Remote Sensing of Environment 194 (2017) 16–32

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Mapping major land cover types and retrieving the age of secondary forests in the Brazilian Amazon by combining single-date optical and radar remote sensing data

João M.B. Carreiras [a,*], Joshua Jones [b], Richard M. Lucas [c], Yosio E. Shimabukuro [d]

[a] NERC National Centre for Earth Observation (NCEO), University of Sheffield, Hicks Building, Hounsfield Road, Sheffield, S3 7RH, United Kingdom
[b] Department of Geography and Earth Sciences, Aberystwyth University, Ceredigion, Aberystwyth, SY23 3DB, Wales, United Kingdom
[c] Centre for Ecosystem Science, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney 2052, NSW, Australia
[d] Remote Sensing Division, National Institute for Space Research, Avenida dos Astronautas 1758, São José dos Campos, 12227-010 São Paulo, Brazil

ARTICLE INFO

Article history:
Received 19 May 2016
Received in revised form 15 February 2017
Accepted 12 March 2017
Available online xxxx

Keywords:
Tropical secondary forests
Amazon
Age of secondary forests
Landsat TM
ALOS PALSAR
Random forests

ABSTRACT

Secondary forests play an important role in restoring carbon and biodiversity lost previously through deforestation and degradation and yet there is little information available on the extent of different successional stages. Such knowledge is particularly needed in tropical regions where past and current disturbance rates have been high but regeneration is rapid. Focusing on three areas in the Brazilian Amazon (Manaus, Santarém, Machadinho d'Oeste), this study aimed to evaluate the use of single-date Landsat Thematic Mapper (TM) and Advanced Land Observing Satellite (ALOS) Phased Arrayed L-band Synthetic Aperture Radar (PALSAR) data in the 2007–2010 period for i) discriminating mature forest, non-forest and secondary forest, and ii) retrieving the age of secondary forests (ASF), with 100 m × 100 m training areas obtained by the analysis of an extensive time-series of Landsat sensor data over the three sites. A machine learning algorithm (random forests) was used in combination with ALOS PALSAR backscatter intensity at HH and HV polarizations and Landsat 5 TM surface reflectance in the visible, near-infrared and shortwave infrared spectral regions. Overall accuracy when discriminating mature forest, non-forest and secondary forest is high (95–96%), with the highest errors in the secondary forest class (omission and commission errors in the range 4–6% and 12–20% respectively) because of misclassification as mature forest. Root mean square error (RMSE) and bias when retrieving ASF ranged between 4.3–4.7 years (relative RMSE = 25.5–32.0%) and 0.04–0.08 years respectively. On average, unbiased ASF estimates can be obtained using the method proposed here (Wilcoxon test, p-value > 0.05). However, the bias decomposition by 5-year interval ASF classes showed that most age estimates are biased, with consistent overestimation in secondary forests up to 10–15 years of age and underestimation in secondary forests of at least 20 years of age. Comparison with the classification results obtained from the analysis of extensive time-series of Landsat sensor data showed a good agreement, with Pearson's coefficient of correlation (R) of the proportion of mature forest, non-forest and secondary forest at 1-km grid cells ranging between 0.97–0.98, 0.96–0.98 and 0.84–0.90 in the 2007–2010 period, respectively. The agreement was lower (R = 0.82–0.85) when using the same dataset to compare the ability of ALOS PALSAR and Landsat 5 TM data to retrieve ASF. This was also dependent on the study area, especially when considering mapping secondary forest and retrieving ASF, with Manaus displaying better agreement when compared to the results at Santarém and Machadinho d'Oeste.

1. Introduction

Land use and land cover change, and particularly conversion from forest to non-forest (i.e., deforestation), is the second largest source of carbon dioxide emissions after fossil fuel combustion, accounting for 9% of annual emissions between 2004 and 2013 (Le Quéré et al., 2015). Deforestation across tropical and subtropical biomes was estimated to account for over 60% of total deforestation between 2000 and 2012 (Hansen et al., 2013), with almost two thirds in areas with high tree cover (>75%). This has severe consequences not only in terms of carbon stocks depletion (Harris et al., 2012) but also losses of biodiversity (Laurance et al., 2014; Lewis et al., 2015).

In the Amazon region, annual deforestation (clear cut) mapping primarily from Landsat data has been carried out by the Brazilian National Institute for Space Research (Instituto Nacional de Pesquisas

* Corresponding author.
E-mail address: j.carreiras@sheffield.ac.uk (J.M.B. Carreiras).

http://dx.doi.org/10.1016/j.rse.2017.03.016

Espaciais, INPE) since 1988 under the *Projeto de Monitoramento do Desmatamento na Amazônia Legal por Satélite* (PRODES) project (INPE, 2015). Reported deforestation rates have been highly variable over time, reaching their highest value in the mid-1990s (~ 30,000 km^2yr^{-1}), but have progressively decreased since the mid-2000s to a record low of ~4500 km^2yr^{-1} in 2012. Several socio-economic factors contributed to these observed changes in deforestation rates (Brondizio and Moran, 2012; Ewers et al., 2008). Large-scale deforestation, which started in the 1970s and continued throughout the 1980s and 1990s, was linked to policies aimed at encouraging colonisation of the region (e.g., fiscal incentives), expanding the road network and granting land titles to settlers (Fearnside, 2005). Since the late 1990s, however, and up until today, global demand for commodities (mainly soy and beef) started playing a stronger role in the temporal variations of annual deforestation rates. Rates of deforestation observed in the late 1990s and early 2000s were successfully reduced by a combination of stronger forest monitoring-based law enforcement, expansion of protected areas, and interventions at the supply chain level (Nepstad et al., 2014).

Deforestation typically results in the replacement of forests by croplands and pastures but these are often abandoned after a few years and replaced with secondary forests. These serve to accumulate carbon and restore biodiversity lost previously during the initial deforestation process (Brown and Lugo, 1990). The age, structure and species composition of secondary forests establishing on abandoned lands are a consequence of several factors, such as land use history, soil fertility and distance to mature primary forests (Chazdon, 2003). For this reason, knowledge of the age and land use history of areas under regeneration is needed to better understand patterns of carbon accumulation and recovery (or otherwise) of biodiversity. Mapping the age of tropical secondary forests often relies on comparing time-series of land cover maps (including a secondary forest class) obtained from classification of high-resolution optical data (Carreiras et al., 2014; Nelson et al., 2000; Prates-Clark et al., 2009). Other approaches use single-date remote sensing (often optical) data to map the age of secondary forests into classes (e.g., initial, intermediate and advanced secondary forests) (Lucas et al., 2000; Vieira et al., 2003). Recently, Chazdon et al. (2016) used an above-ground forest biomass map of the Neotropics (Baccini et al., 2012) in combination with ~1500 plots in secondary forests of known age to derive a large-scale map of the age of secondary forests. All these approaches rely on the availability of reference information about i) areas occupied by secondary forests (time-series approach) or ii) areas of known age or age class (single-date approach). The application of these methods is sometimes severely hampered by frequent cloud cover in tropical regions, thus leading to some regions having poor coverage by optical sensors. For this reason, all-weather Synthetic Aperture Radar (SAR) data are increasingly being used and promoted to improve land use/land cover change monitoring over tropical regions (Reiche et al., 2016). SAR data can provide information related to structural parameters of the forest (above-ground biomass, canopy height) (Carreiras et al., 2012; Cartus et al., 2012; Lucas et al., 2010; Santoro et al., 2011) that could prove useful when retrieving the age of secondary forests and complement that provided by optical sensors.

The objective of this study was to investigate the combined use of single-date optical (Landsat 5 Thematic Mapper) and Advanced Land Observing Satellite (ALOS) Phased Array L-band Synthetic Aperture Radar (PALSAR) to a) map major land cover types (mature forest, non-forest and secondary forest) and b) retrieve the age of secondary forests across three sites (~5000 km^2) in the Brazilian Amazon: Manaus (2.6° S, 60.2° W), Santarém (3.1° S, 54.8° W) and Machadinho d'Oeste (9.5° S, 62.4° W) (Fig. 1A). Since the inception of deforestation in the Amazon, these regions have experienced different but distinct patterns of land use and land cover change thus leading to present day landscapes characterised by a dissimilar spatial arrangement and frequency of land cover classes, namely secondary forests of varying age. By knowing, as a minimum, the age of these forests, a better estimation of the contribution of these landscapes to the regional carbon balance and biodiversity recovery can be achieved.

2. Study areas

The Manaus site, north of Manaus (Amazonas state) (Fig. 1B) encompasses the majority of a Federal conservation unit, the Biological Dynamics of Forest Fragments Project (BDFFP) (Laurance et al., 2011). Other county and state conservation units are also included in this study area. In the early 1970s, the construction of a highway connecting Manaus with Boa Vista (BR-174) was the key cause of deforestation in the region. Agricultural expansion was the main deforestation driver, which occurred across both sides of the highway. The BDFFP was established in 1985, but as far back as 1979, several forest fragments were preserved prior to deforestation of the surrounding forest. These were used to study the impacts of deforestation (and corresponding forest fragmentation) on ecosystem structure and function, thereby informing future conservation programmes in the Amazon (Laurance et al., 2011).

The Santarém site is located approximately 80 km to the south of Santarém (Pará state) (Fig. 1C). The study area is partially within a Federal conservation unit - Tapajós National Forest (FLONA Tapajós) - between the Tapajós River and the BR-163 highway connecting Santarém with Cuiabá (Mato Grosso). This unit was created in 1974 and has been used successfully to implement some novel forest management practices, such as the benefits of reduced impact logging on social welfare and biodiversity (Bacha and Rodriguez, 2007; van Gardingen et al., 2006).

The Machadinho d'Oeste site (Fig. 1D) is mainly located within the Machadinho d'Oeste municipality (Rondônia state). Its origins are a settlement project, initiated by the Brazilian Federal Government in 1982 with the support of the World Bank to colonize some regions of the Amazon (Miranda, 2009). The original vegetation is dominated by open rainforests (Miranda, 2009) and, according to Batistella and Moran (2005), most of its inhabitants depend on subsistence agriculture. This site includes several state-level conservation units, mainly extractive reserves, established in the mid-1990s.

The geographic location of these three sites has implications in terms of climate, although the mean annual temperatures are all around 25 °C to 26 °C (Bierregaard, 2001; Miranda, 2009; Silver et al., 2000). At Manaus, a relatively strong dry season occurs between June and October, with annual rainfall between 1900 and 3500 mm (Laurance et al., 2011). At Santarém, the dry season lasts from May to October and average annual rainfall is approximately 2000 mm (Silver et al., 2000). At Machadinho d'Oeste, the dry season occurs between April and November, with an annual rainfall around 2400 mm (Miranda, 2009). Forest types at the study areas are determined partially by topography and soils. Topography ranges from moderately flat (Manaus: up to 160 m elevation, Santarém: between 50 and 240 m) to moderately hilly (Machadinho d'Oeste: 90–370 m). Overall, soils are nutrient poor; ferralsols at Manaus (Laurance et al., 1999), ferralsols and nitosols in Santarém (Keller et al., 2005; Silver et al., 2000), and ferralsols, nitosols and fluvisols at Machadinho d'Oeste (Miranda, 2009).

3. Data

3.1. Time-series of land cover maps and age of secondary forests (ASF)

Existing time-series of land cover maps depicting three primary classes (mature forest, non-forest, secondary forest) were obtained through automatic classification of Landsat sensor data over the three selected sites. Carreiras et al. (2014) and Prates-Clark et al. (2009) provide detailed information about remote sensing data pre-processing and methods used to generate these land cover maps and the corresponding accuracy assessment. In this section, only a summary is presented.

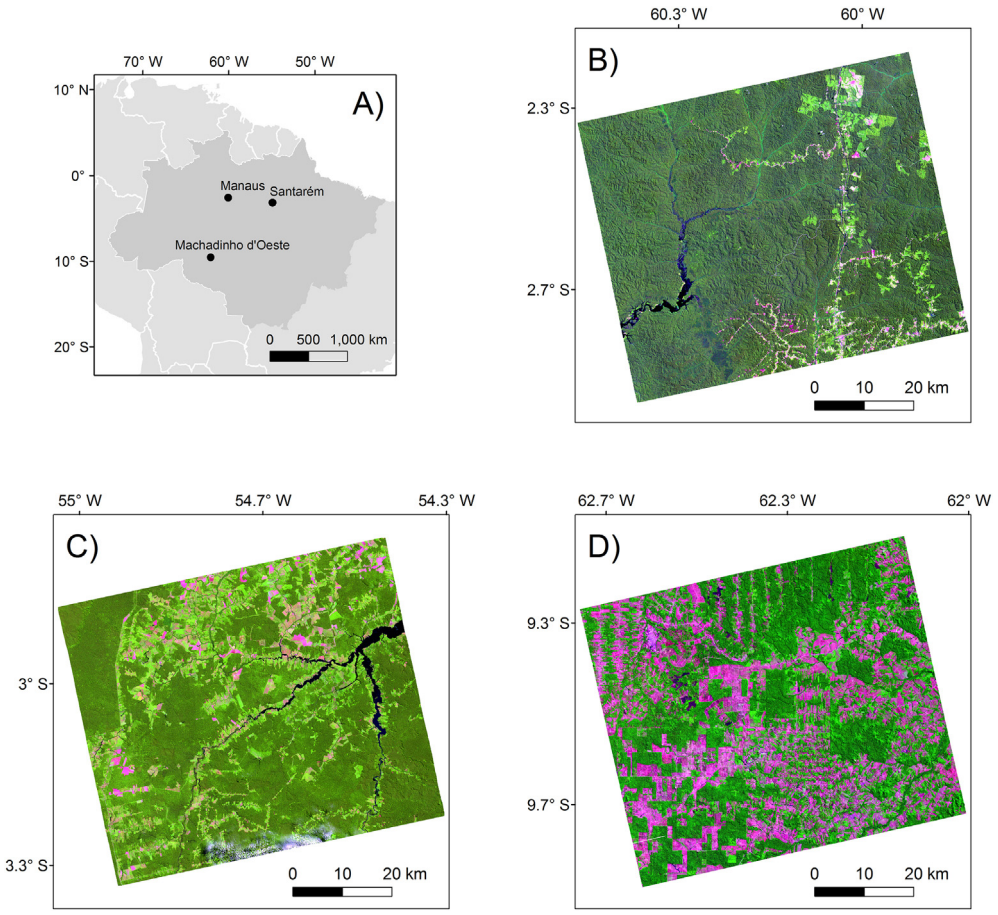


Fig. 1. A) Location of the three sites within the Brazilian Legal Amazon (dark grey). Landsat 5 Thematic Mapper (TM) color composite (RGB: TM band 5, TM band 4, TM band 3) acquired in 2010 over B) Manaus (27 July), C) Santarém (29 June) and D) Machadinho d'Oeste (27 July).

