Detection and characterization of agroforestry systems in the Colombian Andes using sentinel-2 imagery

Sergio Bolívar-Santamaría · Björn Reu

Abstract In the Colombian Andes, agroforestry is a traditional form of agriculture, characterized by a heterogeneous and often diversified composition of trees and crops. This form of land use provides important ecosystem services, such as carbon sequestration, reduction of soil erosion and the maintenance of biodiversity by providing a structural complex habitat. Satellite remote sensing is widely used for studying land use patterns and forest cover, however the discrimination between agroforestry systems and forests is still a challenge, especially in heterogeneous landscapes and in rough terrain. Here, we aim to advance the remote sensing of agroforestry systems using field measurements of vegetation structure in combination with Sentinel-2 images. We use spectral and textural variables derived from Sentinel-2 imagery to predict above ground biomass (AGB), leaf area index (LAI) and canopy closure (CC). The relationship between predicted and observed values obtained from Random Forest regression models showed good fits: for AGB with an $R^2 = 0.92$ and relative RMSE = 34%; for LAI with an $R^2 = 0.91$ and relative

RMSE = 19%; and for CC an $R^2 = 0.89$ and relative RMSE = 9%. This allowed us to map these important ecosystem variables at landscape scale and establish empirical thresholds, with which a discrimination of agroforestry systems from forests was possible with an accuracy of 94%. Our results suggest that the relationship between vegetation structure and the spectral information obtained by Sentinel-2 can contribute to the detection and characterization of agroforestry systems and thus help quantifying the ecosystem services and biodiversity conservation potential provided by this type of tropical agriculture.

Introduction

The term agroforestry system (AFS) generally refers to a traditional agricultural practice, where trees are integrated into the cultivation of crops (Nair 1985). In Latin America approximately 300 million hectares of land are used for agroforestry, which sustains the economy of rural families, mainly due to the production of cacao and coffee (Somarriba et al. 2012). Furthermore, AFS are considered an effective way to

S. Bolívar-Santamaría · B. Reu (✉)
School of Biology, Universidad Industrial de Santander, UIS, Cra 27 Calle 9 Ciudad Universitaria, Bucaramanga, Colombia
e-mail: breu@uis.edu.co

contribute to food security, the conservation of biodiversity and socio-ecological resilience in agricultural landscapes (Sharma et al. 2016; Leimona and Noordwijk, 2017; Mbow et al. 2017; Yapo 2019). Therefore, the last IPCC report (Shukla et al. 2019) suggest that agricultural production using agroforestry represents a strategy to achieve sustainable development goals and climate change mitigation in tropical regions (Garrity 2004). For this reason, countries, such as Colombia, Perú and Nicaragua have started initiatives to promote rural development through the implementation of AFS (Porras et al. 2015). Consequently, AFS should be taken into account in decision-making and policies for sustainable development in tropical countries, which would require a mapping and quantification of their spatial extent (Waldron et al. 2017). However, so far, AFS in tropical countries are not considered in international initiatives on crop assessment (e.g. http://jecam.org) and are not monitored using Earth observation satellites. This is because AFS are often embedded in a small scale and heterogeneous matrix of different land-use types. The CORINE Land Cover methodology adapted for Colombia, for example notes that additional information, such as aerial photography or high-resolution satellite imagery is necessary to identify coffee and cacao crops planted as agroforests (IDEAM 2010). Even more, due to the presence of shade-trees, AFS are often classified as forests in remote sensing products based on automated classification algorithms (i.e. Global Forest Watch, Watch 2002).

There are only a few studies in the tropics that mapped AFS using remote sensing, and even fewer that characterized canopy structure. For example, the mapping of AFS has been mainly done using high resolution imagery such as Quickbird or WorldView 2 using visual interpretation (Bégué et al. 2015), classification algorithms based on textural features (Gomez et al. 2010; Lelong et al. 2014), or a combination of different remote sensing products from Landsat, MODIS or IKONOS (Zomer et al. 2007). Mapping of canopy structure, such as leaf area index or above-ground biomass in tropical AFS however requires multispectral information, such as provided by MODIS and Sentinel-2 (Taugourdeau et al. 2014; Karlson et al. 2020). While the aforementioned studies are distributed around the tropics, there is no study to our knowledge about remote sensing of AFS in the Andean region.

In the Colombian Andes cacao and coffee crops are frequently grown under shade trees, which creates a more complex horizontal and vertical vegetation structure than monocultures, increasing biomass, leaf area index and canopy closure (Isaac et al. 2007; Dossa et al. 2008). This creates habitats that can be exploited by different organisms that interact in their trophic networks (Klein et al. 2006). Furthermore, the presence of crops with trees in a heterogeneous agricultural matrix can enhance connectivity among patches of natural vegetation, thus promoting biodiversity conservation outside protected areas (Jose 2012) and sustainable ecosystem management (Garrity et al. 2006).

Due to the high structural complexity and the presence of native plant species, AFS have been highlighted as biodiversity refuges (Bhagwat et al. 2008). For example, an increased canopy closure exerted by shade trees helps maintain habitat quality, by which biodiversity of different taxa can be preserved (Klein et al. 2002; Brüning et al. 2018). AFS have also been identified as carbon sinks because of their capacity to sequester carbon in both soil and biomass (Cardinael et al. 2018). Some estimations have shown that in the Andean region the biomass in AFS reaches up to 87.37 tons per hectare (Orozco et al. 2015), subject to the species composition and climate. Although the amount of above-ground biomass stored in AFS is in general lower than those in mature forests, they store more biomass than coffee and cacao monocultures. The structural complexity in the canopy of AFS not only promotes stand productivity and biodiversity through complementary resource utilization (Scheper 2019), but also improves the habitat quality for canopy-dwelling and below-ground organisms (Ishii et al. 2004). The presence of a canopy layer in AFS offers multiple regulating ecosystem services, such as microclimate regulation, buffering drought events (DaMatta 2004), the prevention of soil erosion and pest control (Kuyah et al. 2017).

