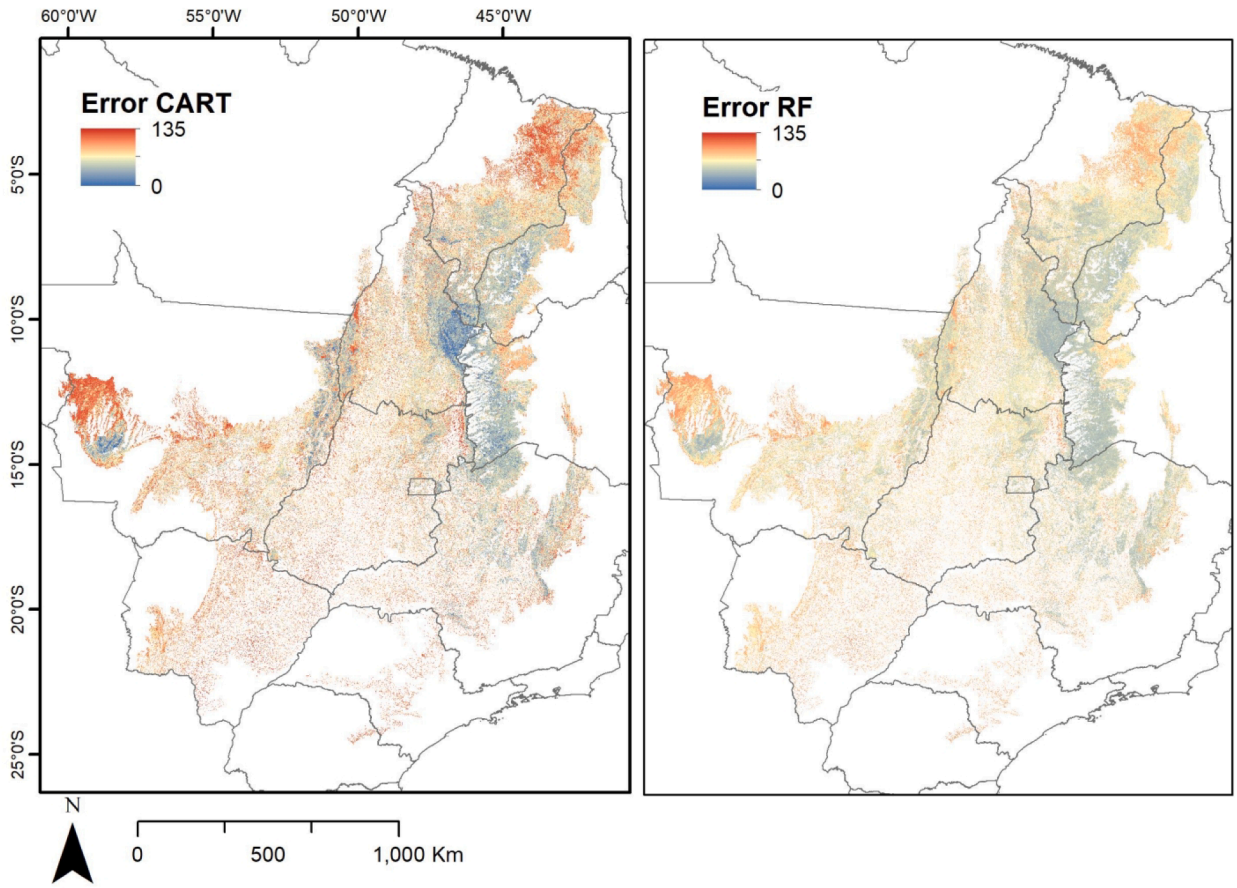

Fig. 6. Error maps of the modeled aboveground woody biomass (error AGWB, in t ha^{-1}) at 30-m resolution for the Cerrado biome, based on two machine learning algorithms tested (Classification And Regression Trees – CART, left panel, and Random Forest – RF, right panel). Predictions are mapped over the native vegetation pixels, classified by the MapBiomas Project Collection 5.0 (forest, savanna, and grasslands).

produced here, even though with a lower proportion of biomass overall (5.6 Gt of AGWB in the biome according to our models versus circa 4 Gt according to the Fourth National Communication). The maps produced by our study can substantially help advance the estimation of carbon stocks and the monitoring of greenhouse gas emissions and removals by the land-use sector in the Cerrado following the example of progressive better biomass maps for the Amazon biome.

The comparison with other widely used biomass reference maps provides further insights into the advancements the current map represents for accounting AGWB in the Cerrado biome. Existing pantropical AGWB maps, as previously mentioned, are focused and calibrated using tropical forest plot data, especially in the Amazon basin, and being biased towards underrepresented biomes such as the Cerrado. As a result, we can observe different estimations of AGWB in the Cerrado biome from map to map. The under-representation of low biomass vegetation types in the calibration of those products might result in overestimation or underestimation of AGWB values in other biomes, such as the Cerrado (da Bispo et al., 2020; Englund et al., 2017). This is what we have observed for most of the reference maps, while the product with the highest agreement to our maps was the one by Saatchi et al. (2011).

4.2. AGWB mapping challenges in the Cerrado

The exclusion of plots in riparian forests from the Tocantins state due to methodological issues (described in Section 2.2.1) may have affected the representation of high-biomass riparian forests in our dataset (maximum levels reached in our maps: 200–250 t ha^{-1}). In any case, the models were driven by spectral bands and indices, mainly, to the detriment of radar predictors. Optical satellite signatures are known to saturate over high levels of vegetation and biomass density (Avitabile et al., 2016; Mutanga & Adam, 2000; Rodríguez-Veiga et al., 2019), so that it is possible that the inclusion of higher biomass plots might not have improved our ability to predict much higher levels of AGWB, and that the stocks available in some forest types in the Cerrado may be larger than the ones estimated in this study.

Our models also showed higher bias in grasslands, in comparison to the other two vegetation types. This is predominantly due to the very low number of grassland plots in our dataset, as well as the models compensating (tendency towards the mean) for the underestimation of high AGWB levels due to saturation (da Bispo et al., 2020). Models in general, and regression trees in particular, are strongly driven by the largest stratum in the dataset (Avitabile et al., 2016). Therefore, an adequate balance between different strata samples is very influential on the model outcome (Baccini et al., 2004; Horning, 2010) In short, we indicate that our models, especially the RF model, should be better at predicting AGWB in the medium-biomass range, between 50 and 100 t ha^{-1}, which is the range that showed less bias in both models and lowest relative error. At low and high woody biomass ranges, the CART model was less biased.