At Manaus, Prates-Clark et al. (2009) used Landsat Multi-spectral Scanner (MSS), Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) data acquired between 1973 and 2003 (path 231, row 62) to analyse the land cover dynamics in the region. Carreiras et al. (2014) extended the analysis up to 2011 using Landsat 5 TM data acquired between 2006 and 2011. At Santarém, mostly Landsat TM data acquired between 1984 and 2003 (path 227, row 62) were classified to generate the time-series of the 3-class land cover maps (Prates-Clark et al., 2009); again, Carreiras et al. (2014) extended the analysis up to 2010 with Landsat 5 TM data acquired in the 2005–2010 period (no cloud-free data were available for 2011). In Machadinho d'Oeste, Landsat 5 TM data from the 1984–2011 period (path 231, row 67) were used to create the time series of the land cover maps (Carreiras et al., 2014). At all sites, most scenes were unaffected by substantive cloud cover and overall gaps in the time-series ranged between one (70%) and 4 years (3%). A solution to solve some of the coverage gaps would have been to create cloud-free annual composites in those years where a single cloud-free scene was not available (e.g., Hermosilla et al. (2015)). However, the vast majority of scenes were acquired in consecutive years and we deemed this to be adequate to provide a good representation of the major land cover dynamics in these study areas.

Several classification algorithms were used to generate the best possible discrimination among mature forest, non-forest and secondary forest at all three sites. At Manaus, Prates-Clark et al. (2009) used both minimum distance and maximum likelihood supervised classification algorithms to generate the time-series of three-class land cover maps up to 2003. Subsequently, Carreiras et al. (2014) followed an object-oriented classification approach to produce the same land cover discrimination between 2006 and 2011. A fuzzy logic approach up to 2003 (Prates-Clark et al., 2009) and a random forests algorithm between 2006 and 2010 (Carreiras et al., 2014) were used to generate the equivalent land cover maps at Santarém. At Machadinho d'Oeste, all Landsat 5 TM data (1984–2011) were classified using a random forests algorithm (Carreiras et al., 2014).

A comparison of the time-series of the three-class land cover maps at each site was carried out on a pixel-by-pixel basis. For example, pixels that were classified as secondary forest or non-forest at a given date and as mature forest on the following date were identified and reassigned to the secondary forest class as these transitions were not considered possible (Carreiras et al., 2014). These pixels were not reclassified as non-forest because it was assumed that spectral confusion between this class and mature forest was unlikely.

Accuracy of the three-class land cover maps for each site was assessed using the most recent very high spatial resolution imagery available (Carreiras et al., 2014). For all areas, and for each class, 200 points were randomly selected from within the centre of contiguous areas greater than 6 ha (Carreiras et al., 2014). However, since this approach (Carreiras et al., 2014) did not take into account the areal extent of each class, we use the guidelines in Olofsson et al. (2014) to provide the area-corrected error matrices (Supplementary Information, Table S1). The overall accuracy was above 90% at all three sites. An omission error of 44% in the non-forest class at Manaus was the consequence of misclassification as mature forest. However, on average, higher omission errors were observed in the secondary forest class: 22%, 13% and 26% in Manaus, Santarém and Machadinho d'Oeste respectively. These were mainly the consequence of misclassification as mature forest in Manaus and Santarem, and as non-forest in Machadinho d'Oeste.

The time-series of land cover maps at each site was compared to generate information about the age of secondary forests (ASF),

estimated by summing the time (in years) that each pixel was occupied by secondary forests since the last clearance event. However, when secondary forests were mapped on the first date of the time-series, and not cleared subsequently, ASF can only be considered as a minimum age as the exact date of land abandonment is not known. Only the maps representing ASF between 2007 and 2010 were used; 2007 was the first year with an available ALOS PALSAR coverage and 2010 the last year of acquisitions of this sensor over the three sites.

3.2. Landsat 5 Thematic Mapper (TM) surface reflectance data

The original 30-m spatial resolution Landsat 5 TM data used to generate the land cover maps in Carreiras et al. (2014) at the three sites between 2007 and 2010 (Table 1) were downloaded from the United States Geological Survey (USGS) EarthExplorer data repository. As mentioned in Section 3.1, all scenes were unaffected by substantive cloud cover and although the surface reflectance product is still flagged as provisional, it was used to eliminate as much as possible the atmospheric contamination from each scene. However, after careful examination of all scenes, atmospheric contamination was evident in the 2010 scene acquired over Machadinho d'Oeste. Therefore, the original Landsat 5 TM acquisition used in Carreiras et al. (2014) (25 June 2010) was replaced by a scene acquired in 27 July 2010. Furthermore, a cloud-free Landsat 5 TM scene acquired over Santarém in 2008 was used, as it was closer in time to the ALOS PALSAR acquisition (23 June 2008, Table 1). Landsat 5 TM spectral bands most likely to be impacted by atmospheric effects (TM band 1, blue: 0.485 μm; TM band 2, green: 0.569 μm) were excluded from the analysis. Detailed information about the algorithm used to generate the surface reflectance product is given in Masek et al. (2006) and USGS (2015). The analysis was carried out with surface reflectance data acquired by TM bands 3 (TM3, red: 0.660 μm), 4 (TM4, near infrared: 0.840 μm), 5 (TM5, shortwave infrared: 1.676) and 7 (TM7, shortwave infrared: 2.222 μm). These spectral regions have different sensitivity to the leaf biophysical components within each resolution cell, with TM3 being sensitive to photosynthetic pigments (mainly chlorophyll), TM4 to cell leaf structure and TM5 and TM7 to leaf water content Nelson et al., 2000.

3.3. ALOS PALSAR dual-pol backscatter intensity

All available ALOS PALSAR dual-pol data (HH and HV polarization) over the three sites between 2007 and 2010 were retrieved in single look complex format. The number of scenes (dates) acquired at each site ranged from eight (Machadinho d'Oeste) to ten (Santarém), with nine dates available over Manaus in the same period. Each ALOS PALSAR dual-pol scene covers an area of approximately 70 km × 70 km, which is less than a quarter of the area covered by a Landsat 5 TM scene. Therefore, only the ALOS PALSAR dual-pol scene covering most of the area at each site in Carreiras et al. (2014) was selected and used throughout the analysis.

All ALOS PALSAR dual-pol data were acquired with an off-nadir angle of 34.3° during ascending orbits, and processed with SARScape (version 4.3.001, Sarmap SA, http://www.sarmap.ch), which followed standard SAR processing (Oliver and Quegan, 1998; Ulaby and Dobson, 1989; Woodhouse, 2006). First, single look complex data were converted to multi-look intensity format, using a multi-look factor of 1 in range and 5 in azimuth to obtain approximately 15-m square pixels in ground range coordinates. The multi-looked data were then transformed into geocoded terrain-corrected data (i.e., transformed from slant-range/azimuth to map projection geometry using the 90-m digital elevation model (DEM) retrieved from the Shuttle Radar Topography Mission, SRTM). Meier et al. (1993) provide a comprehensive description of the geocoded terrain correction method implemented in SARScape.

Radiometric calibration was carried out by following the radar equation and associated corrections for scattering area, antenna gain pattern and range spread loss (Holecz et al., 1993; Holecz et al., 1994). The calibrated value is a normalized dimensionless number (linear units, $m^2 \cdot m^{-2}$), and the corresponding value of gamma nought (γ°) in the dB scale calculated as $10\log_{10}$ of the linear value. Even after rigorous radiometric calibration, backscatter intensity (γ°) variations are clearly identifiable in the range direction and in the presence of topography, thus requiring radiometric normalization. In this study, a cosine correction method was applied to the backscatter intensity (γ°) values to compensate for range variations (Ulaby and Dobson, 1989). Although SAR processing included a multi-look step (5-looks: 1 in range, 5 in azimuth), a visual inspection still revealed a high degree of speckle. Therefore, a multitemporal speckle filter (7 × 7) (Quegan and Yu, 2001) was applied to each time-series (one per site) of ALOS PALSAR dual-pol terrain-corrected data, thus generating a reduced-speckle dataset without significant loss of spatial resolution. In case a multi-temporal SAR dataset was not available, several filtering methods are available to reduce speckle (Lee, 1980; Lee et al., 2009; Touzi, 2002). However, these often have an impact in terms of reducing the spatial resolution of the filtered images (Quegan and Yu, 2001).

Only those scenes acquired close to the Landsat TM acquisitions (to generate the corresponding land cover map) were used in the subsequent analysis (Table 1). All ALOS PALSAR - Landsat 5 TM pairs were acquired within two months of each other and during the dry season, with a view to minimizing the impact of environmental conditions, especially the effects of rainfall events on SAR backscatter intensity.

Table 1
ALOS PALSAR dual-pol (HH + HV polarisations) and Landsat 5 TM data used to discriminate the age of secondary forests (ASF). Date format is yyyymmdd.

	Acquisition date	
Site	ALOS PALSAR	Landsat 5 TM
Manaus	Path: 76	Path: 231
	Frame: 7130	Row: 62
	20070715	20070804
	20080717	20080806
	20090904	20090910
	20100607	20100727
Santarém	Path: 66	Path: 227
	Frame: 7120	Row: 62
	20070613	20070621
	20080430	20080623
	20090803	20090712
	20100621	20100629
Machadinho d'Oeste	Path: 83	Path: 231
	Frame: 6990	Row: 67
	20070811	20070703
	20080813	20080806
	20090701	20090809
	20100704	20100727

4. Methods

4.1. Sampling design

The ability of remote sensing data to discriminate the age of secondary forests (ASF) can only be evaluated where a set of spatially representative areas covered with secondary forests of known age are available. Random sampling was used to generate a spatially representative set of areas covered with secondary forests of known age. First, only areas in Carreiras et al. (2014) classified as secondary forests between 2007 and 2010 were used (e.g., areas mapped as secondary forests in 2007 and converted to non-forest between 2008 and 2010 were discarded). To reduce the effect of unwanted positional errors between the Landsat-derived land cover maps and ALOS PALSAR data, only the 100 largest patches of secondary forests of known age were selected for sampling at each site. The area of these patches ranged from 25 to 963 ha, 21–380 ha and 12–157 ha at Manaus, Santarém and Machadinho d'Oeste respectively. At each patch, a maximum of 10

points was randomly selected subject to the conditions of being i) separated by at least 200 m and ii) located 75 m away from the patch boundary to avoid edge effects. A 100 m × 100 m plot (sampling unit, area = 1 ha) centred at each point was then generated and used to retrieve the average i) ASF, ii) Landsat 5 TM surface reflectance at TM3, TM4, TM5 and TM7 spectral bands, and iii) ALOS PALSAR backscatter intensity at HH and HV polarizations. We chose a sampling unit size of 100 m × 100 m to avoid unwanted positional errors due to an eventual spatial mismatch between ALOS PALSAR and Landsat 5 TM data.

Due to secondary forest patch size and shape differences among sites and restrictions in terms of distance between points and to the patch boundary, the number of sampling units per site was different (Manaus: 205; Santarém: 105; Machadinho d'Oeste: 159; Table 2). At each site, sampling units over mature forest and non-forest persisting in the 2007–2010 period were also randomly selected and followed the same approach described for secondary forests. This was undertaken to further evaluate the capability of combining Landsat 5 TM and ALOS PALSAR data to discriminate amongst mature forest, non-forest and secondary forest. The number of randomly selected sampling units over mature forest, non-forest and secondary forest at each site by year is depicted in Table 2, with the number of samples in secondary forest given in 5-year interval classes. The distribution of sampling units by ASF classes reflects the spatial patterns and dynamics of forest disturbance and recovery at each site, with Manaus clearly dominated by older secondary forests and Santarém and Machadinho d'Oeste by earlier stages of secondary forests (Carreiras et al., 2014). The collection of the training dataset relies on access to extensive time-series of land cover maps over representative areas undergoing a process of forest regeneration. Alternatively, this information can be collected from field campaigns where the age of secondary forests was estimated (e.g., social surveys) (Evans and Moran, 2002).

4.2. Analysis

4.2.1. Mapping mature forest, non-forest and secondary forest

In a first step, we used the training data given in Table 2 to test the ability of combining single-date ALOS PALSAR and Landsat 5 TM data to discriminate major land cover types across the three selected sites by year. For that purpose, we used a well-known non-parametric algorithm: random forests (Breiman, 2001). Non-parametric classification algorithms have been successfully used to discriminate land cover types and retrieve forest biophysical parameters at various spatial scales (Baccini et al., 2012; Carreiras et al., 2006; Carreiras et al., 2012; Friedl et al., 2010). Breiman (2001) random forests algorithm (RF hereafter) is among those most frequently used due to its simplicity and accuracy and can be used to perform either classification (Cutler et al., 2007) or regression (Baccini et al., 2012) analysis. RF is built as an ensemble of binary decision trees, with only two parameters needing to be tuned, i) the number of trees in each RF model (*ntree*) and ii) the number of randomly selected predictors to be used at each decision node (*mtry*) (Breiman, 2001). Each RF tree is fitted to a bootstrap sample of the original training dataset with replacement and grown to full extent. The observations not selected for fitting a RF tree (the out-of-bag sample) are then used to assess its error rate (in classification) or mean error (in regression) and, in the end, combined to give the overall out-of-bag estimate of error rate or mean error. The final prediction is generated as a majority vote (in classification) or as an average (in regression) from prediction over all RF trees. All binary decision trees in a RF are built using independent samples, so this method can also generate an estimate of prediction variability. In addition, an important feature of RF is the ability to generate a measure of the relative influence of each predictor; this is achieved by evaluating how much the out-of-bag error changes when the out-of-bag data for a given variable is permuted while the others are left unchanged (Breiman, 2001).