Therefore, variables of forest stand and canopy structure are essential for quantifying ecological processes and ecosystem services, but are often difficult and demanding to obtain directly. For example, hemispherical photography using canopy pictures with a fish-eye lens (Garrigues et al. 2008) is an inexpensive and efficient way to obtain relatively fast measurements of leaf area index and canopy closure (Garrigues et al. 2008; Fournier and Hall 2017).

Terrestrial laser scanning has also been used to measure canopy openness and canopy gap distributions at high precision (Seidel et al. 2012) and to estimate the distribution of biomass (Calders et al. 2015), but is more expensive. Satellite remote sensing along with ground measurements has been used to efficiently predict structural and biophysical variables of forests (Boyd and Danson 2005; Korhonen et al. 2015, 2017) and to classify different forest types (Laurin et al. 2016; Erinjery et al. 2018; Morin et al. 2019). It even allows to derive continuous variables such as canopy closure (CC), leaf area index (LAI) and above-ground biomass (AGB) that are key variables characterizing forest stand and canopy structure. For example, models based on Sentinel-2 imagery in combination with hemispherical photography as ground truth have shown to accurately predict LAI (Korhonen et al. 2017). Also, above-ground biomass (AGB) has been derived successfully for forests using satellite images at local or regional scales (TSITSI 2016; Korhonen et al. 2017). Even canopy features like tree cover has been estimated for forests worldwide and made available freely (Hansen et al. 2013; Martone et al. 2018), however, these products are limited in their spatial resolution and plantations or AFS are often classified as forests (Tropek et al. 2014).

AFS structure in the tropics is more complex than in temperate regions, where trees are often located as field boundaries in order to function as windbreaks (Nair 1985). Due to the mixed and heterogeneous arrangement of shade trees inside the Andean AFS of coffee and cacao, it is hard to recognize and differentiate them from forests using remote sensing. Nevertheless, density of shade trees causes variation in CC, LAI and AGB, which could be useful to distinguish AFS from forests in a heterogeneous landscape matrix. Even though AFS are poorly characterized in terms of their structure and biomass, their monitoring could support decision making that leads towards the improvement of the benefits provided by AFS to nature and people. The objective of this study is to detect AFS through the mapping of AGB, LAI and CC using Sentinel-2 spectral and textural information in the Colombian Andes.

Material and methods

Study area

The study area is a micro watershed called *Las Cruces* and is located on the Serranía de los Yariguíes, an isolated mountain ridge in the Northern Andes of Colombia (Fig. 1). *Las Cruces* is part of the buffer zone of the Yariguíes National Natural Park in the municipality of San Vicente de Chucurí. The microwatershed has an extension of 5779 ha, ranges in elevation between 570 and 2650 m.a.s.l.; annual precipitation ranges between 1500 and 1700 mm and the mean temperature is around 22.5 °C (Pinilla et al. 2018). *Las Cruces* is very important for the water provisioning of San Vicente de Chucurí and for the production of food, such as cacao, coffee, avocado, citric fruits and vegetables, which are cultivated in a heterogeneous landscape consisting of mainly diversified AFS but also intensified monocultures as well as forests fragments and cattle pastures. *Las Cruces* has been studied within the GEF-Satoyama project (http://gef-satoyama.net/) and proposed as a Socio-Ecological Production Landscape (SEPL) highly valuable for the conservation of biodiversity.

Detection of AFS

In order to distinguish AFS reliably from forests and permanent crops such as coffee and cacao, we implemented a methodology based on forest structure variables such as LAI, AGB and CC. These variables were assessed in 50 on-ground plots and related to a diverse set of spectral and textural variables derived from a Sentinel-2 imagery (Fig. 2). The 50 plots were selected based on a stratified sampling design, in which production systems and forests were classified a priori based on a visual estimation of their canopy closure (i.e. low shade cultivation or monocultures: less than 30% canopy closure; agroforests: between 30 and 70% canopy closure; forests: more than 70% canopy closure). This resulted into a homogeneous ground sampling across the management gradient from low shade monocultures (11 plots) to diversified AFS (29 plots) and conserved forests (10 plots) in the vicinity of the National Natural Park Serranía de los Yariguíes (Fig. 1).
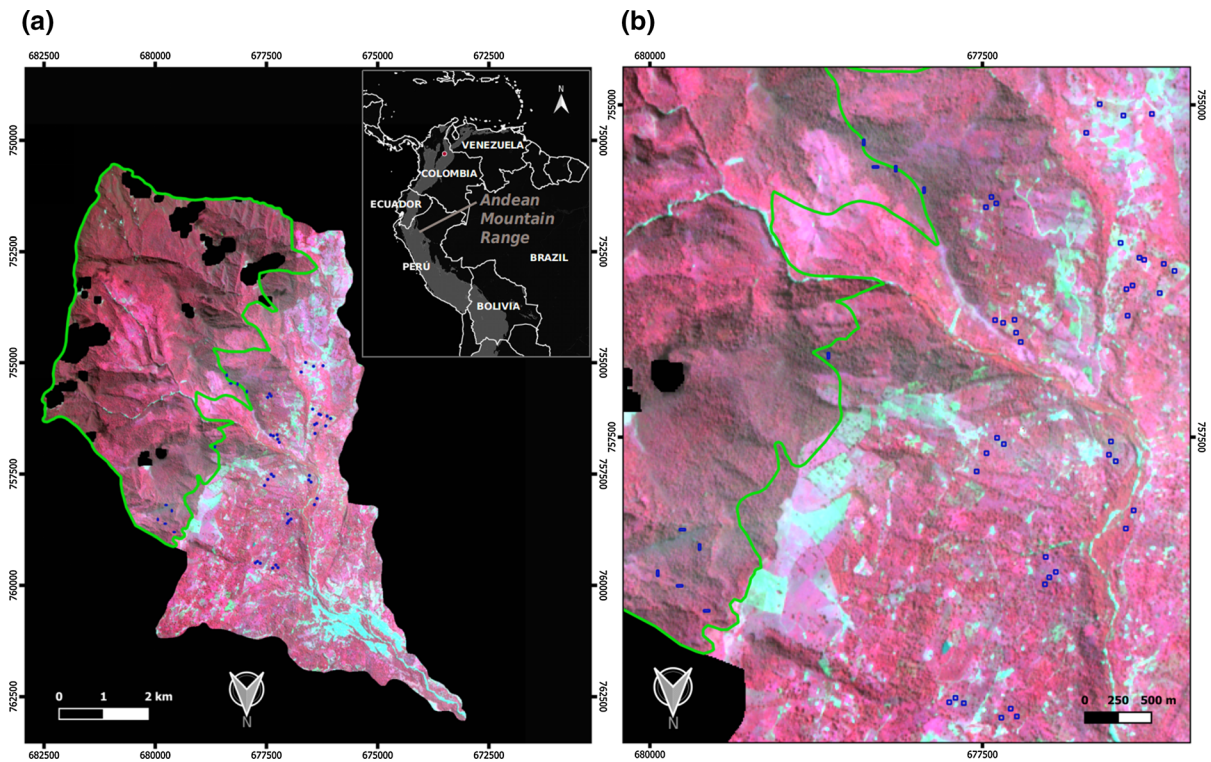
(a)