The strength of optical predictors in detriment of ALOS PALSAR/ALOS-2 PALSAR-2 radar indices have not corroborated other studies [Carreiras et al. (2012) in Guinea-Bissau; Mermoz et al. (2014) in Cameroon], which were successful in predicting biomass in savannas with AGWB lower than 100 t ha^{-1}. With models over such a large extent and considering a high heterogeneity of formations, optical signatures may have proven more efficient in distinguishing between vegetation types. Optical data are more sensitive to shadow and greenness

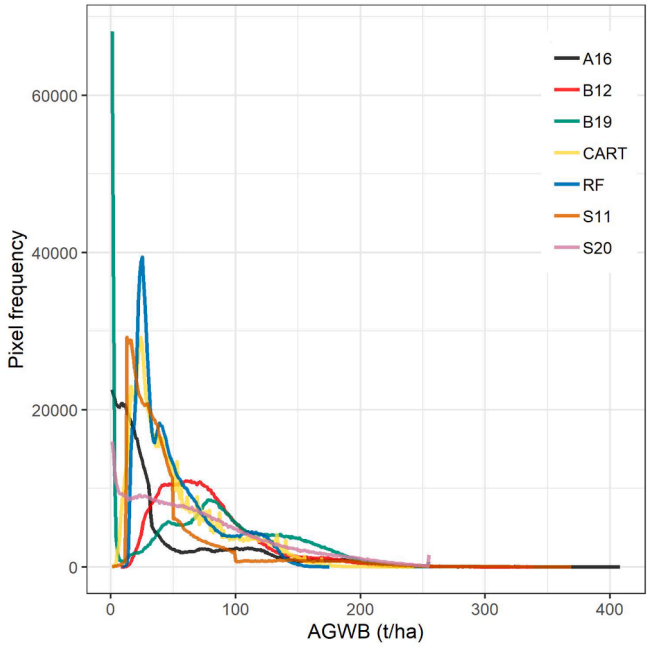

Fig. 7. Histograms of aboveground woody biomass (AGWB) distribution in the Cerrado biome, from the maps produced in this study using both algorithms (Classification And Regression Trees [CART]l, and Random Forest [RF]), and from reference maps by Saatchi et al. (2011) [S11], by Baccini et al. (2012) [B12], by Baccini from Global Forest Watch [B19], by Avitabile et al. (2016) [A16], and by Santoro et al. (2020) [S20]. For the sake of comparability, all maps were resampled to a 1-Km resolution.

(Avitabile et al., 2012; Baccini et al., 2012; Rodríguez-Veiga et al., 2016), which can be relevant to detect sparse and dry from dense and moist vegetation types, while radar data respond better to structure and tree architecture. The relevance of distinguishing between vegetation types for our predictions is corroborated by the fact that most optical factors selected by the models are from the dry season (Green and SWIR-2 bands in the CART model; and Green, Red, and SWIR-1 bands in the RF model), when differences between formations are stronger, which helped to indirectly predict woody biomass, between drier and lower-AGWB vegetation types (e.g. grasslands, savannas, deciduous forests) and moist, higher-AGWB forests. NDMI was the only optical index selected in the rainy season by both models as a very strong predictor. It measures vegetation water content and may have been relevant to distinguish between vegetation types when they are at their greatest potential moisture (Cibula et al., 1992).

We produced maps with relatively high estimated uncertainty, but we expected a high variation in the AGWB estimates, due to our effort's complex and wide-scale nature. The first source of error affecting our models is the measurement error, which consists of uncertainties deriving from field-measured features of the vegetation (especially in the case of height measurements in forests). Sampling errors, which include the plots' geolocation errors, are the highest source of error in our models. Réjou-Méchain et al. (2014) provided a method to estimate the errors resulting from the mismatch between plot-based measurements and the spectral signatures from satellite imagery, and we applied this approach conservatively (considering the smallest plot size available in our dataset). Despite using buffers to minimize the geolocation errors, differences in scale between the ground-based estimates and the pixel-based signatures occur. Allometric equations are also an important source of uncertainty. We strived to use locally and specific equations built for each vegetation type in our dataset, but as with any model, associated errors are to be expected. The report of each type of error and their respective contributions to our model's overall uncertainty are valuable for a conscientious and effective use of our maps.

4.3. Future advancements

The AGWB mapping of the Brazilian Cerrado will likely benefit from an expansion of LiDAR flights over the Cerrado in the future (da Bispo et al., 2020), and the integration with terrestrial LiDAR (Zimbres et al., 2020), both of which are at early stages in the biome. In contrast, the Amazon biome has seen a spread of LiDAR flights in recent years that have allowed the improvement of the biome's biomass stock maps. These improvements have been adopted by the public initiative of GHG estimates by the MCTIC so that the Amazon has the highest-quality estimation of GHG emissions in the country. Hopefully, such advancements will also take place for the Cerrado and other biomes.

However, LiDAR data acquisition, especially LiDAR flights, is costly, making the application to large extents and over time only possible by associating with satellite optical passive or active sensors with broad coverages. The inclusion of the recently available GEDI data into the modeling efforts will also significantly improve the model's reach in scale, and help breach the gaps in sampling in the biome.

We therefore stress that a more accurate AGWB density map for the Brazilian Cerrado will still depend on the exploration of remote sensing applications, but will highly benefit from the advancements in the necessary datasets, from field plots—especially from low biomass ranges—at the correct scales for integration with remote sensing products when feasible, to active sensor datasets (mainly terrestrial, airborne, and space-borne LiDAR). It may also be worthwhile to explore other strategies of extracting spectral signatures to relate to ground-truth information, such as object-based instead of pixel-based methods, which has been shown to improve the accuracy of forest AGWB mapping in highly heterogeneous environments (Silveira et al., 2019).

Finally, a potential vegetation biomass map, or better yet, a time series of spatially-explicit biomass stocks in all Brazilian biomes will certainly improve GHG emission and removal estimates. These estimates are necessary to measure Brazil's contributions to the climate agreements, as well as provide the basis for REDD + policies in the country. Explorations in the available methods of building dependent time-series maps of biomass are still needed, but will be beneficial for accomplishing this goal in the future.