In this study, a sensitivity analysis was carried out to select the combination of *ntree* and *mtry* producing the lowest out-of-bag error in the class displaying the highest error rate. The functions implemented in R Model.Map package (Freeman and Frescino, 2009) were used to build the R code required to carry out the analysis, conducted on a yearly basis between 2007 and 2010. Classification results are reported using the out-of-bag sample to generate the corresponding error matrix, thus making it possible to identify those classes displaying higher omission and commission errors and if classification errors are impacted by the year used to run the RF model. As the number of training samples collected at each class was not proportional to the total area of that class, the traditional error matrix relying on sample counts was corrected using the mapped area of each class, with the proportion of the area in the map that was predicted as *i* and observed as *j* (p_{ij}) calculated with Eq. (1) (Olofsson et al., 2014).

$$p_{ij} = A_i \frac{n_{ij}}{n_i} \quad (1)$$

where A_i is the proportion of area predicted (mapped) as class *i*, n_{ij} the number of samples predicted as class *i* and observed as class *j*, and n_i, the total number of samples predicted as class *i*. Additionally, Olofsson et al. (2014) provide an estimator to calculate the proportion of area of a given class, which is the sum of proportions (p_{ij}) of each class as determined by the reference (observed) dataset. The selected RF algorithm was then applied to the overlapping area covered by the ALOS PALSAR and Landsat 5 TM data acquired at each site (Table 1, Fig. 1). Before applying the selected RF algorithm, these datasets were spatially averaged to the same resolution used to generate the training dataset (1 ha = 100 m × 100 m plots).

Table 2
Number of 1 ha (100 × 100 m) sampling units by land cover type and year at each study area; the number of sampling units over secondary forest is shown in 5-year interval classes.

Site	Year	Number of sampling units						
		Mature forest	Non-forest	Secondary forest (yr)				
				[0,5]	[5,10]	[10,15]	[15,20]	>20	Total
Manaus	2007	202	255	0	0	4	38	163	205
	2008			0	0	4	32	169	
	2009			0	0	4	9	192	
	2010			0	0	4	9	192	
Santarém	2007	155	322	51	36	3	5	10	105
	2008			37	39	12	3	14	
	2009			37	14	37	3	14	
	2010			28	23	36	3	15	
Machadinho d'Oeste	2007	169	834	58	47	34	12	8	159
	2008			47	54	38	12	8	
	2009			47	47	44	11	10	
	2010			2	92	23	32	10	

4.2.2. Discriminating and mapping the age of secondary forests (ASF)

The existence of a significant relationship (either positive or negative) between the age of secondary forests (ASF) and each individual variable (ALOS PALSAR dual-pol backscatter intensity and Landsat 5 TM surface reflectance) was assessed by fitting linear models using generalised least squares, allowing for errors to be either correlated or have unequal variances (Venables and Ripley, 1999). Eq. (2) depicts the linear model when assessing the relationship between ASF and ALOS PALSAR HH and HV backscatter intensity (x_1), with Eq. (3) giving the linear equation when evaluating the relationship between ASF and Landsat 5 TM surface reflectance at TM3, TM4, TM5 and TM7 bands (x_2). These relationships were fitted on a yearly basis, thus allowing us to assess their significance and magnitude between 2007 and 2010. The Akaike Information Criterion (AIC) (Hastie et al., 2009) and the estimate of the slope (a_1) of these fitted models are analysed to assess the impact of each individual variable in terms of retrieving ASF.

$$\ln(\mathrm{ASF}) = a_0 + a_1 x_1, \tag{2}$$

$$\ln(\mathrm{ASF}) = a_0 + a_1 \ln x_2 \tag{3}$$

Carreiras et al. (2014) mapped the extent of ASF by analysing the transitions obtained from time-series of three-class (mature forest, non-forest and secondary forest) land cover maps spanning almost three decades over three selected regions. Maps of ASF obtained from this extensive analysis rely on extensive remote sensing datasets, require high computing capabilities to upscale to larger scales, and can be time consuming. Also, they must rely on post-classification methods to identify and correct for unwanted transitions between consecutive dates, with this being done in the time-series analysis given in Carreiras et al. (2014). Furthermore, frequent cloud cover in tropical regions can introduce large temporal gaps in the time-series analysis, with this possibly impacting the accurate retrieval of the age of secondary forests, especially in highly dynamic regions in terms of forest processes (i.e., deforestation, regeneration). Therefore, we propose an approach based on the compilation of an extensive training dataset relating ASF to variables extracted from single-date remote sensing data (Table 1). This dataset includes information collected over a large ASF range (Table 2) and was tested over a period of 4 years (2007–2010). Furthermore, the interannual variation of each variable (from Landsat 5 TM and ALOS PALSAR data) by ASF class was evaluated using the Anderson-Darling k-sample test (implemented in R with the kSamples package) for the null hypothesis that all observations came from a common population (Scholz and Stephens, 1987). Only the areas identified as secondary forest in Section 4.2.1 will be further refined in terms of mapping their age.

The same non-parametric algorithm used to discriminate and map mature forest, non-forest and secondary forest (see Section 4.2.1) was used in regression mode to estimate ASF as a function of single-date ALOS PALSAR and Landsat 5 TM data. Regression results are reported in terms of the out-of-bag root mean squared error (RMSE) and bias (Hastie et al., 2009). Furthermore, a detailed analysis of the RMSE and bias by ASF class and year was also carried out, with the contribution of bias to the RMSE calculated with Eq. (4) (Hastie et al., 2009):

$$\mathrm{RMSE} = \sqrt{\mathrm{bias}^2 + \mathrm{variance}}. \tag{4}$$

Additionally, we tested for the null hypothesis of bias equal to zero using the one-sample Wilcoxon test (Hollander and Wolfe, 1999). Scatterplots of the relationship between observed and (out-of-bag) predicted ASF are used to characterize the accuracy of the retrieval method. An asymptotic regression model (Eq. (5)) (Pinheiro and Bates, 2000) was used to analytically estimate the asymptote (saturation) associated with this relationship:

$$ASF_{pre} = b0 + b1\left(1 - e^{\left(-e^{b2} ASF_{obs}\right)}\right), \tag{5}$$

where ASF_{pre} and ASF_{obs} are the predicted and observed ASF values respectively, $b0$ the estimated intercept on the y-axis, $b1$ the estimated difference between the asymptote and the y-intercept, and $b2$ the estimated logarithm of the rate constant. The *NLSstAsymptotic* command in the R *stats* package was used for this purpose.

4.2.3. Comparison with existing datasets

The results from the single-date analysis depicted here (100 m spatial resolution) were compared with those obtained from Carreiras et al. (2014) using a time-series analysis of Landsat TM data at 30 m spatial resolution. Several methods can be employed when comparing two (or more) maps either depicting land cover information or a given biophysical parameter. In this study, we used a systematic grid of 1 km × 1 km covering the entire overlapping area between the time-series analysis (Carreiras et al., 2014) and the output from this study. A similar approach was used in several studies comparing multiple land cover maps or maps estimating a given biophysical variable (Carreiras et al., 2012; Hansen and Reed, 2000; Mayaux et al., 2004; Pisek and Chen, 2007).

The proportion of each class and the mean ASF inside each 1 km × 1 km grid cells was estimated from the two datasets (time-series and this study) and the Pearson's coefficient of correlation computed to assess the level of agreement between them. Only the 1 km × 1 km grid cells with at least 25% of pixels with ASF values (from this study and Carreiras et al. (2014)) were used when comparing ASF maps.

Table 3
Error matrices by year, depicting the classification results of using a random forests (RF) algorithm to discriminate mature forest (MF), non-forest (NF) and secondary forest (SF) with single-date ALOS PALSAR and Landsat 5 TM data. OA – overall accuracy, OE – omission error, CE – commission error. The reported values refer to the out-of-bag sample and are given in terms of proportion, following the guidelines provided in Olofsson et al. (2014), and using the combined mapped area of each class in the three study areas.

			Observed			Total	CE (%)
			MF	NF	SF		
2007							
OA = 96%	Predicted	MF	0.5774	0.0011	0.0271	0.6056	4.66
		NF	0.0000	0.1827	0.0006	0.1834	0.35
		SF	0.0064	0.0023	0.2023	0.2110	4.14
		Total	0.5838	0.1861	0.2300	1.0000	
		OE (%)	1.10	1.84	12.05		
2008							
OA = 95%	Predicted	MF	0.5551	0.0011	0.0382	0.5945	6.62
		NF	0.0000	0.2128	0.0012	0.2140	0.56
		SF	0.0077	0.0017	0.1822	0.1916	4.91
		Total	0.5628	0.2156	0.2216	1.0000	
		OE (%)	1.37	1.30	17.80		
2009							
OA = 96%	Predicted	MF	0.5653	0.0000	0.0334	0.5987	5.58
		NF	0.0000	0.2034	0.0003	0.2037	0.14
		SF	0.0078	0.0013	0.1885	0.1976	4.58
		Total	0.5731	0.2047	0.2222	1.0000	
		OE (%)	1.36	0.63	15.15		
2010							
OA = 95%	Predicted	MF	0.5412	0.0000	0.0412	0.5830	7.17
		NF	0.0000	0.2299	0.0008	0.2307	0.36
		SF	0.0086	0.0033	0.1744	0.1863	6.39
		Total	0.5498	0.2332	0.2170	1.0000	
		OE (%)	1.57	1.41	19.64		

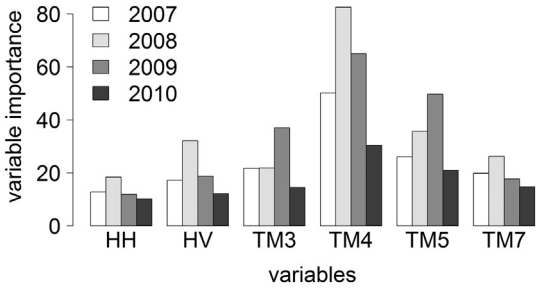


Fig. 2. Variable importance score by year (2007–2010) of the fitted random forests models using ALOS PALSAR HH and HV backscatter intensity and Landsat 5 TM surface reflectance (TM3, TM4, TM5 and TM7 bands) to discriminate mature forest, non-forest and secondary forest.

5. Results

5.1. Mapping mature forest, non-forest and secondary forest

Table 3 gives the (out-of-bag) area-corrected error matrices by year when using a random forests (RF) algorithm to discriminate mature forest, non-forest and secondary forest. The overall accuracy is high across all years (95–96%) with omission errors in the secondary forests class ranging between 12 and 20% because of misclassification as mature forest. Mature forest and non-forest have classification errors consistently below 10% and 2% respectively.

The variable importance score by year is given in Fig. 2. Landsat 5 TM surface reflectance in the near infrared spectral region (TM4) contributes the most to discriminating mature forest, non-forest and secondary forest, whereas surface reflectance in the shortwave infrared (TM5) ranks second in all years. According to this metric, variables obtained from ALOS PALSAR have lower discrimination capability when compared to those obtained from Landsat 5 TM.