(b)

Fig. 1 **a** False-color Sentinel-2 composite image of the study area from June 2018, green line represents the limits of Yariguíes National Natural Park. **b** Zoom in showing the locations of vegetation plots in blue within the *Las Cruces* micro-watershed

Ground measurements and calculation of vegetation structure variables

Stand structure of AFS, including tree density, tree height, basal area and species composition was assessed using the Point Centered Quarter method (PCQM) (Mitchell 2015). The PCQM has shown to be an efficient and reliable method to characterize forest stand structure (Jafari et al. 2013; Manduell et al. 2012). Also PCQM has proved to be appropriate for ground-truth data in remote sensing analyses (Satyanarayana et al. 2018) and we consider it as an efficient sampling method in cultivation systems that exhibit some sort of planting design. In order to apply PCQM to AFS, we did two modifications. First, linear transects were modified to a grid-based design (i.e. rectangular plots), with a size of 30 by 30 m, with 9 sampling points each 15 m (Fig. 3a). Second, at each sampling point we applied PCQM sampling twice, once for the crop trees (referred to as midstory), and once for the shade trees (referred to as overstory). In the forest, plots consisted of two linear transects with 8

points and 15 m between them (Fig. 3b), because of the inaccessibility and steep slope of the terrain. The mid- and overstory in forest were defined in an arbitrary manner separating trees below and above 10 m. Tree density (for mid- and overstory) was estimated based on the average distance from nearest neighbor trees. Correction factors were applied in case when an overstory tree appeared in more than one quadrant as a nearest neighbor tree (Mitchell 2015; Warde and Petranka 1981). Basal area (BA) of each tree was calculated from its diameter at breast height (DBH). Additional to PCQM sampling we assessed LAI and CC using hemispherical canopy photography. To do so, on each plot we took four pictures in the middle of each quadrant spanned up by the 9 PCQM sampling points using a Canon Powershot G6 camera and a Fish eye lens Raynox DCR-CF 185PRO. The hemispherical pictures were taken at twilight, avoiding the direct sun to achieve a good contrast between the canopy and the sky and were analyzed using Gap Light Analyzer (GLA, Frazer et al. 1999) to estimate LAI and CC. Finally, we calculated AGB for each
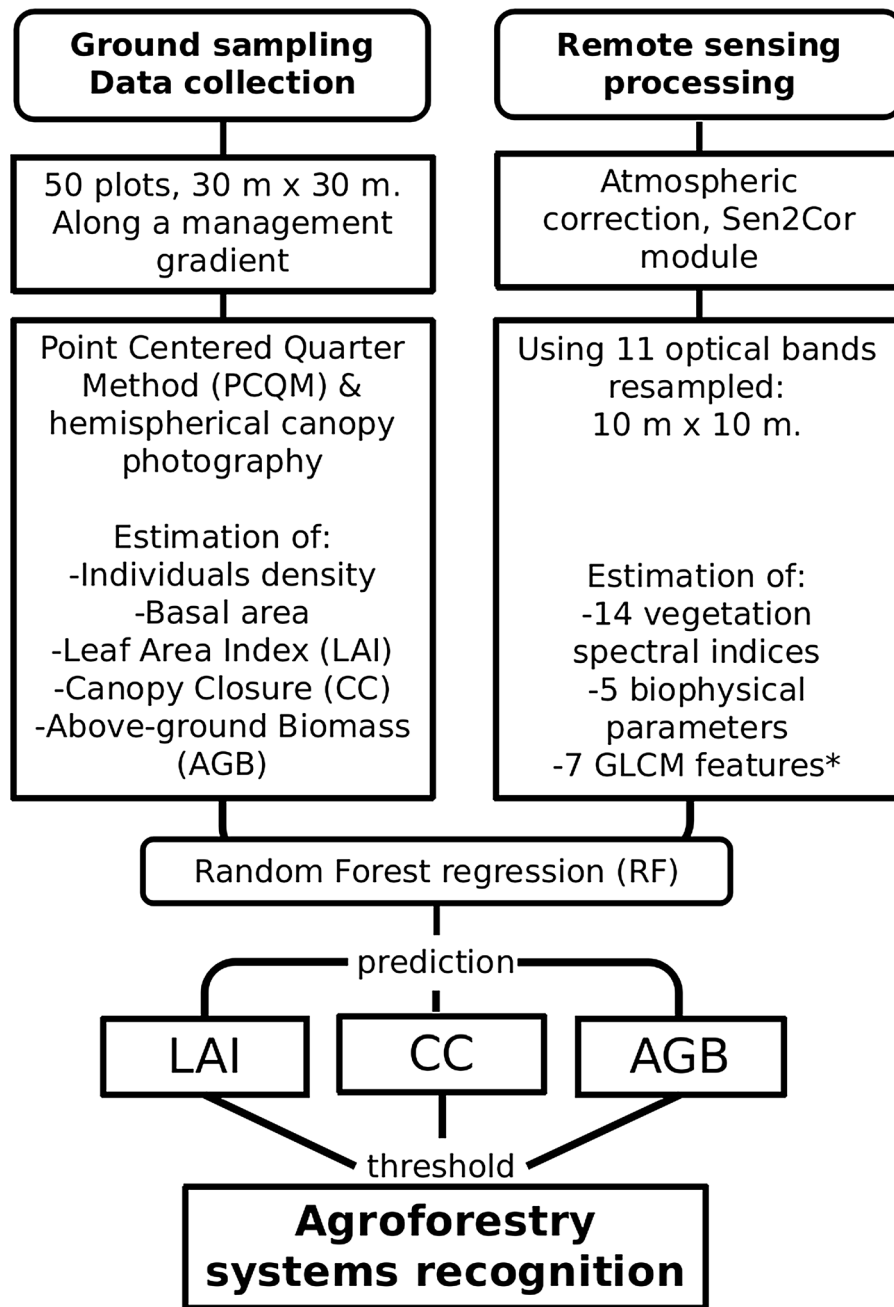
Fig. 2 Workflow summarizing the recognition of agroforestry systems (AFS) by means of field measurements and remote sensing using Sentinel-2 imagery. *GLCM refers to texture features calculated from Grey Level Co-ocurrence Matrices