5. Conclusions

Since most of the currently available AGWB maps that cover the Cerrado's extent were calibrated for tropical rainforests, they do not properly characterize the amount and spatial distribution of AGWB in this biome. Our locally calibrated maps are a valuable contribution to understanding the current state of AGWB stocks in the Cerrado biome, over different native vegetation types, at an unprecedented resolution and scale. This study was possible due to the application of state-of-the-art multi-sensor and analytical techniques, but mainly due to the compilation of a large and unique plot dataset from across the biome's extent, representative of the structural heterogeneity present in the Cerrado's vegetation. Our model's predictors were derived from optical multispectral signatures acquired at different seasons and L-band SAR backscatter data. The phenological differences between seasons and vegetation types appear to be more suited for the estimations AGWB in this ecosystem. This approach may also be suitable for other savannas, which are structurally complex and heterogeneous. Despite our limitations to accurately represent AGWB in very low and very high biomass vegetation types, we believe we have improved on the existing knowledge on the woody biomass stocks present in the Brazilian Cerrado, especially in the savanna woodlands, the predominant and most threatened vegetation type in the biome.

CRediT authorship contribution statement

Barbara Zimbres: Conceptualization, Methodology, Formal analysis, Writing - original draft. **Pedro Rodríguez-Veiga:**

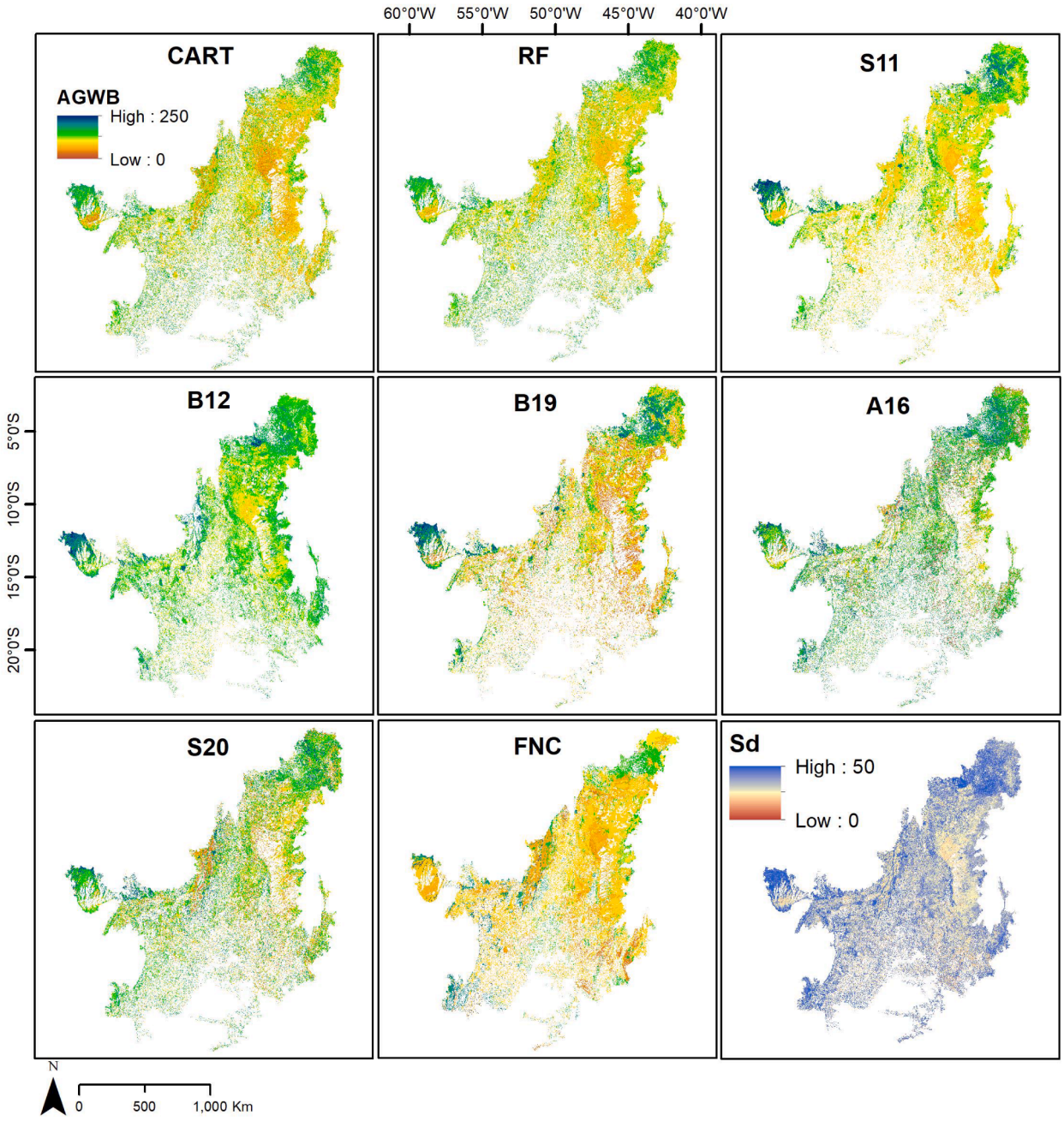

Fig. 8. Aboveground woody biomass (AGWB, in t ha^{-1}) maps for the Cerrado biome, including the ones produced by this study using both algorithms (Classification And Regression Trees [CART]l, and Random Forest [RF]), and reference maps by Saatchi et al. (2011) [S11], by Baccini et al. (2012) [B12], by Baccini from Global Forest Watch [B19], by Avitabile et al. (2016) [A16], by Santoro et al. (2020) [S20], and by the MCTIC (2020) [FNC]. A standard deviation of AGWB (t ha^{-1}) considering all maps is also presented.

Conceptualization, Methodology, Software, Formal analysis, Supervision, Writing - review & editing. **Julia Z. Shimbo:** Conceptualization, Supervision, Writing - review & editing. **Polyanna Conceição Bispo:** Conceptualization, Methodology, Formal analysis, Writing - review & editing. **Heiko Balzter:** Conceptualization, Supervision, Writing - review & editing. **Mercedes Bustamante:** Conceptualization, Supervision, Writing - review & editing. **Iris Roitman:** Data curation, Investigation, Writing - review & editing. **Ricardo Haidar:** Data curation, Investigation, Writing - review & editing. **Sabrina Miranda:** Data curation, Investigation, Writing - review & editing. **Letícia Gomes:** Data curation, Investigation, Writing - review & editing. **Fabrício Alvim Carvalho:** Data curation, Investigation, Writing - review & editing. **Eddie Lenza:** Data curation, Investigation, Writing - review & editing. **Leonardo Maracahipes-Santos:** Data curation, Investigation, Writing - review & editing. **Ana Clara Abadia:** Data curation, Investigation, Writing - review & editing. **Jamir Afonso Prado Júnior:** Data curation, Investigation, Writing - review & editing. **Evandro Luiz Mendonça Machado:** Data curation, Investigation, Writing - review & editing. **Anne Priscila Dias Gonzaga:** Data curation, Investigation, Writing - review & editing. **Marcela Castro Nunes Santos Terra:** Data curation, Investigation, Writing - review & editing. **José Marcio Mello:** Data curation, Investigation, Writing - review & editing. **José Roberto Soares Scolforo:** Data curation, Investigation. **José Roberto Rodrigues Pinto:** Data curation, Investigation, Writing - review & editing. **Ane Alencar:** Conceptualization, Writing - review & editing.