An example of the application of the selected RF algorithm over the three study sites in 2010 is depicted in Fig. 3. These maps cover the overlapping area between the selected ALOS PALSAR and Landsat 5 TM scenes (Table 1 and Fig. 1). Overall areal estimates of each class at each site were obtained by using the information contained in the error matrix (Table 3) to correct the estimates obtained by pixel counting (Olofsson et al., 2014). Most of the deforested areas in Manaus were covered by secondary forest in 2010 (83%), with these representing only a small fraction of the study area, which continues to be dominated by mature forest (85%). At Santarém, 64% of the

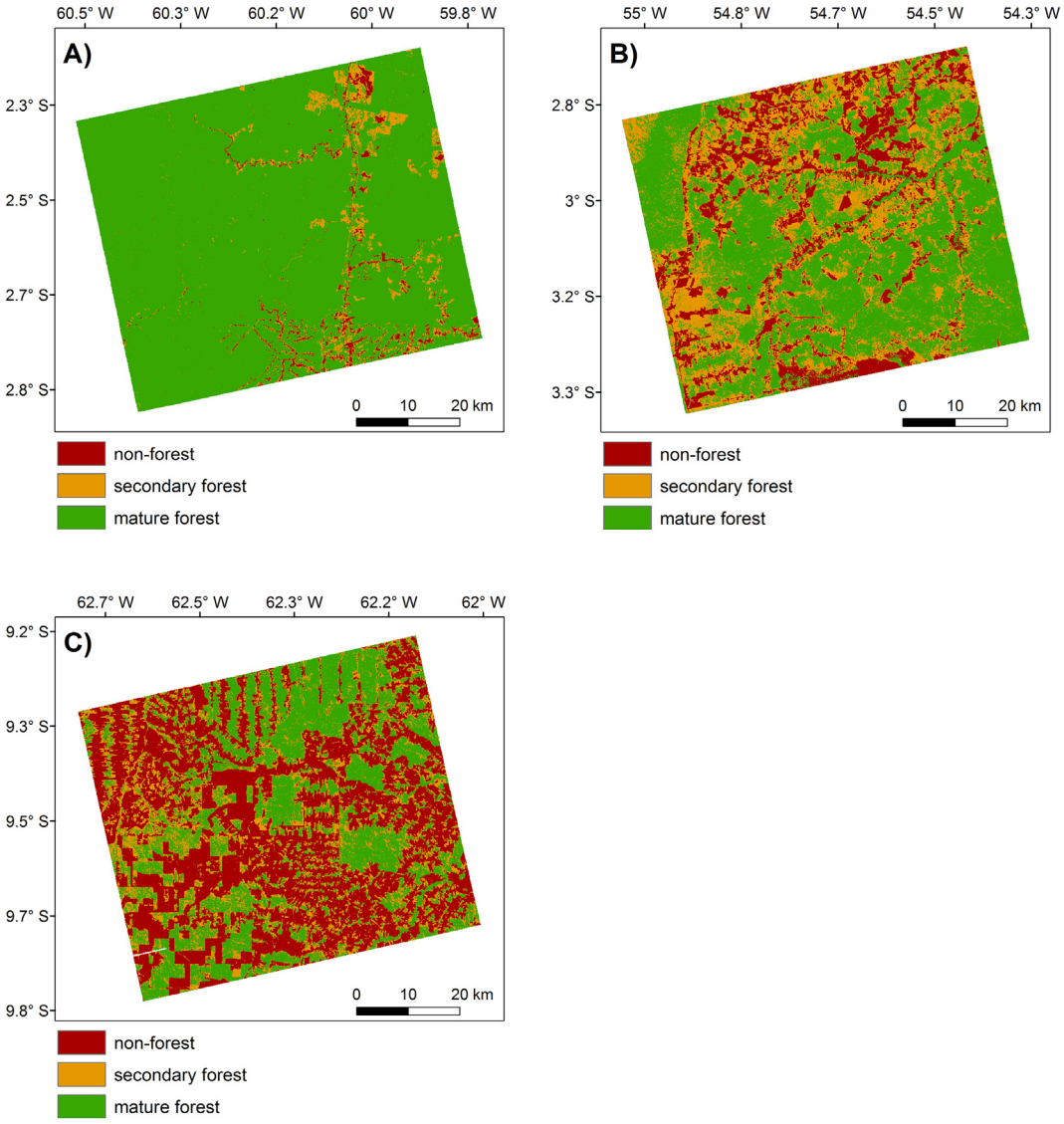


Fig. 3. Random forests (RF) based maps of non-forest, secondary forest and mature forest in 2010 over A) Manaus, B) Santarém, and C) Machadinho d'Oeste. The mapped regions correspond to the overlapping area between Landsat 5 TM and ALOS PALSAR data at each site (Fig. 1 and Table 1).

Table 4
Akaike Information Criterion (AIC) and a1 (slope, from Eqs. 1 and 2) estimate of linear models (fitted with generalised linear squares) between the age of secondary forest (ASF) and individual variables obtained from remote sensing data, by year; the estimates of the a1 parameter marked in bold are not significantly different from zero for an alpha of 5%.

Independent variable	2007		2008		2009		2010	
	AIC	a_1	AIC	a_1	AIC	a_1	AIC	a_1
ALOS PALSAR HH	1119.2	0.750	982.6	0.631	905.6	0.553	791.5	0.549
ALOS PALSAR HV	978.0	0.629	910.8	0.628	802.4	0.553	674.1	0.560
Landsat 5 TM3	1194.6	−3.812	1117.8	−1.649	1055.1	**0.226**	953.4	**−0.386**
Landsat 5 TM4	1295.4	−3.287	1126.5	−2.924	990.5	−2.943	825.1	−4.097
Landsat 5 TM5	815.7	−7.187	825.1	−5.764	783.6	−5.747	612.0	−5.817
Landsat 5 TM7	806.5	−4.602	917.0	−3.397	848.2	−3.718	610.5	−3.722

deforested area up to 2010 was covered by secondary forest, with mature forest covering approximately half (48%) of this site. At Machadinho d'Oeste, 49% of the site was occupied by non-forest in 2010, with less than one third (30%) of the deforested area supporting secondary forest; at this site, mature forest accounted only for 30% of the area.

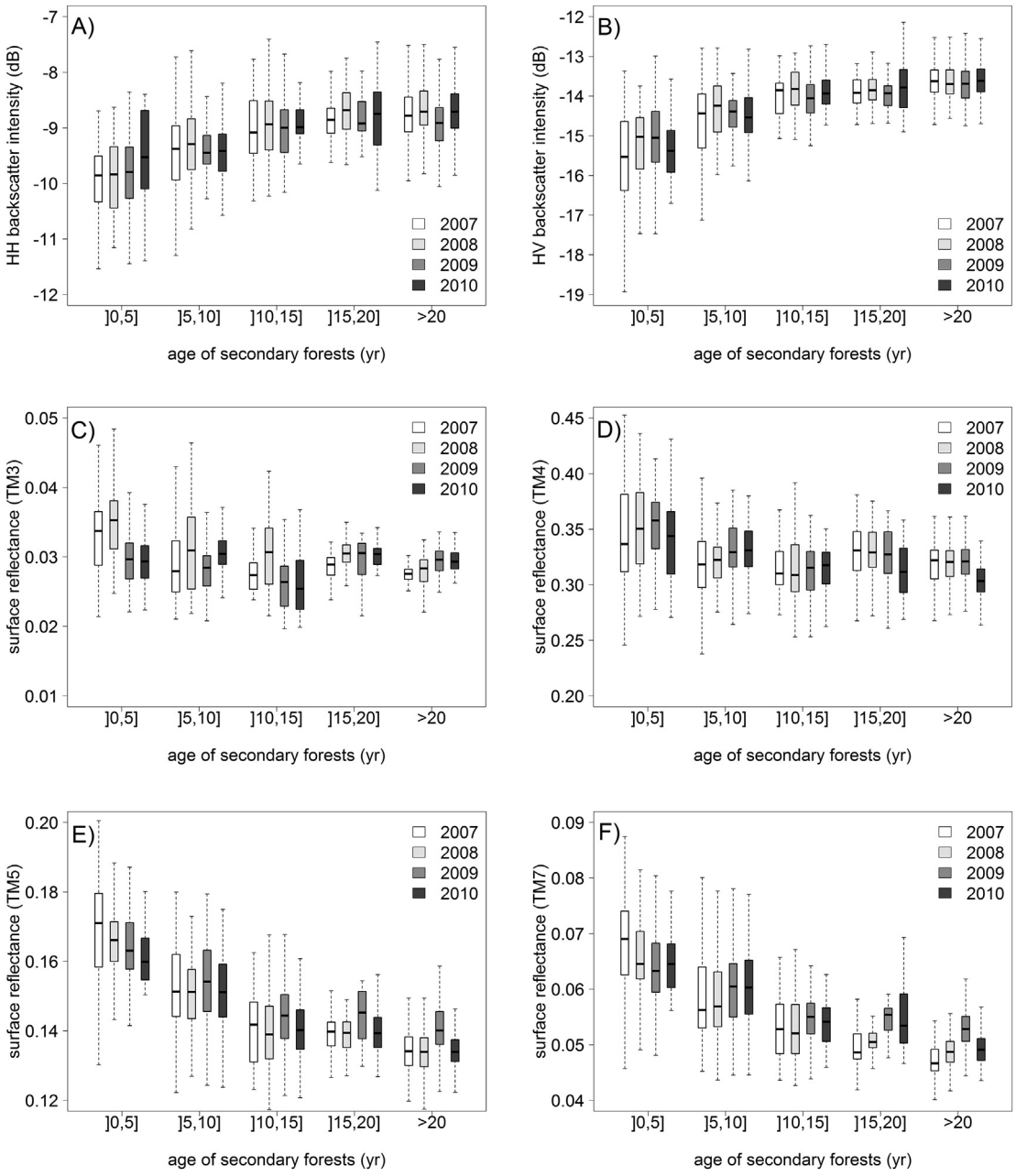


Fig. 4. Distribution of the values of variables obtained from ALOS PALSAR and Landsat 5 TM by age of secondary forests and year: A) ALOS PALSAR HH, B) ALOS PALSAR HV, C) Landsat 5 TM3, D) Landsat 5 TM4, E) Landsat 5 TM5, F) Landsat 5 TM7; data collected over the three Amazonian sites in secondary forest ranging from 1 to 27 years; a box-and-whisker plot displaying the minimum, first quartile, median, third quartile and maximum value of each variable (y-axis) is shown by age of secondary forests (x-axis).

Table 5
Anderson-Darling *k*-sample test for the null hypothesis that values from annual variables in the 2007–2010 period aggregated by 5-year interval age of secondary forest (ASF) were drawn from the same population.

Variables	ASF class (yr)				
	[0,5] p-Value	[5,10]	[10,15]	[15,20]	>20
ALOS PALSAR HH	0.024	0.045	0.155	0.100	<0.001
ALOS PALSAR HV	0.031	0.007	0.104	0.289	0.202
Landsat 5 TM3	<0.001	<0.001	<0.001	<0.001	<0.001
Landsat 5 TM4	0.020	0.001	0.412	0.026	<0.001
Landsat 5 TM5	<0.001	0.316	0.018	0.001	<0.001
Landsat 5 TM7	<0.001	0.043	0.009	<0.001	<0.001

5.2. Discriminating and mapping the age of secondary forests (ASF)

The values of the Akaike Information Criterion (AIC) and the estimate of the slope of the fitted models from Eq. 2 and 3 relating the age of secondary forests (ASF) against the individual variables obtained from ALOS PALSAR and Landsat TM data are given in Table 4. Across all years, the best quality models (as measured by AIC) were those relating ASF to Landsat 5 TM surface reflectance at the shortwave infrared spectral regions (TM5 and TM7). The relationship between ASF and ALOS PALSAR backscatter intensity at HV polarization also performed well. Landsat 5 TM surface reflectance at the red spectral region (TM3) was a poor predictor of ASF; this relationship was not significant in 2009 and 2010. An increase in ASF resulted always in increased ALOS PALSAR HH and HV backscatter intensity and decreased Landsat 5 TM surface reflectance at the red, near infrared and shortwave infrared spectral regions.

The distribution of ALOS PALSAR dual-pol backscatter intensity and Landsat 5 TM surface reflectance values by ASF and year is given in Fig. 4. In terms of ALOS PALSAR dual-pol HH and HV backscatter intensity, the relationship is markedly logarithmic (Fig. 4A and Fig. 4B), levelling-off (saturating) around 15–20 years of age for median HH and HV values of [−8.92, −8.68] dB and [−13.93, −13.85] dB, respectively, in the 2007–2010 period. Many studies assumed this type of relationship when modelling forest biophysical parameters as a function of SAR data (Carreiras et al., 2012; Lucas et al., 2006; Mitchard et al., 2009). Data acquired at HV polarisation yielded better quality models when compared to that at HH polarisation, with the former displaying lower AIC values (Table 4). A decreasing relationship between Landsat 5 TM surface reflectance and ASF was observed with all spectral bands (the exception was the non-significant relationship with TM3 in 2009 and 2010). This relationship was stronger in the shortwave infrared spectral region (TM5 and TM7) (Fig. 4E and Fig. 4F). However, the lack of sensitivity was evident around 15–20 years of age. Median values of Landsat 5 TM surface reflectance at the shortwave infrared were 0.139–0.143 and 0.049–0.055 in the 2007–2010 period for TM5 and TM7 respectively. Furthermore, a higher dispersion of ALOS PALSAR HH backscatter intensity values by ASF class is observed when compared to HV. The latter shows decreased variability with increasing ASF, especially in secondary forest of at least 10 years of age. This was confirmed with the Anderson-Darling *k*-sample test (Table 5). At the

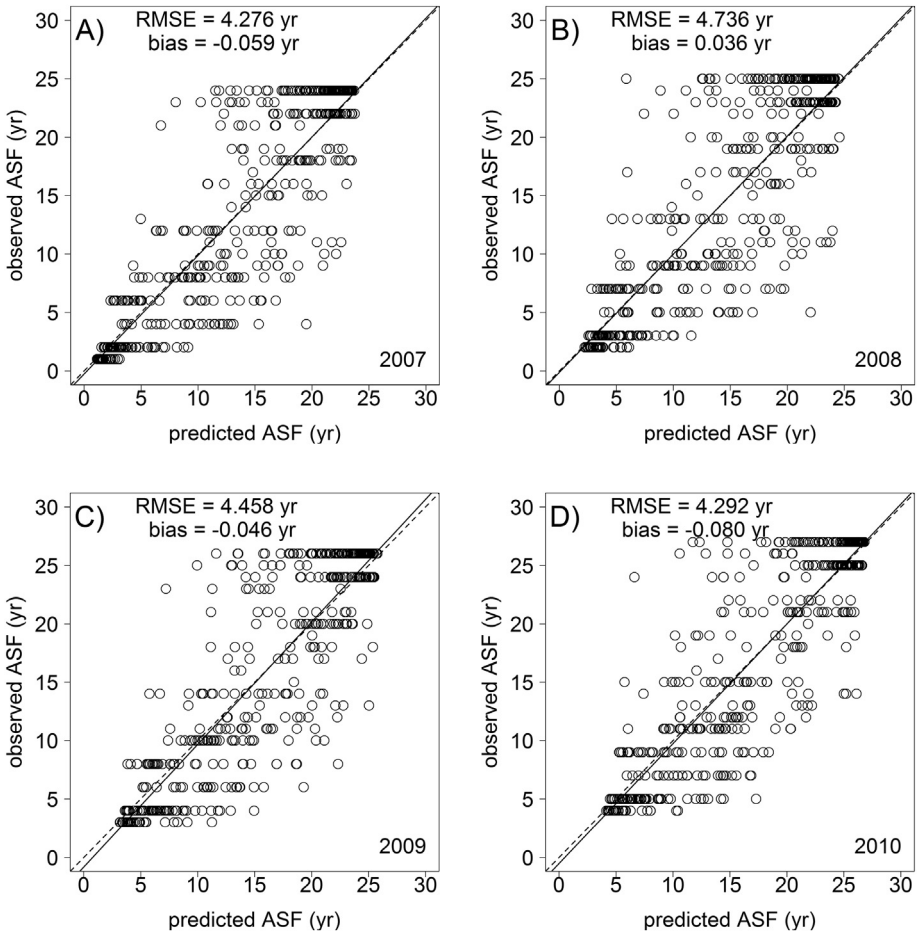


Fig. 5. Relationship between observed and predicted age of secondary forests (ASF) resulting from fitting random forests (RF) models with data from A) 2007, B) 2008, C) 2009 and D) 2010. Root mean square error (RMSE) and bias is indicated by year, with the solid line representing the fitted linear model between predicted and observed ASF and the dashed line the perfect agreement.