plot, and mid- and overstory separately, using allometric equations for coffee (Segura et al. 2006), for cacao, fruit tree species and Musaceae plants (Somarriba et al. 2013) as well as for shade trees (Alvarez et al. 2012) by multiplying the mean biomass of a mid- and overstory tree with the tree density of each layer.

In order to explore the structural variability across plots, we use Principal Component Analysis (PCA) considering all vegetation structure variables. Field work was done from April to October 2018 and subsequent calculations of structural variables were
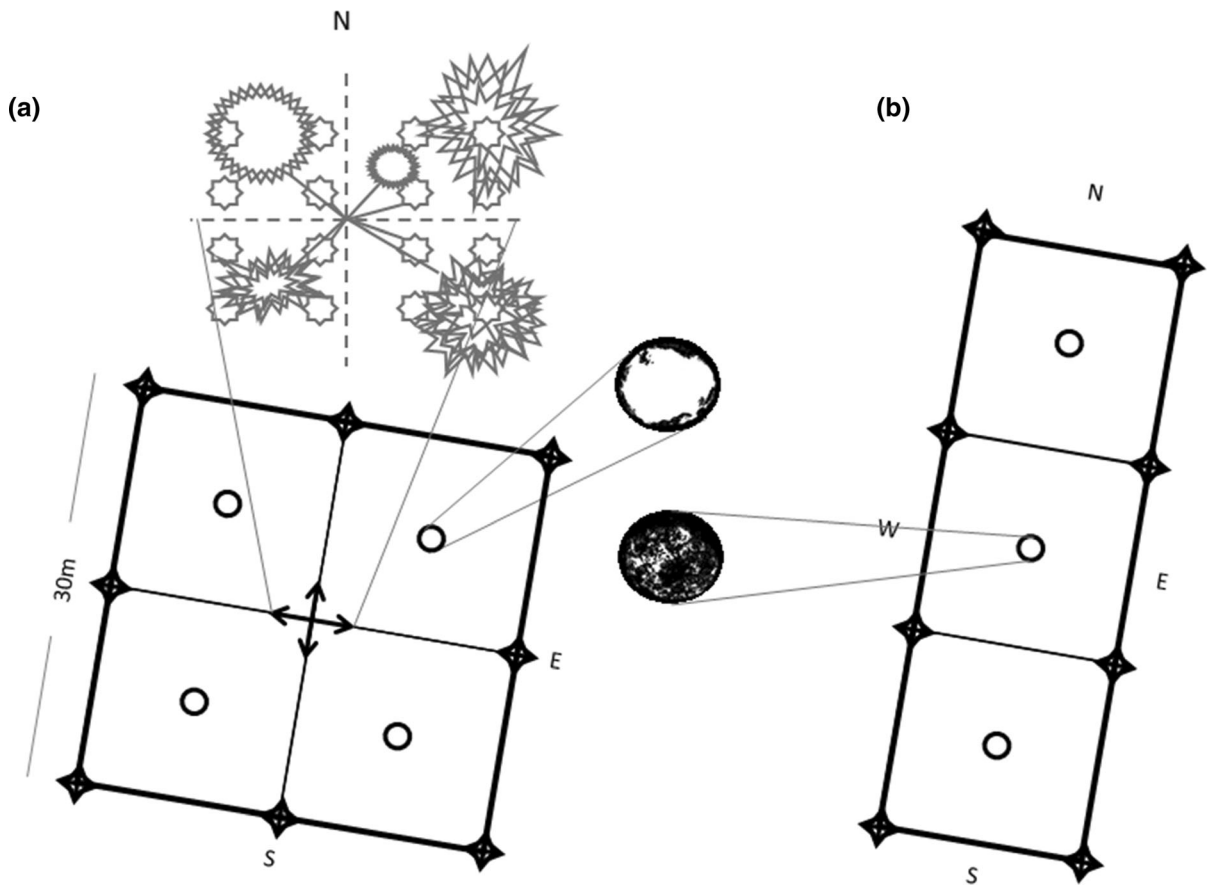
Fig. 3 Plot design using the PCQM. Distribution of PCQM points (shown as crosses) and hemispheric canopy photograph (circles) in **a** AFS and **b** forests

done using the R environment for statistical computing, Version 3.6.2. (Development Core Team 2011).

Remote sensing variables derived from Sentinel-2 imagery

We downloaded a Sentinel-2 scene from 23rd of June 2018 using ESA's platform Copernicus scihub (https://scihub.copernicus.eu/dhus/#/home). This scene was the only almost cloud-free image that could be obtained between January and October 2018 for the study area. We applied atmospheric correction using the Sen2Cor processor from the Sentinel Application Platform (SNAP v6.0. 2019). For further analyses all spectral bands were spatially resampled to 10 m using as a reference the blue band, and reprojected to UTM 18 N zone using nearest neighbor resampling for both steps. Finally, the area of the image was masked to the area of the *Las Cruces* micro-watershed (Fig. 1).