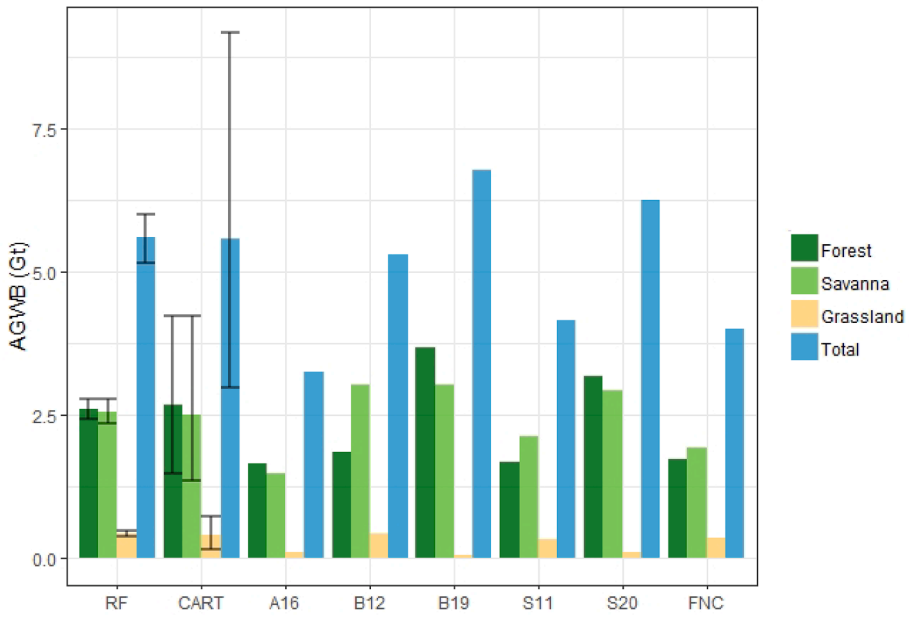

Fig. 9. Aboveground woody biomass (AGWB, in Gt or billion tonnes) in the Cerrado biome (total), and each vegetation type considered in this study, derived from the maps produced by this study using both algorithms (Classification And Regression Trees [CART]l, and Random Forest [RF]), and from reference maps by Saatchi et al. (2011) [S11], by Baccini et al. (2012) [B12], by Baccini from Global Forest Watch [B19], by Avitabile et al. (2016) [A16], by Santoro et al. (2020) [S20], and by the Fourth National Communication [FNC]. Confidence intervals (95%) are shown for our modeled maps.

Table 5
Mean [standard deviation] of aboveground woody biomass (AGWB, in t ha^{-1}) In each of the three vegetation types, according to the MapBiomas Collection 5 native vegetation mask for 2019. Statistics were derived from the maps produced by this study using both algorithms [RF and CART], and from the reference maps by Saatchi et al. (2011) [S11], by Baccini et al. (2012) [B12], by Baccini from Global Forest Watch [B19], by Avitabile et al. (2016) [A16], by Santoro et al. (2020) [S20], and by the Fourth National Communication [FNC].

Vegetation type	RF	CART	A16	B12	B19	S11	S20	FNC
Forest	86.27 [30.38]	88.36 [36.03]	55.33 [66.01]	64.8 [67.18]	121.99 [51.25]	56.05 [62.19]	105.76 [61.22]	65.80 [57.36]
Savanna	41.29 [20.8]	40.4 [25.24]	24.04 [36.77]	51.87 [45.42]	49.03 [45.88]	34.5 [36.26]	47.43 [43.4]	34.28 [35.24]
Grassland	24.99 [12.09]	23.61 [17.19]	5.9 [14.27]	27.15 [28.53]	3.5 [15.7]	18.89 [14.66]	6.63 [18.33]	22.38 [23.69]

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by the Forests 2020 project of the UK Space Agency's International Partnerships Programme (IPP) under the Global Challenge Research Fund (GCRF). Pedro Rodríguez-Veiga and Heiko Balzter were also supported by the Natural Environment Research Council's National Centre for Earth Observation (NCEO), and the European Space Agency (ESA) Biomass Climate Change Initiative (CCI+) project. M Bustamante received support from the National Institute of Science and Technology for Climate Change Phase 2 under CNPq Grant 465501/2014-1, FAPESP Grant 2014/50848-9, and the National Coordination for High Level Education and Training (CAPES) Grant 16/2014. The authors JRR Pinto (Process Number 307340/2017-1), M Bustamante (Process Number 307768/2017-1), and E Lenza (Process Number 308236/2017-3) are supported with scholarship productivity award. Fabrício Alvim Carvalho is supported with a scholarship productivity award by the Brazilian National Research Council (CNPq).

References

Abdala, G.C., Caldas, L.S., Haridasan, M., Eiten, G., 1998. Above and belowground organic matter and root: shoot ratio in a cerrado in Central Brazil. Braz. J. Ecol. 2 (1), 11–23.

Alencar, A., Z. Shimbo, J., Lenti, F., Balzani Marques, C., Zimbres, B., Rosa, M., Arruda, V., Castro, I., Fernandes Márcico Ribeiro, J. P., Varela, V., Alencar, I., Piontekowski, V., Ribeiro, V., M. C. Bustamante, M., Eyji Sano, E., Barroso, M., 2020. Mapping Three Decades of Changes in the Brazilian Savanna Native Vegetation Using Landsat Data Processed in the Google Earth Engine Platform. Remote Sens. 12(6), 924. https://doi.org/10.3390/rs12060924.

de Almeida, C.T., Galvão, L.S., Cruz, D.O., Pierre, J., Balbaud, H., Daniele, A., Rocha, F., Pereira, D.S., Yumie, L., Pontes, A., Maurício, P., Alencastro, L. De, Valéria, C., Silva, D.J., Longo, M., 2019. Combining LiDAR and hyperspectral data for aboveground biomass modeling in the Brazilian Amazon using different regression algorithms. Remote Sens. Environ. 232 (July) https://doi.org/10.1016/j.rse.2019.111323.