HV polarization this test showed that all annual samples came from the same population when considering secondary forests older than 10 years. The same was true at HH polarization but only for ASF between 10 and 20 years. In the optical region there is divergence across years, with the exception of Landsat 5 TM surface reflectance at the near infrared (band 4) in the [10,15] yr ASF class and shortwave infrared (band 5) in the [5,10] yr ASF class.

Out-of-bag validation of the fitted random forests (RF) models by year is given in Fig. 5. Root mean square error (RMSE) ranged between 4.3 and 4.7 years, corresponding to 25.5–32.0% of the mean observed ASF. Bias, which in absolute value ranged from 0.04 to 0.08 years, was not significantly different from zero for an alpha of 5% (non-parametric Wilcoxon test). This indicated that on average most of the error was a consequence of higher variance around the ASF estimates. The coefficient of determination (R^2) between the observed and predicted ASF ranged from 0.71 (2008) to 0.76 (2007 and 2010), with all regressions significant for an alpha of 5%. However, some degree of overestimation at lower ASF levels and underestimation at higher ASF values is observed, with the method showing lack of sensitivity with increasingly higher ASF values. The estimated value of the asymptote ($b0 + b1$, Eq. (5)) ranged between 20.0 years (2007) and 24.5 years (2010), which agrees with the values suggested from the observation of the distribution of individual variables (from ALOS PALSAR dual-pol and Landsat 5 TM data) as a function of ASF (Fig. 4). ASF estimates in secondary forests younger than 10 years were biased towards higher values and with a RMSE of 3–5 years (Table 6). On the contrary, they were biased towards lower values in secondary forests older than 20 years. Unbiased ASF estimates are generally obtained in the intermediate age classes (10–20 years).

As in Fig. 2, Fig. 6 depicts the variable importance score from the RF regression models by year. There is not a single variable contributing the most for retrieving ASF by year. However, Landsat 5 TM surface reflectance at the red and shortwave infrared spectral regions (TM3, TM5 and TM7) and ALOS PALSAR backscatter intensity at HV polarization were the variables with higher importance scores in the 2007–2010 period.

Fig. 7 displays the maps showing the estimated ASF by study area in 2010 when applied to the same region depicted in Fig. 3. The map of the coefficient of variation around the ASF estimates obtained from the RF models is also indicated. At Manaus, 59% of the area mapped as secondary forests supported forests with at least 20 years of age, with a significant proportion (67%) of the ASF estimates being obtained with a coefficient of variation lower than 30%. The mapped secondary forests at Santarém and Machadinho d'Oeste were younger than at Manaus, with 94% and 84% of these forests being younger than 20 years, respectively; 78–85% of the ASF estimates at these two sites were obtained with a coefficient of variation between 20 and 50%.

Table 6
Mean age of secondary forests (ASF, yr), root mean square error (RMSE; yr and %), bias (yr and %) and p-value of the one-sample Wilcoxon test (H_0: bias = 0) by ASF class and year.

ASF class (yr)	n	Mean ASF	RMSE		Bias		
			(yr)	(%)	(yr)	(%)	Wilcoxon test (p-value)
2007							
[0, 5]	109	2.183	3.691	169.0	−2.346	40.4	<0.001
[5, 10]	93	7.602	4.917	64.7	−2.385	23.5	<0.001
[10, 15]	40	12.650	4.694	37.1	−1.344	8.2	0.092
[15, 20]	46	17.761	3.248	18.3	−0.555	2.9	0.192
>20	181	23.271	4.382	18.8	2.923	44.5	<0.001
Overall	469	13.817	4.276	31.0	−0.059	0.0	0.837
2008							
[0, 5]	109	3.183	4.197	131.8	−2.817	45.1	<0.001
[5, 10]	83	8.313	4.842	58.2	−2.236	21.3	<0.001
[10, 15]	41	12.488	6.284	50.3	−2.438	15.1	0.018
[15, 20]	55	18.309	3.675	20.1	0.112	0.0	0.877
>20	181	24.271	4.869	20.1	3.333	46.9	<0.001
Overall	469	14.817	4.736	32.0	0.036	0.0	0.516
2009							
[0, 5]	84	3.643	3.503	96.2	−2.593	54.8	<0.001
[5, 10]	93	8.151	4.832	59.3	−2.434	25.4	<0.001
[10, 15]	54	12.704	5.123	40.3	−2.546	24.7	<0.001
[15, 20]	47	18.809	3.054	16.2	−0.804	6.9	0.042
>20	191	25.047	4.724	18.9	3.129	43.9	<0.001
Overall	469	15.817	4.458	28.2	−0.046	0.0	0.704
2010							
[0, 5]	84	4.643	3.403	73.3	−2.320	46.5	<0.001
[5, 10]	61	8.180	4.750	58.1	−2.501	27.7	<0.001
[10, 15]	85	12.659	4.730	37.4	−2.173	21.1	<0.001
[15, 20]	23	18.348	4.587	25.0	0.267	0.3	0.870
>20	216	25.463	4.252	16.7	2.260	28.3	<0.001
Overall	469	16.817	4.292	25.5	−0.080	0.0	0.503

5.3. Comparison with existing datasets

5.3.1. Land cover maps of mature forest, non-forest and secondary forest

Pearson's coefficient of correlation between the proportion of each land cover class (mature forest, non-forest and secondary forest) inside every 1 km × 1 km grid cell from Carreiras et al. (2014) and this study is given in Table 7. Fig. 8 depicts the scatterplots of these proportions by land cover class in 2010. The coefficient of correlation between the proportion of each land cover class obtained from Carreiras et al. (2014) and this study ranged between 0.84 and 0.98. Consistently lower values of agreement were obtained in the secondary forest class (0.84–0.90) when compared to mature forest (0.97–0.98) and non-forest (0.96–0.98). On average, the estimates of MF proportion at each 1 km × 1 km grid cell from this study are consistently lower (~9%) than those from Carreiras et al. (2014). In terms of proportion of non-forest, the estimates from this study are on average higher than those from Carreiras et al. (2014), and diverge further with increased proportion of this class (slope = 1.242). The estimates of proportion of secondary forest obtained from this study and those obtained in Carreiras et al. (2014) are more scattered when compared to mature forest and non-forest.

The relationship between the mean ASF at each 1 km × 1 km grid cell obtained from Carreiras et al. (2014) and this study by year is shown in Fig. 9. Correlation between the estimates from both studies ranged between 0.82 and 0.85. ASF estimates from this study were on average higher than those obtained from the time-series analysis given in (Carreiras et al., 2014), with these diverging with decreasing ASF values (slope = 0.80–0.95), but more markedly in 2010.

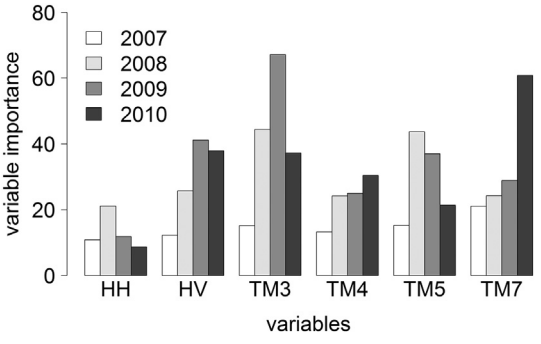
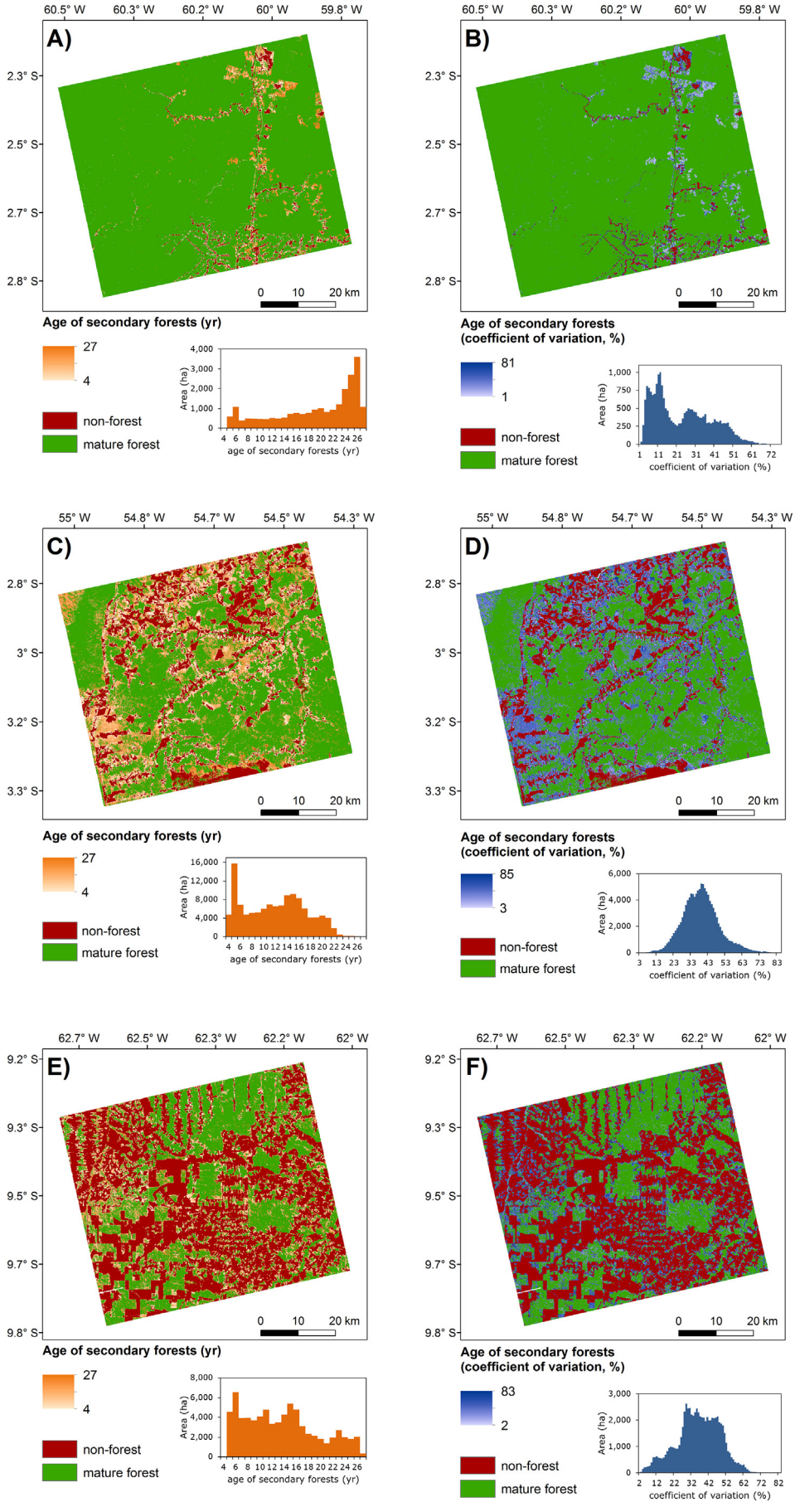


Fig. 6. Variable importance score by year (2007–2010) of the fitted random forests (RF) models using ALOS PALSAR HH and HV backscatter intensity and Landsat 5 TM surface reflectance at the TM3, TM4, TM5 and TM7 spectral regions to predict the age of secondary forests (ASF).


A)
B)
C)
D)
E)
F)
Age of secondary forests (yr)
Age of secondary forests (coefficient of variation, %)
non-forest
mature forest
Area (ha)
age of secondary forests (yr)
coefficient of variation (%)
0 10 20 km

6. Discussion

6.1. Discrimination of mature forest, non-forest and secondary forest

The omission and commission errors of the secondary forest class given in Carreiras et al. (2014) ranged between 13–26% and 2–8% respectively, with these being particularly high in Manaus and Machadinho d'Oeste, namely the omission errors: 22% and 26% respectively (Supplementary Information, Table S1). In this study, omission and commission errors in the secondary forest class considering all years in the 2007–2010 period were of similar magnitude: 12–20% and 4–6% respectively (Table 3). The accuracy assessment given in Carreiras et al. (2014) refers to the 2007 land cover map over Manaus and the 2010 maps in Santarém and Machadinho d'Oeste, with these maps being produced with image segmentation followed by decision-rule classification (Manaus) and pixel-based random forests classification (Santarém and Machadinho d'Oeste), all using only Landsat 5 TM data. Therefore, the major methodological differences between the classification approach used here and in Carreiras et al. (2014) is related to i) different training datasets, ii) using Landsat 5 TM surface reflectance data, and iii) including also ALOS PALSAR dual-pol data. The size of the training dataset might have an impact in terms of the accuracy in random forests models only if sample size reduction is greater than 50% (Rodriguez-Galiano et al., 2012). However, this is valid in cases where the original dataset can be considered representative of the land cover classes being mapped. Song et al. (2001) suggested that correction for atmospheric effects have a positive impact in terms of classification accuracy. Additionally, ALOS PALSAR data has been shown to generate accurate land cover maps and detection of forest dynamics in tropical forested regions (Longepe et al., 2011; Reiche et al., 2015; Walker et al., 2010), and, therefore, it is expected that its inclusion in this study had a positive impact when discriminating mature forest, non-forest and secondary forest.