Using the Thematic Land Processing tool in SNAP, we estimated 14 radiometric derived vegetation indices (VIs, see Supplementary Table 1), and 5 biophysical parameters: leaf area index, fraction of absorbed photosynthetically active radiation (fapar), fraction of vegetation cover (fcover), Chlorophyll content in the leaf (Cab) and Canopy Water Content (CW). The biophysical parameters were estimated based on neural networks where 5 radiative transfer models (or neurons in the hidden layer) were trained to estimate each of the before mentioned variables using 11 normalized spectral bands as input layer (Weiss and Baret 2016). In order to consider image textures, we calculated texture layers using Grey Level Co-ocurrence Matrix (Haralick and Shanmugam 1973) using a moving window size of 5 by 5 pixels which resulted into 7 textural features known as contrast, dissimilarity, homogeneity, entropy, mean, variance and correlation (Haralick and Shanmugam 1973). Texture

layers were calculated using the NDVI band with the Texture Analysis tool in SNAP. Hence, taking into account 11 spectral bands of Sentinel 2 (excluding B1 and B10, because their coarse spatial resolution of 60 m and little information beyond their help to atmospheric correction), 37 layers were generated to be used as variables for predicting LAI, AGB and CC measured on ground in the 50 plots.

Predicting vegetation structure variables from Sentinel-2 spectral, biophysical and textural variables

Using the coordinates in the center of the AFS plots and at the corner points of the forest plots, we generated polygons for the plots and extracted the mean of pixel values within each polygon in the 37 raster layers. We generated a matrix of 37 predictor variables for the 50 plots (Fig. 2), which were used to predict LAI, AGB and CC using Random Forest regression. For Random Forest regression we used default values for the parameters *ntree* (number of trees) and *mtry* (number of variables randomly sampled as candidates at each split). To evaluate the model quality, we report the out-of-bag error of the Random Forest model, the Pearson's correlation coefficient (R^2), the root mean squared error (RMSE) as well as the relative RMSE (in %) resulting from the comparison of observed versus predicted values. Before generating maps for the three response variables LAI, AGB and CC, we realized a supervised land-use classification to mask pixels of water, urban settlement, bare soil and pastures. To do so, we used Random Forest classification with 324 training points (109 for forest, 47 for pasture, 43 for crops, 52 for water, 39 for bare soil and 34 for urban cover) mapped during field work and the 11 spectral bands of the Sentinel-2 image (excluding B1 and B10, because of their coarse resolution). The Random Forest classification resulted in an overall accuracy of 86% derived from the confusion matrix. For the remaining pixels, we predicted LAI, AGB and CC using Random Forest regression as explained above. Feature extraction, model building and predictions were done using the packages "raster", "rgdal" and "RandomForest" within the R language for statistical computing (Hijmans and Van Etten., 2012; Bivand et al. 2015; Liaw and Wiener 2002; R Development Core Team

2011). The final maps were laid out using QGIS 3.6.2 Noosa (QGIS Development Team 2015).

To detect AFS, we defined empirical thresholds for AGB, LAI and CC based on our model predictions (Fig. 5), the on-ground measurements (see Supplementary Fig. 1) and literature values about AGB and CC in tropical AFS (Nair et al. 2010; Dhyani et al. 2020). Then, we validated the accuracy of the classification using the confusion matrices of two classifications, one with two categories (AFS vs. forest) and another with three (monocultures vs. AFS vs. forest). Finally, in order to evaluate the CC of AFS in the context of global forest cover, we compared our predictions of CC with tree cover derived from the global forest change data-set (Hansen et al. 2013) for the year 2018.

Results

Vegetation structure

A principal component analysis of the vegetation structure variables (see Table 1) derived from the PCQM sampling and hemispherical photography uncovered a gradient from monocultures, over AFS to natural forests that is summarized in two dimensions representing more than 72% of the variance (Fig. 4). Forests differentiate from the remaining vegetation because of their higher AGB, LAI, CC and tree density of the overstory. However, the differentiation of AFS from monocultures is less pronounced. Furthermore, structural variation within the AFS emerge from the density and basal area of midstory trees, which however is less well captured by the a priori classification.

Predicting vegetation structure variables from Sentinel-2 variables

Random Forest regression (RF) allowed very good predictions of LAI, AGB and CC from Sentinel-2 imagery derived variables (Fig. 5). Linear models between predicted and observed values for LAI, AGB and CC showed a high accuracy of predictions with R^2 of 0.91 for LAI, 0.92 for AGB, and 0.9 for CC, respectively (Fig. 5a, b, c). This translates into relative root mean squared errors (RRMSE) of 19% for LAI, 34% for AGB and 9% for CC (Fig. 5). In general, for