Alvares, C.A., Stape, J.L., Sentelhas, P.C., De Moraes Gonçalves, J.L., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorol. Z. 22 (6), 711–728. https://doi.org/10.1127/0941-2948/2013/0507.

Alvarez, E., Duque, A., Saldarriaga, J., Cabrera, K., de las Salas, G., del Valle, I., Lema, A., Moreno, F., Orrego, S., Rodríguez, L., 2012. Tree above-ground biomass allometries for carbon stocks estimation in the natural forests of Colombia. Forest Ecol. Manage., 267, 297–308. https://doi.org/10.1016/j.foreco.2011.12.013.

Arantes, A.E., Ferreira, L.G., Coe, M.T., 2016. The seasonal carbon and water balances of the Cerrado environment of Brazil: past, present, and future influences of land cover and land use. ISPRS J. Photogramm. Remote Sens. 117, 66–78. https://doi.org/10.1016/j.isprsjprs.2016.02.008.

Assad, E.D., 1994. Chuva nos cerrados: análise e espacialização. Embrapa-CPAC.

Avitabile, V., Baccini, A., Friedl, M.A., Schmullius, C., 2012. Capabilities and limitations of Landsat and land cover data for aboveground woody biomass estimation of Uganda. Remote Sens. Environ. 117, 366–380. https://doi.org/10.1016/j.rse.2011.10.012.

Avitabile, V., Herold, M., Heuvelink, G.B.M., Lewis, S.L., Phillips, O.L., Asner, G.P., Armston, J., Ashton, P.S., Banin, L., Bayol, N., Berry, N.J., Boeckx, P., de Jong, B.H. J., Devries, B., Girardin, C.A.J., Kearsley, E., Lindsell, J.A., Lopez-Gonzalez, G., Lucas, R., Willcock, S., 2016. An integrated pan-tropical biomass map using multiple reference datasets. Glob. Change Biol. 22 (4), 1406–1420. https://doi.org/10.1111/gcb.13139.

Azevedo, T., Rosa, M.R., Shimbo, J.Z. Oliveira, M.G., 2021. Annual report on deforestation in Brazil 2020. MapBiomas. 93 p.

Baccini, A., Friedl, M.A., Woodcock, C.E., Warbington, R., 2004. Forest biomass estimation over regional scales using multisource data. Geophys. Res. Lett. 31, L10501. https://doi.org/10.1029/2004GL019782.

Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P.S.A., Dubayah, R., Friedl, M.A., Samanta, S., Houghton, R.A.,

2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nat. Clim. Change 2 (3), 182–185. https://doi.org/10.1038/nclimate1354.

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 (4), 045011 https://doi.org/10.1088/1748-9326/3/4/045011.

Baskerville, G.L., 1972. Use of logarithmic regression in the estimation of plant biomass. Can. J. Forestry 2, 49–53.

Birth, G.S., McVey, G.R., 1968. Measuring the color of growing turf with a reflectance spectrophotometer 1. Agron. J. 60 (6), 640–643. https://doi.org/10.2134/agronj1968.00021962006000060016x.

Bispo, P.C., Santos, J.R., Valeriano, M.M., Touzi, R., Seifert, F.M., 2014. Integration of polarimetric PALSAR attributes and local geomorphometric variables derived from SRTM for forest biomass modeling in central amazonia. Can. J. Remote Sens. 40 (1), 26–42. https://doi.org/10.1080/07038992.2014.913477.

Bispo, P. D. C., Pardini, M., Papathanassiou, K. P., Kugler, F., Balzter, H., Rains, D., dos Santos, J. R., Rizaev, I. G., Tansey, K., dos Santos, M. N., & Spinelli Araujo, L., 2019. Mapping forest successional stages in the Brazilian Amazon using forest heights derived from TanDEM-X SAR interferometry. Remote Sens. Environ. 232(July 2018), 111194. https://doi.org/10.1016/j.rse.2019.05.013.

Bispo, P. da C., Rodríguez-Veiga, P., Zimbres, B., do Couto de Miranda, S., Henrique Giusti Cezare, C., Fleming, S., Baldacchino, F., Louis, V., Rains, D., Garcia, M., Del Bon Espírito-Santo, F., Roitman, I., Pacheco-Pascagaza, A. M., Gou, Y., Roberts, J., Barrett, K., Ferreira, L. G., Shimbo, J. Z., Alencar, A., … Balzter, H., 2020. Woody Aboveground Biomass Mapping of the Brazilian Savanna with a Multi-Sensor and Machine Learning Approach. Remote Sensing, 12(17), 2685. https://doi.org/10.3390/rs12172685.

Bourlière, F., Hadley, M., 1983. Present-day savannas: an overview. In: Bourlière, F. (Ed.), Ecosystems of the world. 13. Tropical savannas. Elsevier Scientific Publishing Company, pp. 1–17.

Breiman, L., 2001. Random forests. Machine Learning. 45, 5–32.

Carreiras, J.M.B., Vasconcelos, M.J., Lucas, R.M., 2012. Understanding the relationship between aboveground biomass and ALOS PALSAR data in the forests of Guinea-Bissau (West Africa). Remote Sens. Environ. 121, 426–442. https://doi.org/10.1016/j.rse.2012.02.012.

Castro, L. H. R., Moreira, A. M., Assad, E. D. (1994). Definição e regionalização dos padrões pluviométricos dos cerrados brasileiros. In: ASSAD, E.D. Chuvas nos cerrados: análise e espacialização. Brasília, Embrapa-CPAC/Embrapa-SPI, 423 p.

Chang, J., Shoshany, M. (2016). Mediterranean shrublands biomass estimation using Sentinel-1 and Sentinel-2. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, pp. 5300–5303.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.P., Nelson, B.W., Ogawa, H., Puig, H., Riéra, B., Yamakura, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145 (1), 87–99. https://doi.org/10.1007/s00442-005-0100-x.

Chave, Jérôme, Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W. B.C., Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martínez-Yrízar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Glob. Change Biol. 20 (10), 3177–3190. https://doi.org/10.1111/gcb.12629.

Chen, L., Ren, C., Zhang, B., Wang, Z., Xi, Y., 2018. Estimation of forest above-ground biomass by geographically weighted regression and machine learning with sentinel imagery. Forests 9 (10), 582. https://doi.org/10.3390/f9100582.