A range of studies carried out in the Amazon showed how problematic discrimination between mature forest and secondary forest can be (Carreiras et al., 2006; Lucas et al., 2000). Additionally, some have also pointed out to higher misclassification errors between non-forest and secondary forest. Secondary forest is a highly dynamic land cover class, which might display at early stages structural and spectral characteristics similar to some non-forest classes (e.g., cropland and pasture). On the other hand, older secondary forests can be structurally and spectrally more similar to mature forests. Variations in the reflectance characteristics of secondary forests also occur as a function of the dominant species composition of the overstorey (Lucas et al., 2002). Omission and commission errors in the secondary forest class ranging from 8–11% to 7–10% were reported by Metzger (2002) when using Landsat data in Pará in the mid-1990s to map a range of land cover classes, with this being the consequence of misclassification mainly with mature forest. Vieira et al. (2003) used Landsat data from the late-1990s to map stages of secondary forest also in Pará and achieved omission and commission errors in the range of 17–25% to 22–40% respectively. Kuplich (2006) using a combination of Landsat TM and L- and C-band SAR data from the mid-1990s around Manaus reported omission and commission errors in the secondary forest class of 78% and 25%, respectively.

Comparison between the proportion of each land cover class from this study and that given in Carreiras et al. (2014) (Table 7 and Fig. 8) resulted in a good agreement in the mature forest and non-forest classes, decreasing in the secondary forest class. This is possibly a consequence of the higher misclassification errors in this class. Values reported in Table 7 and Fig. 8 are from aggregating data from the three selected sites in 2010, with Figure S1 (Supplementary Information) giving that same information at the site level. This shows that the agreement is better at Manaus than at Santarém and Machadinho d'Oeste. At Manaus, Pearson's coefficient of correlation when comparing the proportion of mature forest, non-forest and secondary forest from the two studies was 0.98, 0.96 and 0.95 respectively. The same correlation coefficients were similar or lower at Santarém and Machadinho d'Oeste, especially in the secondary forest class, with correlation markedly decreasing to 0.77 and 0.60, respectively. In Santarém, considering the time-series approach, a large area covered by secondary forests of around 10 years of age resulted from the extensive El-Nino driven fire of 1997 (Carreiras et al., 2014). That area was classified as non-forest in the subsequent year (1998) and as secondary forest since then, which can also be seen in Fig. 3B in Carreiras et al. (2014) as the extensive irregular orange secondary forest patch of intermediate age. It is possible that the 1997 fire did not impact the mature forest significantly but only its understorey (enough to leave a burnt area spectral signature) and thus when using the single-date method this area was essentially classified as mature forest. At Machadinho d'Oeste, the proportion of secondary forest at each 1 km grid cell obtained from this study is consistently lower than that obtained from the time-series analysis in Carreiras et al. (2014). A possible explanation relates to the size of the secondary forest patches in this region. The average secondary forest patch at Machadinho d'Oeste according to the time-series analysis (at 30-m spatial resolution) ranged from 1.6 ha to 2.9 ha, with this value increasing considerably to 8.3–10.0 ha when using the single-date approach depicted here (100-m spatial resolution). Interestingly, and according to the time-series analysis, Manaus displayed the highest average patch area of secondary forests (13.7–18.1 ha), with these values decreasing to 7.4–14.9 ha when the single-date method was applied. Landscape structure plays a role in the ability of remote sensing data to detect patches of a given land cover type when these have dimensions that are similar or even smaller than the spatial resolution of the analysis being carried out. It is worth noting again that the 100-m spatial resolution used in this study was the consequence of using training areas (secondary forests of known age) of this size. Some studies have pointed to patch size being an important factor (alongside heterogeneity) impacting classification accuracy, with the probability of correctly classifying a given area positively correlated to patch size and negatively correlated with heterogeneity (Lechner et al., 2009; Smith et al., 2003).

6.2. Retrieving the age of secondary forests (ASF)

6.2.1. Inter-annual variability of ALOS PALSAR and Landsat 5 TM data

At younger ages, the canopy of secondary forests is less developed, both in terms of coverage and size of tree branches and stems, thus exposing a greater proportion of soil to the sensor's field of view. Different weather conditions prior to the ALOS PALSAR acquisitions could have resulted in different soil moisture content. This has been shown to have an impact on the backscatter intensity due to differing dielectric properties, and more markedly on the HH polarization (Fig. 4) due to its strong dependence on ground scattering (Lucas et al., 2010; Mermoz et al., 2014). However, we cannot rule out that some of the variation could be due to the number of samples at each age class not being the same between 2007 and 2010 due to samples transitioning to older classes in that period (Table 2). These results are also indicative of a good absolute calibration of the data acquired by this sensor (Shimada et al., 2009), especially in areas with higher tree cover, making it useful for monitoring forest dynamics globally (Shimada et al., 2014).

Decreased variability around ALOS PALSAR HV backscatter intensity with increasing ASF (Fig. 4B) is possibly related to a decreasing contribution from ground scattering and increased contribution from volume scattering in the signal return due to increased tree branch density with the age of secondary forests. Landsat 5 TM surface reflectance acquired over secondary forests in the red to shortwave infrared spectral range indicated a much stronger dependency on acquisition date (year).

Fig. 7. Random forests (RF) based maps of age of secondary forests and corresponding coefficient of variation over Manaus (A and B), Santarém (C and D) and Machadinho d'Oeste (E and F). Non-forest and mature forest (from Fig. 3) is also shown. The histograms depict the area (ha) by age of secondary forests (yr) and corresponding coefficient of variation (%).

Table 7
Pearson's coefficient of correlation between the proportion of mature forest, non-forest and secondary forest at each 1 km × 1 km grid cell (n = 4701) overlaid on the land cover maps obtained in Carreiras et al. (2014) and from this study.

Year	Pearson's coefficient of correlation		
	Mature forest	Non-forest	Secondary forest
2007	0.97	0.97	0.85
2008	0.97	0.98	0.88
2009	0.98	0.98	0.90
2010	0.97	0.96	0.84

Only surface reflectance of TM4 and TM5 bands over secondary forests ranging between 10–15 years and 5–10 years, respectively, were not significantly different across the 2007–2010 period (Table 5). Data acquired by optical sensors are affected by, among other factors, atmospheric contamination and varying sun-sensor geometry (Li et al., 2010), both of which are related to acquisition date. The atmospheric-corrected Landsat 5 TM data used in this study are still considered provisional due to non-optimal correction over some specific regions (USGS, 2015), namely in areas with extensive cloud cover that are typical of tropical regions. Sun-sensor geometry effects occur due to changing spectral reflectance because of differing illumination and observation geometry with relation to the observed object, which was shown to impact the analysis of vegetation dynamics relying on time-series of satellite optical data (Morton et al., 2014). Therefore, as variables obtained from Landsat 5 TM data were contributing the most to the discrimination of ASF (Fig. 6), caution must be placed when analysing transitions of ASF across years. Additionally, decreasing Landsat TM5 and TM7 surface reflectance with regrowth age were also observed by other studies (Boyd et al., 1996; Lucas et al., 2002; Nelson et al., 2000) and might be explained by an increasing proportion of water-rich photosynthetic components (leaves) within each resolution cell with age, which in turn leads to increased water absorption at that spectral region and consequent lowered reflectance. To our knowledge, the results depicted in Fig. 4 give the first overview of the capability of L-band SAR and high-resolution optical data to characterize the dynamics of tropical secondary forests spanning almost 30 years.

6.2.2. Error and bias when retrieving the age of secondary forests (ASF)

Many estimates of the prediction error associated with applications of remote sensing data to retrieve biophysical parameters fail to give information about the corresponding bias-variance decomposition of the RMSE by interval of the observed values (Avitabile et al., 2016; Carreiras et al., 2012; Dubayah et al., 2010; Lefsky, 2010). As seen in this study, an overall unbiased error estimate can be a consequence of biased and unbiased estimates along the range of values of the variable being predicted (Table 6).

Comparison between the ASF estimates obtained from this study and Carreiras et al. (2014) was disaggregated by site and is given in Fig. S2 (Supplementary Information). A better agreement was consistently obtained at Manaus between 2007 and 2010, with Pearson's coefficient of correlation in the range 0.78–0.82, followed by Santarém (0.59–0.76) and Machadinho d'Oeste (0.26–0.37). Again, and similarly

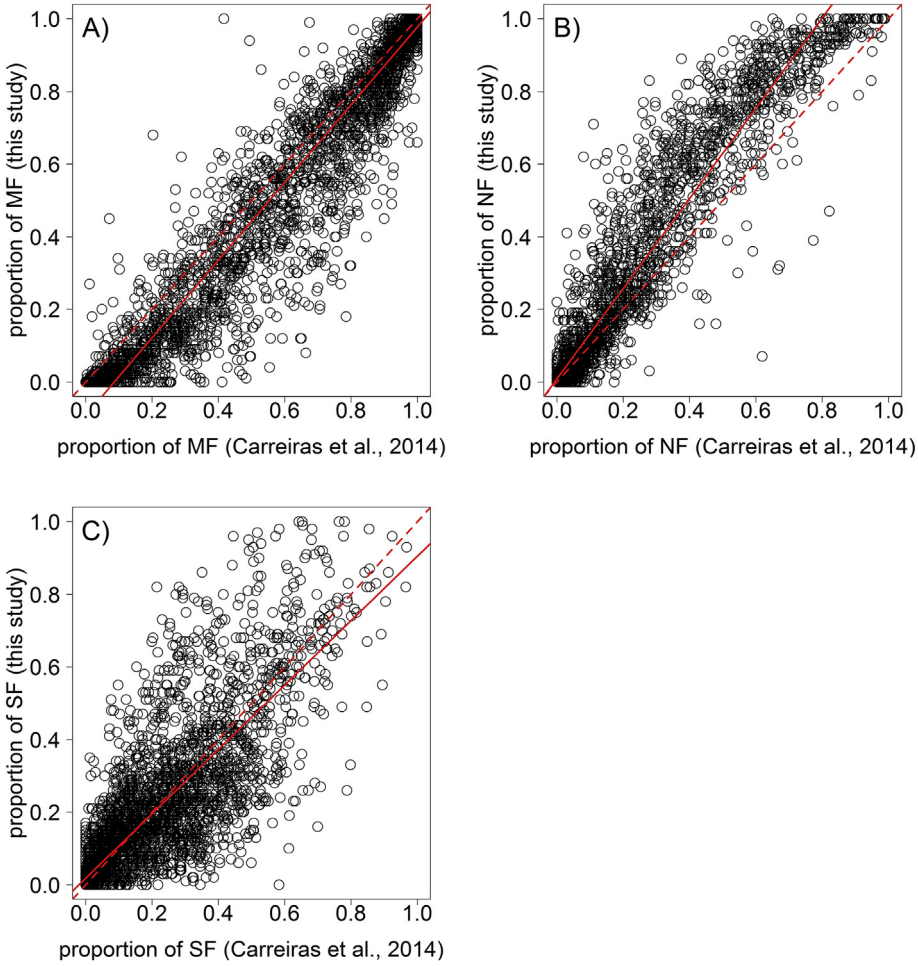


Fig. 8. Relationship between the proportion of A) mature forest (MF), B) non-forest (NF) and C) secondary forest (SF) obtained from the land cover maps produced by Carreiras et al. (2014) and in this study at each 1 km x 1 km grid cell over the 2010 maps. The solid line represents the fitted linear model between the proportion of each land cover class from Carreiras et al. (2014) and this study and the dashed line the perfect agreement.

to the results obtained during land cover classification (Section 5.1 and corresponding discussion in Section 6.1), a better agreement at Manaus is likely to be the result of comparing the results over a site that has on average larger patches of secondary forest when compared to Santarém and Machadinho d'Oeste.

6.2.3. The age of secondary forests (ASF) and biomass accumulation

The relationship between the individual variables obtained either from ALOS PALSAR dual-pol or Landsat 5 TM data and the age of secondary forests (ASF) (Fig. 4) indicate saturation around 15–20 years. Furthermore, an asymptote around 20–24 years was estimated when analysing the relationship between observed and predicted ASF from fitting random forest models (Fig. 5). This indicates that the method used to retrieve ASF loses sensitivity around 20–25 years. Tropical secondary forests have on average much higher productivity when compared to (old growth) mature forests, with growth rates being highly dependent of several factors (Brown and Lugo, 1990). The structure of tropical secondary forests (e.g., species composition, canopy height, biomass) is strongly affected by i) climate, ii) soil fertility, iii) type, duration and intensity of prior land use and iv) distance to remnant mature forests (Chazdon et al., 2007). A meta-analysis was carried out to understand the variability of above-ground biomass productivity in these forests, defined as the increment of above ground dry matter per unit area and time. Table S2 (Supplementary Information) depicts above-ground biomass productivity estimates of tropical secondary forests from 17 studies covering major tropical regions. Several factors influencing these estimates are also indicated, namely climate, prior land use intensity, and age class. The median value of tropical secondary forests growth rates by ASF class varies between 4.9 Mg ha^{-1} yr^{-1} (>20 years) and 5.9 Mg ha^{-1} yr^{-1} ([10,15] yr) (Fig. S3, Supplementary Information). The dispersion of estimates is higher for younger ASF classes (e.g., minimum and maximum equal to 0.1 and 18.0 Mg ha^{-1} yr^{-1}, respectively, in the [0,5] yr age class), possibly indicating a greater dependence on soil fertility, climate, prior land use intensity and distance to remnant forests when regeneration starts to establish on abandoned deforested areas. Growth rates in the older ASF class (>years) advanced age class have the lower dispersion (minimum and maximum equal to 1.3 and 10.6 Mg ha^{-1} yr^{-1} respectively). Considering all studies together and regardless of ASF class, the mean growth rate in secondary forests is ~6 Mg ha^{-1} yr^{-1}. If the combination of ALOS PALSAR dual-pol and Landsat 5 TM data to predict ASF loses sensitivity around 20 years, then this indicates an above-ground biomass saturation value around 120 Mg ha^{-1}.