Table 1 Vegetation structure variables assessed in the 50 field plots

| Variable	Abbreviation	Units	Mean	Range
Tree density of midstorey layer	D_mid	Number of individuals per plot	250.89	46.15–831.66
Tree density of overstorey layer	D_over	Number of individuals per plot	17.90	0.68–64
Total basal area	BA	$m^2\ ha^{-1}$	1.923	0.25–7.08
Basal area in midstorey	BA_mid	$m^2\ ha^{-1}$	0.29	0.03–0.95
Basal area in overstorey	BA_over	$m^2\ ha^{-1}$	1.63	0.13–7.03
Mean height of crops	H_mid	m	3.35	1.13–6.86
Mean height of shade trees	H_over	m	12.24	2.84–24.50
Leaf area index	LAI	Unitless	1.63	0.07–3.60
Canopy closure	CC	%	70.02	19.45–93.25
Above ground biomass	AGB	$Mg\ ha^{-1}$	103.36	4.38–405.22
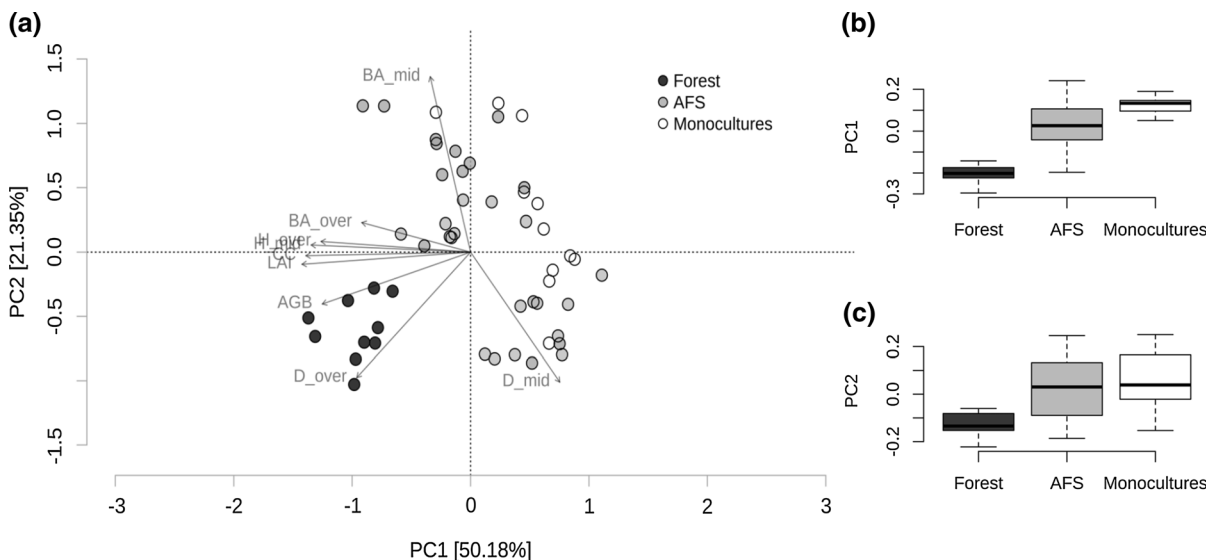
Fig. 4 Results of the principal components analysis (PCA) showing the structural variation across field plots (**a**). Black dots represent forest plots, grey dots agroforestry systems (AFS) and white dots monocultures under low shade. Variables and abbreviations are the same as in Table 1, 'mid' refers to midstory (i.e. crop layer), while 'over' refers to overstory (i.e. shade tree layer). The proportion of variance explained by each principal component is provided in brackets at each axis. The two inserts on the right (**b** and **c**) show the distribution of PC coordinates among the three land-use types, clearly separating forests from AFS on PC1

all three variables, models slightly underestimate high values and overestimate small values. The most important variables for LAI, AGB and CC predictions were the atmospherically corrected optical bands of the Sentinel-2 image, followed by some radiometric derived biophysical parameters: canopy water content (lai_cw), leaf area index (lai) and fraction of photo-synthetically active radiation (fapar). Vegetation spectral indices (VIs) and Grey Level Co-ocurrence Matrix (GLCM) features did not show a strong influence in the model predictions (Fig. 5). Finally, based upon the predictions of the models obtained for the three variables, we established in a pragmatic manner thresholds that permit to distinguish forests from AFS, which is a LAI above 2.25, and AGB above 139 Mg ha^{-1} and a CC above 82%. These thresholds
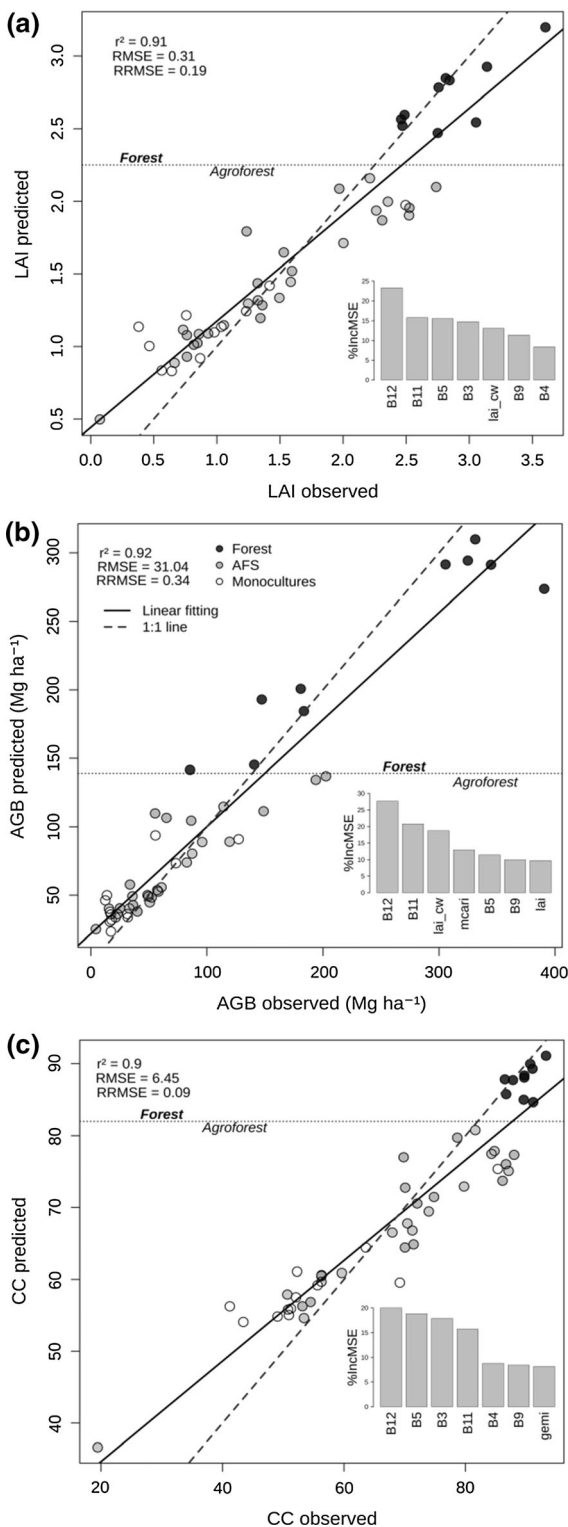
Fig. 5 Scatter plots between observed and predicted values of LAI, AGB, CC. The insert on the bottom right shows the most important predictor variables for each model based on the Mean Decrease in Accuracy (%IncMSE) derived from the out-of-bag procedure in the Random Forest model. 'B' stands for Sentinel-2 optical bands, while 'lai_cw' stands for canopy water content; 'gndvi' stands for green normalized difference vegetation index; 'mcari' stands for modified chlorophyll absorption ratio index; 'lai' stands for leaf area index; 'fapar' stands for fraction of photosynthetically active radiation and 'gemi' stands for global environment monitoring index (see Supplementary Information Table 1 for more detailed information)

a data-driven adjustment of the a priori defined thresholds in the field (compare Supplementary Fig. 1).