Cibula, W.G., Zetka, E.F., Rickman, D.L., 1992. Response of thematic mapper bands to plant water stress. Int. J. Remote Sens. 13(10), 1869–1880. https://doi.org/10.1080/01431169208904236.

Coutinho, L.M., 2002. O bioma do cerrado. In: Eugen Warming e o cerrado brasileiro: um século depois (pp. 77–91). UNESP.

Englund, O., Sparovek, G., Berndes, G., Freitas, F., Ometto, J.P., Oliveira, P.V.D.C.E., Costa, C., Lapola, D., 2017. A new high-resolution nationwide aboveground carbon map for Brazil. Geo: Geogr. Environ. 4 (2), 1–12. https://doi.org/10.1002/geo2.45.

Feldpausch, T.R., Banin, L., Phillips, O.L., Baker, T.R., Lewis, S.L., Quesada, C.A., Affum-Baffoe, K., Arets, E.J.M.M., Berry, N.J., Bird, M., Brondizio, E.S., De Camargo, P., Chave, J., Djagbletey, G., Domingues, T.F., Drescher, M., Fearnside, P.M., França, M. B., Fyllas, N.M., Lloyd, J., 2011. Height-diameter allometry of tropical forest trees. Biogeosciences 8 (5), 1081–1106. https://doi.org/10.5194/bg-8-1081-2011.

Felfili, J.M., Silva Junior, M.C., Rezende, A.V., Haridasan, M., Filgueiras, T.S., Mendonça, R.C., Walter, B.M.T., Nogueira, P.E., 2001. O projeto Biogeografia do Bioma Cerrado: hipóteses e padronização da metodologia. In: Garay, I., Dias, B. (Eds.), Conservação da biodiversidade em ecossistemas tropicais: avanços conceituais e revisão de metodologias de avaliação e monitoramento. Editora Vozes, pp. 157–173.

Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Joseph Hughes, M., Laue, B., 2017. Cloud detection algorithm comparison and validation for operational Landsat data products. Remote Sens. Environ. 194, 379–390. https://doi.org/10.1016/j.rse.2017.03.026.

Françoso, R.D., Brandão, R., Nogueira, C.C., Salmona, Y.B., Machado, R.B., Colli, G.R., 2015. Habitat loss and the effectiveness of protected areas in the Cerrado Biodiversity Hotspot. Natureza & Conservação 13 (1), 35–40. https://doi.org/10.1016/j.ncon.2015.04.001.

Gao, B., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58 (3), 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3.

Gitelson, A.A., Viña, A., Arkebauer, T.J., Rundquist, D.C., Keydan, G., Leavitt, B., 2003. Remote estimation of leaf area index and green leaf biomass in maize canopies. Geophys. Res. Lett. 30 (5), n/a-n/a. https://doi.org/10.1029/2002GL016450.

Horning, N., 2010. Random Forests: An algorithm for image classification and generation of continuous fields data sets. International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences.

Houghton, R.A., 2005. Aboveground forest biomass and the global carbon balance. Glob. Change Biol. 11 (6), 945–958. https://doi.org/10.1111/j.1365-2486.2005.00955.x.

Houghton, R.A., Butman, D., Bunn, A.G., Krankina, O.N., Schlesinger, P., Stone, T.A., 2011. Mapping Russian forest biomass with data from satellites and forest inventories. https://doi.org/10.1088/1748-9326/2/4/045032.

Huete, A., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25 (3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.

IBGE, 2019. Mapa de Biomas e Sistema Costeiro-Marinho do Brasil - 1:250 000. IBGE, Coordenação de Recursos Naturais e Estudos Ambientais, https://www.ibge.gov.br/geociencias/cartas-e-mapas/informacoes-ambientais/15842-biomas.html?=&t=acesso-ao-produto. Accessed 5 July 2020.

IBGE, 2012. Manual técnico da vegetação brasileira: sistema fitogeográfico, inventário das formações florestais e campestres, técnicas e manejo de coleções botânicas, procedimentos para mapeamentos. IBGE, Coordenação de Recursos Naturais e Estudos Ambientais. 2. ed., Rio de Janeiro, 276 p.

Jiang, Z., Huete, A., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sens. Environ. 112 (10), 3833–3845. https://doi.org/10.1016/j.rse.2008.06.006.

Joshi, N., Mitchard, E.T.A., Brolly, M., Schumacher, J., Fernández-Landa, A., Johannsen, V.K., Marchamalo, M., Fensholt, R., 2017. Understanding 'saturation' of radar signals over forests. Sci. Rep. 7 (1), 3505. https://doi.org/10.1038/s41598-017-03469-3.

Le Toan, T., Beaudoin, A., Riom, J., Guyon, D., 1992. Relating forest biomass to SAR data. IEEE Trans. Geosci. Remote Sens. 30 (2), 403–411. https://doi.org/10.1109/36.134089.

Leitão, P. J., Schwieder, M., Pötzschner, F., Pinto, J. R. R., Teixeira, A. M. C., Pedroni, F., Sanchez, M., Rogass, C., van der Linden, S., Bustamante, M. M. C., & Hostert, P., 2018. From sample to pixel: multi-scale remote sensing data for upscaling aboveground carbon data in heterogeneous landscapes. Ecosphere, 9(8). https://doi.org/10.1002/ecs2.2298.

MapBiomas Project – Collection 5.0 of the Annual Time-series of Land Cover and Land Use Maps in Brazil, https://mapbiomas.org. Accessed 31 October 2020.

MCTI – Ministério da Ciência e Tecnologia e Inovação, (2020). Quarta Comunicação Nacional do Brasil à Convenção-Quadro das Nações Unidas sobre Mudança do Clima, https://sirene.mctic.gov.br/portal/export/sites/sirene/backend/galeria/arquivos/2020/12/2020_12_22_4CN_v5_Ingles.pdf. Accessed 10 January 2021.

Mermoz, S., Le Toan, T., Villard, L., Réjou-Méchain, M., Seifert-Granzin, J., 2014. Biomass assessment in the Cameroon savanna using ALOS PALSAR data. Remote Sens. Environ. 155, 109–119. https://doi.org/10.1016/j.rse.2014.01.029.

Miranda, S. do C. de, De-Carvalho, P.S., Bustamante, M.M. da C., Silva Junior, M.C.da, 2017. Variação temporal na estrutura da vegetação lenhosa de cerrado sentido restrito sobre Neossolos Quartzarênicos. Espacios, 38(4), 3.