Optical data alone have limited value to directly retrieve the above-ground biomass of tropical forests (Gibbs et al., 2007), since the upper layers of the canopy are contributing the most to the at-sensor reflectance. However, this is most useful when assessing disturbance and recovery processes (Hansen et al., 2013). On the other hand, radar data have been used extensively to estimate the above-ground biomass of forests across most tropical regions, including secondary forests. Imhoff (1995) analysed the relationship between National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL) multi-frequency AIRSAR data and field measurements over a range of forest types and observed a saturation value of ~40 Mg ha^{-1} in the above ground biomass of broadleaved evergreen forests in Hawaii. Luckman et al. (1997) used Japanese Earth Resources Satellite (JERS-

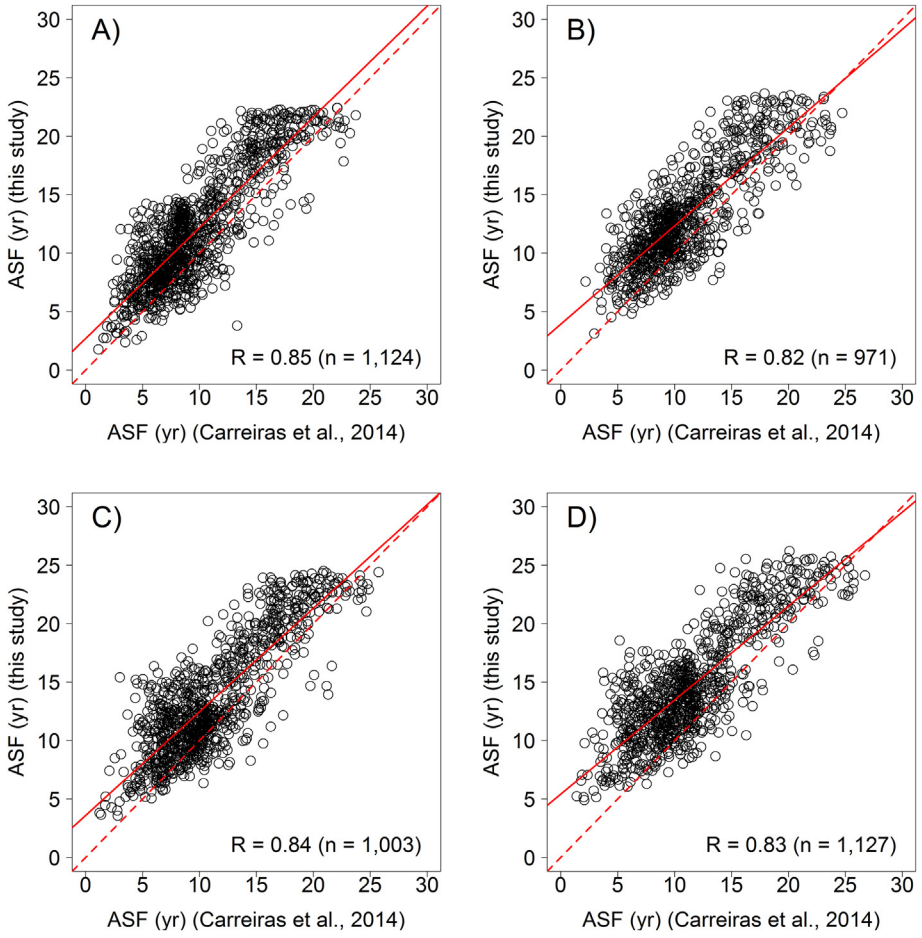


Fig. 9. Relationship between the average age of secondary forests (ASF) obtained from the maps produced by Carreiras et al. (2014) and in this study at each 1 km × 1 km grid cells in A) 2007, B) 2008, C) 2009 and D) 2010. The solid line represents the fitted linear model between the proportion of each land cover class from Carreiras et al. (2014) and this study and the dashed line the perfect agreement. Pearson's coefficient of correlation (R) and the number of 1 km × 1 km grid cells (n) are also shown.

1) single-pol (HH) and Space Shuttle Imaging Radar mission (SIR-C) dual-pol (HH and HV) L-band data acquired in 1993–1994 over tropical secondary forests in Brazil (around our Santarém site) alongside field-based above ground biomass estimates and observed a saturation limit around 60 Mg ha^{-1}. Kuplich et al. (2005) used L-band JERS-1 HH and SIR-C HH and HV data acquired over regenerating tropical forests around our Manaus and Santarém sites in 1994–1995 together with field-based above-ground biomass estimates and observed saturation also around 60 Mg ha^{-1}. More recently, Mitchard et al. (2011) used ALOS PALSAR dual-pol data acquired in 2007 to estimate the above-ground biomass content in a region in Cameroon (Central Africa) dominated by forest-savanna transition; they reported a loss of sensitivity around 150–200 Mg ha^{-1} for ALOS PALSAR HV data and lower (~100–150 Mg ha^{-1}) at HH polarization. Here, the above-ground biomass value at which the loss of sensitivity occurs (~120 Mg ha^{-1}) is similar to that reported by Mitchard et al. (2011).

7. Conclusions

The study recognized the ability of combining ALOS PALSAR dual-pol and Landsat 5 TM surface reflectance data to map mature forest, non-forest and secondary forest, with overall accuracy of 95–96% across the Brazilian Amazon in the 2007–2010 period, but with higher errors in the secondary forest class (omission and commission errors in the range 12–20% and 4–6% respectively) because of misclassification as mature forest. The method used to retrieve the age of secondary forests (ASF) generated root mean square errors in the range 4.3–4.7 years (25.5–32.0%) for forests aged up to ~30 years, with these estimates being on average unbiased. However, overestimation at younger ages (<10–15 years) and underestimation at older ages (>20 years) was observed, and varying with study area. The predictive ability over Manaus, which had on average larger patches of secondary forest, was higher, as a consequence of most ASF estimates being obtained with lower variability when compared to those at Santarém and Machadinho d'Oeste.

The method depends on having access to the location of a set of representative points covered by secondary forests of known age. Furthermore, varying illumination/observation geometry on Landsat 5 TM surface reflectance across study areas at any given year will have an impact in terms of changing the distribution of reflectance values by age of secondary forests as will the different and changing composition of the tree species dominating the upper canopy. This could be mitigated by normalizing the Landsat 5 TM surface reflectance data encompassing larger areas by using coarser, wide-swath, high-temporal resolution optical data, such as that acquired by the Moderate Resolution Imaging Spectroradiometer (MODIS) (Gao et al., 2006).

The study highlighted that ASF maps (up to a certain age) can be generated without the need for extensive time-series analysis of Landsat and/or SAR data. Moreover, the method allowed also assigning a per-pixel estimate of the prediction variability to each ASF estimate. The resulting maps in combination with above-ground biomass accumulation curves can contribute to better understand the carbon dynamics of secondary forests in the Amazon.

Acknowledgements

This study was supported by the REGROWTH-BR project (PTDC/AGR-CFL/114908/2009, Foundation for Science and Technology, FCT, Portugal). J.M.B. Carreiras' work was funded by NERC/RCUK support of the National Centre for Earth Observation (NCEO). This work has also been undertaken within the framework of the JAXA Kyoto & Carbon Initiative (K&C # 1621) and JAXA ALOS-2 RA-4 (PI-1208). ALOS PALSAR data were provided by JAXA EORC. Landsat Surface Reflectance products courtesy of the United States Geological Survey Earth Resources Observation and Science Center. The authors would also like to thank Shaun Quegan (U. Sheffield/NCEO) for his comments on earlier versions of this manuscript. We are also deeply acknowledged to the valuable comments from two anonymous reviewers who led us to improve the analysis presented here.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2017.03.016.

References

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., Malhi, Y., Morel, A., Mitchard, E.T.A., Nagy, L., Qie, L., Quinones, M.J., Ryan, C.M., Ferry, S.J.W., Sunderland, T., Laurin, G.V., Gatti, R.C., Valentini, R., Verbeeck, H., Wijaya, A., Willcock, S., 2016. An integrated pan-tropical biomass map using multiple reference datasets. Glob. Chang. Biol. 22, 1406–1420.

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. Chang. 2, 182–185.

Bacha, C.J.C., Rodriguez, L.C.E., 2007. Profitability and social impacts of reduced impact logging in the Tapajos National Forest, Brazil - a case study. Ecol. Econ. 63, 70–77.

Batistella, M., Moran, E.F., 2005. Human dimensions of land use and land cover in the Amazon: a contribution from LBA. Acta Amazon. 35, 239–247.

Bierregaard, R.O., 2001. Lessons from Amazonia: the ecology and conservation of a fragmented forest. Yale University Press, New Haven, Conn.

Boyd, D.S., Foody, G.M., Curran, P.J., Lucas, R.M., Honzak, M., 1996. An assessment of radiance in Landsat TM middle and thermal infrared wavebands for the detection of tropical forest regeneration. Int. J. Remote Sens. 17, 249–261.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brondizio, E.S., Moran, E.F., 2012. Level-dependent deforestation trajectories in the Brazilian Amazon from 1970 to 2001. Popul. Environ. 34, 69–85.

Brown, S., Lugo, A.E., 1990. Tropical secondary forests. J. Trop. Ecol. 6, 1–32.

Carreiras, J.M.B., Pereira, J.M.C., Campagnolo, M.L., Shimabukuro, Y.E., 2006. Assessing the extent of agriculture/pasture and secondary succession forest in the Brazilian Legal Amazon using SPOT VEGETATION data. Remote Sens. Environ. 101, 283–298.

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.

Carreiras, J.M.B., Jones, J., Lucas, R.M., Gabriel, C., 2014. Land use and land cover change dynamics across the Brazilian Amazon: insights from extensive time-series analysis of remote sensing data. PLoS One 9.

Cartus, O., Kellndorfer, J., Rombach, M., Walker, W., 2012. Mapping canopy height and growing stock volume using Airborne LiDAR, ALOS PALSAR and Landsat ETM. Remote Sens. (Basel) 4, 3320–3345.

Chazdon, R.L., 2003. Tropical forest recovery: legacies of human impact and natural disturbances. Perspect. Plant Ecol. Evol. Syst. 6, 51–71.

Chazdon, R.L., Letcher, S.G., van Breugel, M., Martinez-Ramos, M., Bongers, F., Finegan, B., 2007. Rates of change in tree communities of secondary neotropical forests following major disturbances. Philos. Trans. R. Soc., B 362, 273–289.

Chazdon, R.L., Broadbent, E.N., Rozendaal, D.M.A., Bongers, F., Zambrano, A.M.A., Aide, T.M., Balvanera, P., Becknell, J.M., Boukili, V., Brancalion, P.H.S., Craven, D., Almeida-Cortez, J.S., Cabral, G.A.L., de Jong, B., Denslow, J.S., Dent, D.H., DeWalt, S.J., Dupuy, J.M., Duran, S.M., Espirito-Santo, M.M., Fandino, M.C., Cesar, R.G., Hall, J.S., Hernandez-Stefanoni, J.L., Jakovac, C.C., Junqueira, A.B., Kennard, D., Letcher, S.G., Lohbeck, M., Martinez-Ramos, M., Massoca, P., Meave, J.A., Mesquita, R., Mora, F., Munoz, R., Muscarella, R., Nunes, Y.R.F., Ochoa-Gaona, S., Orihuela-Belmonte, E., Pena-Claros, M., Perez-Garcia, E.A., Piotto, D., Powers, J.S., Rodriguez-Velazquez, J., Romero-Perez, I.E., Ruiz, J., Saldarriaga, J.G., Sanchez-Azofeifa, A., Schwartz, N.B., Steininger, M.K., Swenson, N.G., Uriarte, M., van Breugel, M., van der Wal, H., Veloso, M.D.M., Vester, H., Vieira, I.C.G., Bentos, T.V., Williamson, G.B., Poorter, L., 2016. Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. Sci. Adv. 2.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792.

Dubayah, R.O., Sheldon, S.L., Clark, D.B., Hofton, M.A., Blair, J.B., Hurtt, G.C., Chazdon, R.L., 2010. Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. J. Geophys. Res. Biogeosci. 115.

Evans, T.P., Moran, E.F., 2002. Spatial integration of social and biophysical factors related to landcover change. Popul. Dev. Rev. 28, 165–186.

Ewers, R.M., Laurance, W.F., Souza, C.M., 2008. Temporal fluctuations in Amazonian deforestation rates. Environ. Conserv. 35, 303–310.

Fearnside, P.M., 2005. Deforestation in Brazilian Amazonia: history, rates, and consequences. Conserv. Biol. 19, 680–688.

Freeman, E., Frescino, T., 2009. ModelMap: modeling and map production using Random Forest and Stochastic Gradient Boosting. USDA Forest Service, Rocky Mountain Research Station, Ogden, UT, USA.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X.M., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182.

Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. IEEE Trans. Geosci. Remote Sens. 44, 2207–2218.
van Gardingen, P.R., Valle, D., Thompson, I., 2006. Evaluation of yield regulation options for primary forest in Tapajos National Forest, Brazil. For. Ecol. Manage. 231, 184–195.
Gibbs, H.K., Brown, S., Niles, J.O., Foley, J.A., 2007. Monitoring and estimating tropical forest carbon stocks: making REDD a reality. Environ. Res. Lett. 2.
Hansen, M.C., Reed, B., 2000. A comparison of the IGBP DISCover and University of Maryland 1 km global land cover products. Int. J. Remote Sens. 21, 1365–1373.
Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853.
Harris, N.L., Brown, S., Hagen, S.C., Saatchi, S.S., Petrova, S., Salas, W., Hansen, M.C., Potapov, P.V., Lotsch, A., 2012. Baseline map of carbon emissions from deforestation in tropical regions. Science 336, 1573–1576.
Hastie, T., Tibshirani, R., Friedman, J.H., 2009. The elements of statistical learning: data mining, inference, and prediction. second ed. Springer, New York, NY.
Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2015. An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. Remote Sens. Environ. 158, 220–234.
Holecz, F., Meier, E., Piesbergen, J., Nüesch, D., 1993. Topographic effects on radar cross section. Proceedings of the CEOS SAR Calibration Workshop. ESA-ESTEC, Noordwijk, The Netherlands, pp. 23–28.
Holecz, F., Meier, E., Piesbergen, J., Nüesch, D., Moreira, J., 1994. Rigorous derivation of backscattering coefficient. CEOS SAR Calibration Workshop. University of Michigan, Ann Arbor, Michigan, USA, pp. 143–155.
Hollander, M., Wolfe, D.A., 1999. Nonparametric Statistical Methods. second ed. Wiley, New York.
Imhoff, M.L., 1995. Radar backscatter and biomass saturation - ramifications for global biomass inventory. IEEE Trans. Geosci. Remote Sens. 33, 511–518.
INPE, 2015. Projeto PRODES. Monitoramento da Floresta Amazônica Brasileira por Satélite. Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, Brazil http://www.obt.inpe.br/prodes/index.php.
Keller, M., Varner, R., Dias, J.D., Silva, H., Crill, P., de Oliveira, R.C., 2005. Soil-atmosphere exchange of nitrous oxide, nitric oxide, methane, and carbon dioxide in logged and undisturbed forest in the Tapajos National Forest, Brazil. Earth Interact. 9, 1–28.
Kuplich, T.M., 2006. Classifying regenerating forest stages in Amazonia using remotely sensed images and a neural network. For. Ecol. Manage. 234, 1–9.
Kuplich, T.M., Curran, P.J., Atkinson, P.M., 2005. Relating SAR image texture to the biomass of regenerating tropical forests. Int. J. Remote Sens. 26, 4829–4854.
Laurance, W.F., Fearnside, P.M., Laurance, S.G., Delamonica, P., Lovejoy, T.E., Rankin-de Merona, J., Chambers, J.Q., Gascon, C., 1999. Relationship between soils and Amazon forest biomass: a landscape-scale study. For. Ecol. Manage. 118, 127–138.
Laurance, W.F., Camargo, J.L.C., Luizao, R.C.C., Laurance, S.G., Pimm, S.L., Bruna, E.M., Stouffer, P.C., Williamson, G.B., Benitez-Malvido, J., Vasconcelos, H.L., Van Houtan, K.S., Zartman, C.E., Boyle, S.A., Didham, R.K., Andrade, A., Lovejoy, T.E., 2011. The fate of Amazonian forest fragments: a 32-year investigation. Biol. Conserv. 144, 56–67.
Laurance, W.F., Sayer, J., Cassman, K.G., 2014. Agricultural expansion and its impacts on tropical nature. Trends Ecol. Evol. 29, 107–116.
Le Quéré, C., Moriarty, R., Andrew, R.M., Peters, G.P., Ciais, P., Friedlingstein, P., Jones, S.D., Sitch, S., Tans, P., Arneth, A., Boden, T.A., Bopp, L., Bozec, Y., Canadell, J.G., Chini, L.P., Chevallier, F., Cosca, C.E., Harris, I., Hoppema, M., Houghton, R.A., House, J.I., Jain, A.K., Johannessen, T., Kato, E., Keeling, R.F., Kitidis, V., Klein Goldewijk, K., Koven, C., Landa, C.S., Landschützer, P., Lenton, A., Lima, I.D., Marland, G., Mathis, J.T., Metzl, N., Nojiri, Y., Olsen, A., Ono, T., Peng, S., Peters, W., Pfeil, B., Poulter, B., Raupach, M.R., Regnier, P., Rödenbeck, C., Saito, S., Salisbury, J.E., Schuster, U., Schwinger, J., Séférian, R., Segschneider, J., Steinhoff, T., Stocker, B.D., Sutton, A.J., Takahashi, T., Tilbrook, B., van der Werf, G.R., Viovy, N., Wang, Y.P., Wanninkhof, R., Wiltshire, A., Zeng, N., 2015. Global carbon budget 2014. Earth Syst. Sci. Data 7, 47–85.
Lechner, A.M., Stein, A., Jones, S.D., Ferwerda, J.G., 2009. Remote sensing of small and linear features: quantifying the effects of patch size and length, grid position and detectability on land cover mapping. Remote Sens. Environ. 113, 2194–2204.
Lee, J.S., 1980. Digital image-enhancement and noise filtering by use of local statistics. IEEE Trans. Pattern Anal. Mach. Intell. 2, 165–168.
Lee, J.S., Wen, J.H., Ainsworth, T.L., Chen, K.S., Chen, A.J., 2009. Improved sigma filter for speckle filtering of SAR imagery. IEEE Trans. Geosci. Remote Sens. 47, 202–213.
Lefsky, M.A., 2010. A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. Geophys. Res. Lett. 37.
Lewis, S.L., Edwards, D.P., Galbraith, D., 2015. Increasing human dominance of tropical forests. Science 349, 827–832.
Li, F.Q., Jupp, D.L.B., Reddy, S., Lymburner, L., Mueller, N., Tan, P., Islam, A., 2010. An evaluation of the use of atmospheric and BRDF correction to standardize landsat data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 257–270.
Longepe, N., Rakwatin, P., Isoguchi, O., Shimada, M., Uryu, Y., Yulianto, K., 2011. Assessment of ALOS PALSAR 50 m orthorectified FBD data for regional land cover classification by support vector machines. IEEE Trans. Geosci. Remote Sens. 49, 2135–2150.
Lucas, R.M., Honzak, M., Curran, P.J., Foody, G.M., Milne, R., Brown, T., Amaral, S., 2000. Mapping the regional extent of tropical forest regeneration stages in the Brazilian Legal Amazon using NOAA AVHRR data. Int. J. Remote Sens. 21, 2855–2881.
Lucas, R.M., Honzak, M., Amaral, I.D., Curran, P.J., Foody, G.M., 2002. Forest regeneration on abandoned clearances in central Amazonia. Int. J. Remote Sens. 23, 965–988.
Lucas, R.M., Cronin, N., Lee, A., Moghaddam, M., Witte, C., Tickle, P., 2006. Empirical relationships between AIRSAR backscatter and LiDAR-derived forest biomass, Queensland, Australia. Remote Sens. Environ. 100, 407–425.
Lucas, R., Armston, J., Fairfax, R., Fensham, R., Accad, A., Carreiras, J., Kelley, J., Bunting, P., Clewley, D., Bray, S., Metcalfe, D., Dwyer, J., Bowen, M., Eyre, T., Laidlaw, M., Shimada, M., 2010. An EVALUATION of the ALOS PALSAR L-band backscatter-above ground biomass relationship Queensland, Australia: impacts of surface moisture condition and vegetation structure. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 576–593.
Luckman, A., Baker, J., Kuplich, T.M., Yanasse, C.D.F., Frery, A.C., 1997. A study of the relationship between radar backscatter and regenerating tropical forest biomass for spaceborne SAR instruments. Remote Sens. Environ. 60, 1–13.
Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., Lim, T.K., 2006. A Landsat surface reflectance dataset for North America, 1990–2000. IEEE Geosci. Remote Sens. Lett. 3, 68–72.
Mayaux, P., Bartholome, E., Fritz, S., Belward, A., 2004. A new land-cover map of Africa for the year 2000. J. Biogeogr. 31, 861–877.
Meier, E., Frei, U., Nüesch, D., 1993. Precise terrain corrected geocoded images. In: Schreier, G. (Ed.), SAR Geocoding: Data and Systems. Herbert Wichmann Verlag, Karlsruhe, Germany, pp. 173–185.
Mermoz, S., Toan, T.L., Villard, L., Rejou-Mechain, M., Seifert-Granzin, J., 2014. Biomass assessment in the Cameroon savanna using ALOS PALSAR data. Remote Sens. Environ. 155, 109–119.
Metzger, J.P., 2002. Landscape dynamics and equilibrium in areas of slash-and-burn agriculture with short and long fallow period (Bragantina region, NE Brazilian Amazon). Landsc. Ecol. 17, 419–431.
Miranda, E.E., 2009. Sustentabilidade Agrícola na Amazonia - 23 anos de monitoramento da agricultura em Machadinho d'Oeste (RO). Embrapa Monitoramento por Satélite, Campinas, SP, Brazil http://www.machadinho.cnpm.embrapa.br/index.html.
Mitchard, E.T.A., Saatchi, S.S., Woodhouse, I.H., Nangendo, G., Ribeiro, N.S., Williams, M., Ryan, C.M., Lewis, S.L., Feldpausch, T.R., Meir, P., 2009. Using satellite radar backscatter to predict above-ground woody biomass: A consistent relationship across four different African landscapes. Geophys. Res. Lett. 36.
Mitchard, E.T.A., Saatchi, S.S., Lewis, S.L., Feldpausch, T.R., Woodhouse, I.H., Sonke, 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, 2861–2873.
Morton, D.C., Nagol, J., Carabajal, C.C., Rosette, J., Palace, M., Cook, B.D., Vermote, E.F., Harding, D.J., North, P.R.J., 2014. Amazon forests maintain consistent canopy structure and greenness during the dry season. Nature 506, 221–224.
Nelson, R.F., Kimes, D.S., Salas, W.A., Routhier, M., 2000. Secondary forest age and tropical forest biomass estimation using thematic mapper imagery. Bioscience 50, 419–431.
Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., da Motta, R.S., Armijo, E., Castello, L., Brando, P., Hansen, M.C., McGrath-Horn, M., Carvalho, O., Hess, L., 2014. Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. Science 344, 1118–1123.
Oliver, C., Quegan, S., 1998. Understanding synthetic aperture radar images. Artech House, Boston.
Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57.
Pinheiro, J.C., Bates, D.M., 2000. Mixed-effects models in S and S-PLUS. Springer, New York.
Pisek, J., Chen, J.M., 2007. Comparison and validation of MODIS and VEGETATION global LAI products over four BigFoot sites in North America. Remote Sens. Environ. 109, 81–94.
Prates-Clark, C.D., Lucas, R.M., dos Santos, J.R., 2009. Implications of land-use history for forest regeneration in the Brazilian Amazon. Can. J. Remote. Sens. 35, 534–553.
Quegan, S., Yu, J.J., 2001. Filtering of multichannel SAR images. IEEE Trans. Geosci. Remote Sens. 39, 2373–2379.
Reiche, J., Verbesselt, J., Hoekman, D., Herold, M., 2015. Fusing Landsat and SAR time series to detect deforestation in the tropics. Remote Sens. Environ. 156, 276–293.
Reiche, J., Lucas, R., Mitchell, A.L., Verbesselt, J., Hoekman, D.H., Haarpaintner, J., Kellndorfer, J.M., Rosenqvist, A., Lehmann, E.A., Woodcock, C.E., Seifert, F.M., Herold, M., 2016. Combining satellite data for better tropical forest monitoring. Nat. Clim. Chang. 6, 120–122.
Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J. Photogramm. Remote Sens. 67, 93–104.
Santoro, M., Beer, C., Cartus, O., Schmullius, C., Shvidenko, A., McCallum, I., Wegmuller, U., Wiesmann, A., 2011. Retrieval of growing stock volume in boreal forest using hyper-temporal series of Envisat ASAR ScanSAR backscatter measurements. Remote Sens. Environ. 115, 490–507.
Scholz, F.W., Stephens, M.A., 1987. K-sample Anderson-Darling tests. J. Am. Stat. Assoc. 82, 918–924.
Shimada, M., Isoguchi, O., Tadono, T., Isono, K., 2009. PALSAR radiometric and geometric calibration. IEEE Trans. Geosci. Remote Sens. 47, 3915–3932.
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.
Silver, W.L., Neff, J., McGroddy, M., Veldkamp, E., Keller, M., Cosme, R., 2000. Effects of soil texture on belowground carbon and nutrient storage in a lowland Amazonian forest ecosystem. Ecosystems 3, 193–209.
Smith, J.H., Stehman, S.V., Wickham, J.D., Yang, L.M., 2003. Effects of landscape characteristics on land-cover class accuracy. Remote Sens. Environ. 84, 342–349.
Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., Macomber, S.A., 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? Remote Sens. Environ. 75, 230–244.

Touzi, R., 2002. A review of speckle filtering in the context of estimation theory. IEEE Trans. Geosci. Remote Sens. 40, 2392–2404.
Ulaby, F.T., Dobson, M.C., 1989. Handbook of radar scattering statistics for terrain. Artech House, Norwood, MA.
USGS, 2015. Product Guide. Landsat 4-7 Climate Data Record (CDR). Surface Reflectance. Version 6.1. United States Geological Survey, Department of the Interior.
Venables, W.N., Ripley, B.D., 1999. Modern applied statistics with S-PLUS. third ed. Springer, New York.
Vieira, I.C.G., de Almeida, A.S., Davidson, E.A., Stone, T.A., de Carvalho, C.J.R., Guerrero, J.B., 2003. Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazonia. Remote Sens. Environ. 87, 470–481.
Walker, W.S., Stickler, C.M., Kellndorfer, J.M., Kirsch, K.M., Nepstad, D.C., 2010. Large-area classification and mapping of forest and land cover in the Brazilian Amazon: a comparative analysis of ALOS/PALSAR and Landsat data sources. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 594–604.
Woodhouse, I.H., 2006. Introduction to Microwave Remote Sensing. Taylor & Francis, Boca Raton.