Mapping vegetation structure variables and AFS

Using the Random Forest regression models, we mapped LAI, AGB and CC (see Fig. 6 a, b, c) for the *Las Cruces* micro-watershed and identified the distribution of AFS therein based upon the established thresholds (Fig. 6d). While the classification of AFS *vs.* forest is almost perfect with an overall accuracy of 94% and a kappa of 0.82 (Table 2), the classification taking into account low-shade monocultures (i.e. monocultures *vs.* AFS *vs.* forests) is fair, overestimating the occurrence of monocultures yielding an overall accuracy of 50% and a kappa of 0.25 (Table 3). Hence, we estimate that around two-thirds of *Las Cruces* is used under AFS management (excluding low shade monocultures) with a LAI between 1.10 and 2.25, an AGB between 60 and 139 Mg ha^{-1} and a CC between 50 and 82%. Based on our analysis the CC of forest is always above 82% (Fig. 7; see Supplementary Fig. 1); this aspect is especially interesting, since the Global Forest Change product for 2018 (Hansen et al. 2013, https://earthenginepartners.appspot.com/science-2013-global-forest) recognizes AFS as forests with a tree cover near to 100%, which indicates a overestimation of forest area where there are production systems with a high shade cover.

Discussion

Using field measurements of essential vegetation structure variables in combination with Sentinel-2

are derived from the model predictions within the study area (Fig. 5, pointed line), and can be considered
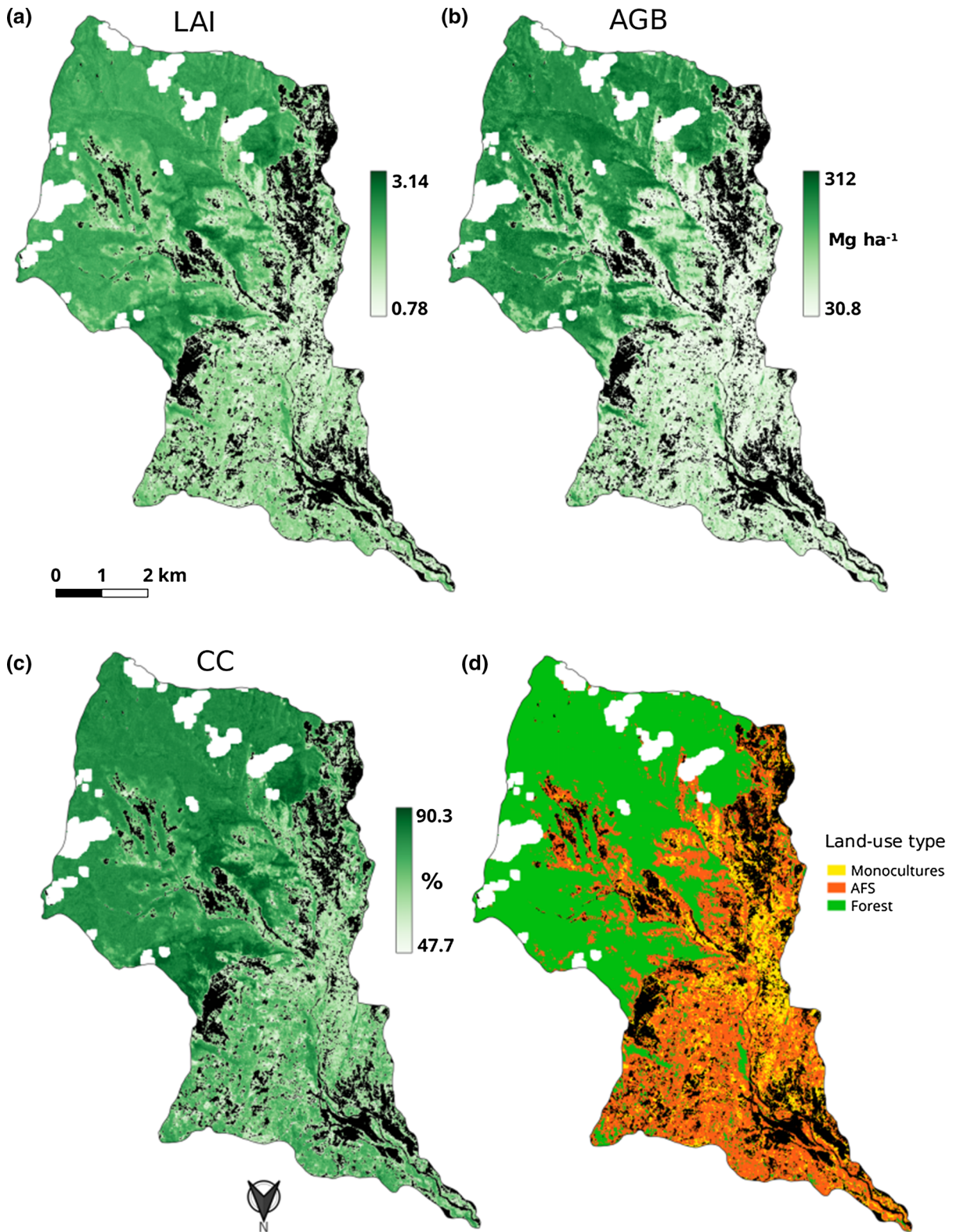
Table 2 Values of overall accuracy and Cohen's Kappa obtained from confusion matrices generated from AFS *vs.* forest