Mitchard, E.T.A., Saatchi, S.S., Lewis, S.L., Feldpausch, T.R., Woodhouse, I.H., Sonké, B., Rowland, C., Meir, P., 2011. Measuring biomass changes due to woody encroachment and deforestation/degradation in a forest–savanna boundary region of central Africa using multi-temporal L-band radar backscatter. Remote Sens. Environ. 115 (11), 2861–2873. https://doi.org/10.1016/j.rse.2010.02.022.

Mitchard, E.T.A., Saatchi, S.S., White, L.J.T., Abernethy, K.A., Jeffery, K.J., Lewis, S.L., Collins, M., Lefsky, M.A., Leal, M.E., Woodhouse, I.H., Meir, P., 2012. Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: Overcoming problems of high biomass and persistent cloud. Biogeosciences 9 (1), 179–191. https://doi.org/10.5194/bg-9-179-2012.

Morandi, P.S., Marimon, B.S., Marimon-Junior, B.H., Ratter, J.A., Feldpausch, T.R., Colli, G.R., Munhoz, C.B.R., da Silva Júnior, M.C., de Souza Lima, E., Haidar, R.F., Arroyo, L., Murakami, A.A., de Góis Aquino, F., Walter, B.M.T., Ribeiro, J.F., Françoso, R., Elias, F., de Oliveira, E.A., Reis, S.M., Phillips, O.L., 2018. Tree diversity and above-ground biomass in the South America Cerrado biome and their conservation implications. Biodivers. Conserv. 29 (5), 1519–1536. https://doi.org/10.1007/s10531-018-1589-8.

Mutanga, O., Adam, E., 2000. High density biomass estimation: testing the utility of Vegetation Indices and the Random Forest Regression algorithm. 1–4.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858. https://doi.org/10.1038/35002501.

Nogueira, E.M., Fearnside, P.M., Nelson, B.W., Barbosa, R.I., Keizer, E.W.H., 2008. Estimates of forest biomass in the Brazilian Amazon: New allometric equations and adjustments to biomass from wood-volume inventories. For. Ecol. Manage. 256 (11), 1853–1867. https://doi.org/10.1016/j.foreco.2008.07.022.

Oliveira Filho, A.T., Scolforo, J.R., 2008. Compilação e caracterização das espécies arbóreas da flora nativa de Minas Gerais. In: Oliveira Filho, A.T., Scolforo, J.R. (Eds.), Espécies Arbóreas da Flora Nativa (pp. 1–8). Editora UFLA.

Omar, H., Misman, M.A., Kassim, A.R., 2017. Synergetic of PALSAR-2 and Sentinel-1A sar polarimetry for retrieving aboveground biomass in dipterocarp forest of Malaysia. Appl. Sci. 7 (7), 675. https://doi.org/10.3390/app7070675.

Ottmar, R.D., Vihnanek, R.E., Miranda, H.S., Sato, M.N., Andrade, S.M.A., 2001. Séries de estereo-fotografias para quantificar a biomassa da vegetação do Cerrado do Brasil Central, vol. I. USDA/USAID/UnB. Gen. Tech. Rep, Portland: US Department of Agriculture, Forest Service.

Ouchi, K., 2013. Recent trend and advance of synthetic aperture radar with selected topics. Remote Sens. 5 (2), 716–807. https://doi.org/10.3390/rs5020716.

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 (5), 1053–1068. https://doi.org/10.1016/j.rse.2009.12.018.

Quegan, S., Yu, J.J., 2001. Filtering of multichannel SAR images. IEEE Trans. Geosci. Remote Sens. 39 (11), 2373–2379. https://doi.org/10.1109/36.964973.

Réjou-Méchain, M., Muller-Landau, H.C., Detto, M., Thomas, S.C., Le Toan, T., Saatchi, S. S., Barreto-Silva, J.S., Bourg, N.A., Bunyavejchewin, S., Butt, N., Brockelman, W.Y., Cao, M., Cárdenas, D., Chiang, J.-M., Chuyong, G.B., Clay, K., Condit, R., Dattaraja, H.S., Davies, S.J., Chave, J., 2014. Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks. Biogeosciences 11 (23), 6827–6840. https://doi.org/10.5194/bg-11-6827-2014.

Rezende, A.V., Do Vale, A.T., Sanquetta, C.R., Figueiredo Filho, A., Felfili, J.M., 2006. Comparison of mathematical models to volume, biomass and carbon stock estimation of the woody vegetation of a cerrado sensu stricto in Brasília, DF. Scientia Forestalis/Forest Sciences 71, 65–76.

Ribeiro, J.F., Walter, B.M.T., 2008. As principais fitofisionomias do bioma Cerrado. In: Sano, S.M., Almeida, S.P., Ribeiro, J.F. (Eds.), Cerrado: ecologia e flora (2nd ed., pp. 151–212). Embrapa Informação Tecnológica.

Rodríguez-Veiga, P., Barbosa-Herrera, A. P., Barreto-Silva, J. S., Bispo, P. C., Cabrera, E., Capachero, C., Galindo, G., Gou, Y., Moreno, L. M., Louis, V., Lozano, P., Pacheco-Pascagaza, A. M., Pachon-Cendales, I. P., Phillips-Bernal, J. F., Roberts, J., Salinas, N. R., Vergara, L., Zuluaga, A. C., Balzter, H, 2019. MAPPING THE SPATIAL DISTRIBUTION OF COLOMBIA'S FOREST ABOVEGROUND BIOMASS USING SAR AND OPTICAL DATA. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-3/W7(3/W7), 57–60. https://doi.org/10.5194/isprs-archives-XLII-3-W7-57-2019.

Rodríguez-Veiga, Pedro, Carreiras, J., Smallman, T.L., Exbrayat, J.F., Ndambiri, J., Mutwiri, F., Nyasaka, D., Quegan, S., Williams, M., Balzter, H., 2020. Carbon stocks and fluxes in kenyan forests and wooded grasslands derived from earth observation and model-data fusion. Remote Sensing 12 (15). https://doi.org/10.3390/RS12152380.

Rodríguez-Veiga, Pedro, Saatchi, S., Tansey, K., Balzter, H., 2016. Magnitude, spatial distribution and uncertainty of forest biomass stocks in Mexico. Remote Sens. Environ. 183, 265–281. https://doi.org/10.1016/j.rse.2016.06.004.

Roitman, I., Bustamante, M.M.C., Haidar, R.F., Shimbo, J.Z., Abdala, G.C., Eiten, G., Fagg, C.W., Felfili, M.C., Felfili, J.M., Jacobson, T.K.B., Lindoso, G.S., Keller, M., Lenza, E., Miranda, S.C., Pinto, J.R.R., Rodrigues, A.A., Delitti, W.B.C., Roitman, P., Sampaio, J.M., 2018. Optimizing biomass estimates of savanna woodland at different spatial scales in the Brazilian Cerrado: Re-evaluating allometric equations and environmental influences. PLoS ONE 13 (8), e0196742. https://doi.org/10.1371/journal.pone.0196742.

Rouse, J.W., Haas, R.S., Schell, J.A., Deering, D.W., 1973. Monitoring vegetation systems in the Great Plains with ERTS. Proc. 3rd ERT S Symposium 1, 48–62.

Russo, G., Alencar, A., Ribeiro, V., Amorim, C., Shimbo, J. Lenti, F., Castro, I., 2018. Cerrado: The Brazilian savanna's contribution to GHG emissions and to climate solutions. Policy Brief. December , 2018, IPAM. https://ipam.org.br/wp-content/uploads/2018/12/Policy-Brief-Cerrado-COP24-en-1.pdf.

Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M., Morel, A., 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. Proc. Natl. Acad. Sci. 108 (24), 9899–9904. https://doi.org/10.1073/pnas.1019576108.

Sano, E.E., Rosa, R., Brito, J.L.S., Ferreira, L.G., 2010. Land cover mapping of the tropical savanna region in Brazil. Environ. Monit. Assess. 166 (1–4), 113–124. https://doi.org/10.1007/s10661-009-0988-4.

Santoro, M., Cartus, O., Carvalhais, N., Rozendaal, D., Avitabile, V., Bruin, S. De, Herold, M., Quegan, S., Veiga, P.R., Balzter, H., 2020. Science 5174 (July), 1–38.

dos Santos, E.G., Shimabukuro, Y.E., de Moura, Y.M., Gonçalves, F.G., Jorge, A., Gasparini, K.A., Arai, E., Duarte, V., Ometto, J.P., 2019. Multi-scale approach to estimating aboveground biomass in the Brazilian Amazon using Landsat and LiDAR data. Int. J. Remote Sens. 40 (22), 8635–8645. https://doi.org/10.1080/2150704X.2019.1619955.

Schultz, M., Clevers, J.G.P.W., Carter, S., Verbesselt, J., Avitabile, V., Vu, H., Herold, M., 2016. Performance of vegetation indices from Landsat time series in deforestation monitoring. Int. J. Appl. Earth Observat. Geoinform. 52, 318–327. https://doi.org/10.1016/j.jag.2016.06.020.

Schwieder, M., Leitão, P.J., da Cunha Bustamante, M.M., Ferreira, L.G., Rabe, A., Hostert, P., 2016. Mapping Brazilian savanna vegetation gradients with Landsat time series. Int. J. Appl. Earth Obs. Geoinf. 52, 361–370. https://doi.org/10.1016/j.jag.2016.06.019.

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.

Scolforo, J.R., Rufini, A.L., Mello, J.M.de, Oliveira, A.D. de, Silva, C.P. de C., 2008. EQUAÇÕES PARA ESTIMAR O VOLUME DE MADEIRA DAS FISIONOMIAS, EM MINAS GERAIS. In: Inventario Florestal Minas Gerais (pp. 67–101).

SEEG Project – Greenhouse Gas Emission and Removal Estimating System, https://seeg.eco.br. Accessed 5 January 2021.

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., Lucas, R., 2014. New global forest/non-forest maps from ALOS PALSAR data (2007–2010). Remote Sens. Environ. 155, 13–31. https://doi.org/10.1016/j.rse.2014.04.014.

Silveira, E.M.O., Silva, S.H.G., Acerbi-Junior, F.W., Carvalho, M.C., Carvalho, L.M.T., Scolforo, J.R.S., Wulder, M.A., 2019. Object-based random forest modelling of aboveground forest biomass outperforms a pixel-based approach in a heterogeneous and mountain tropical environment. Int. J. Appl. Earth Obs. Geoinf. 78, 175–188. https://doi.org/10.1016/j.jag.2019.02.004.

Sinha, S., Jeganathan, C., Sharma, L.K., Nathawat, M.S., Das, A.K., Mohan, S., 2016. Developing synergy regression models with space-borne ALOS PALSAR and Landsat TM sensors for retrieving tropical forest biomass. J. Earth Syst. Sci. 125 (4), 725–735. https://doi.org/10.1007/s12040-016-0692-z.

Soares-Filho, B., Rajão, R., Macedo, M., Carneiro, A., Costa, W., Coe, M., Rodrigues, H., Alencar, A., 2014. Cracking Brazil's forest code. Science 344 (6182), 363–364. https://doi.org/10.1126/science.1246663.

Spera, S.A., Galford, G.L., Coe, M.T., Macedo, M.N., Mustard, J.F., 2016. Land-use change affects water recycling in Brazil's last agricultural frontier. Glob. Change Biol. 22 (10), 3405–3413. https://doi.org/10.1111/gcb.13298.

Strassburg, B.B.N., Brooks, T., Feltran-Barbieri, R., Iribarrem, A., Crouzeilles, R., Loyola, R., Latawiec, A.E., Oliveira Filho, F.J.B., Scaramuzza, C.A.de M., Scarano, F.R., Soares-Filho, B., Balmford, A., 2017. Moment of truth for the Cerrado hotspot. Nat. Ecol. Evolut. 1(4), 0099. https://doi.org/10.1038/s41559-017-0099.

Wu, W., 2014. The Generalized difference vegetation index (GDVI) for dryland characterization. Remote Sensing 6 (2), 1211–1233. https://doi.org/10.3390/rs6021211.

Zanne, A.E., Lopez-Gonzalez, G., Coomes, D.A., Ilic, J., Jansen, S., Lewis, S.L.., Miller, R. B., Swenson, N.G., Wiemann, M.C., Chave, J. & LopezGonzalez, G., 2009. Global Wood Density database. http://hdl.handle.net/ 10255/dryad.235. Accessed 6 October 2009.

Zhang, G., Lu, Y., 2012. Bias-corrected random forests in regression. J. Appl. Statist. 39 (1), 151–160. https://doi.org/10.1080/02664763.2011.578621.

Zimbres, B., Shimbo, J., Bustamante, M., Levick, S., Miranda, S., Roitman, I., Silvério, D., Gomes, L., Fagg, C., Alencar, A., 2020. Savanna vegetation structure in the Brazilian Cerrado allows for the accurate estimation of aboveground biomass using terrestrial laser scanning. For. Ecol. Manage. 458 https://doi.org/10.1016/j.foreco.2019.117798.