		Reference			Sum
		AFS	Forest	Pasture	
Prediction	AFS	38	1	0	39
	Forest	1	9	0	10
	Pasture	1	0	0	1
	Sum	40	10	0	

Overall accuracy: 94.0%, Kappa coefficient of 0.82

reflectance data, and derived spectral and texture variables, it was possible to predict LAI, AGB and CC for the entire study area. Prediction of the models achieved coefficients of determination (R2) above 0.9 and relative root mean square errors of 19%, 34% and 9% respectively (Fig. 5). These predictions allowed the detection of AFS and an accurate distinction from forests with an accuracy of 94% (see Tables 2, 3). However, also taking into account monocultures planted below few shade trees, resulted in a lower accuracy of 50%. This may be because the spatial information on coffee crops at full exposure (without

Table 3 Values of overall accuracy and Cohen's Kappa obtained from confusion matrices generated from monocultures *vs.* AFS *vs.* forest

		Reference				Sum
		AFS	Forest	Mono-culture	Pasture	
Prediction	AFS	11	1	5	0	17
	Forest	1	9	0	0	10
	Monoculture	17	0	5	0	22
	Pasture	0	0	1	0	1
	Sum	29	10	11	0	

Overall accuracy: 50.0%, Kappa coefficient of 0.25

shade trees) has not been fully considered in the sampling design because of its rarity. Consequently, the presence of few shade trees within a coffee plantation can influence the spectral fingerprint of pixels in these areas causing significant overlap with those of AFS.

Based upon pragmatically defined thresholds, we have developed a novel approach towards detecting AFS from satellite imagery. Whether the thresholds presented here are generally applicable across Colombia has to be demonstrated in future studies, nevertheless they agree with values found in the literature (Nair et al. 2010; Dhyani et al. 2020). For example, Marín et al. (2016) consider AFS with an AGB of 122 Mg ha^{-1} while for Zapata Arango (2019) the higher value of CC found in AFS was 75%. On the other hand, these thresholds can be applied in a flexible manner using LAI, AGB and CC values adjusted to the particular AFS design and management at different study areas. The estimation of canopy variables across landscapes can aid the discrimination between forest and AFS and provide a better understanding of canopy properties in agricultural areas.

Field measurements of AGB revealed a maximum of 405.2 Mg ha^{-1} in a dense cacao AFS, which even surpasses the AGB of some natural forest plots and highlights the great potential of agroforests to store carbon and mitigate climate change (Nair et al. 2010). This result could be due to the presence of *Citrus* fruit trees in the overstory layer of this cacao plot, which increases the amount of AGB because of the massive trunk of *Citrus* fruit trees. Nevertheless, the mean AGB of the forests evaluated in this study was around 200 Mg ha^{-1}, while for AFS it was in average 77 Mg ha^{-1}, which agrees well with previous studies of AFS that report ranges between 12 and 228 Mg ha^{-1} (Albrecht and Kandji 2003).

PCQM has been developed to characterize the structure of mangrove forests and has been used as ground-truth sampling method for remote sensing without the need to establish fixed area plots (Satyanarayana et al. 2011, 2018). It has been successfully implemented in savannas (Satyanarayana et al. 2018) and has also been shown to be useful to characterize the structure of AFS in Andean forests to study the relationship between forest structure and amphibian community composition (Brüning et al. 2018). Together with hemispherical photography and allometric equations developed for tropical AFS (Segura
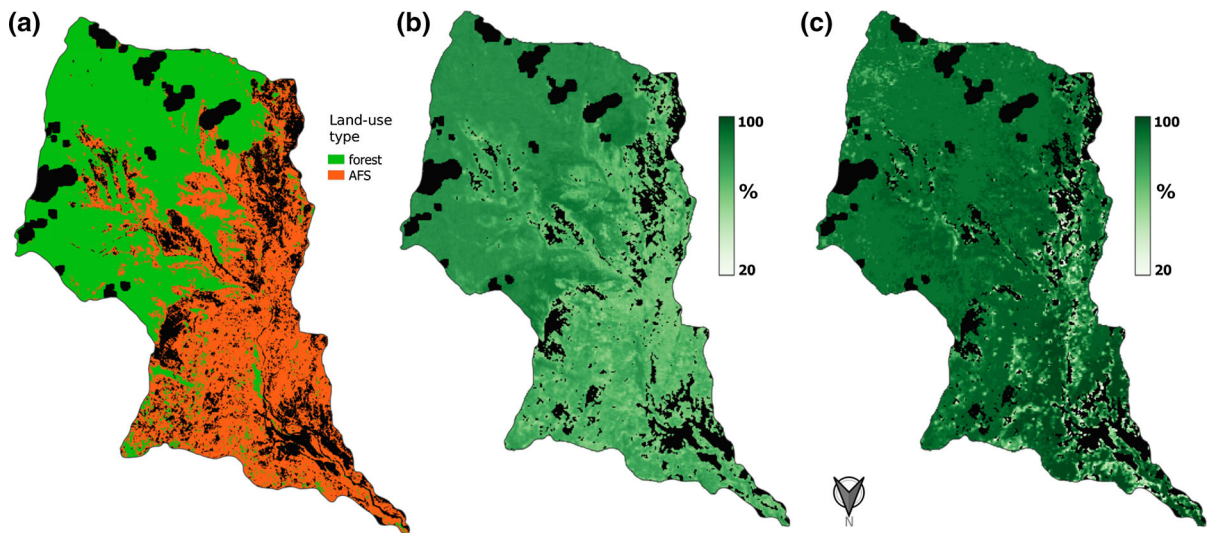
Fig. 7 Map of **a** forest *vs.* AFS, **b** canopy closure prediction from this study, and **c** tree cover derived from the global forest change dataset (Hansen et al. 2013) for the year 2018

et al. 2006; Somarriba et al. 2013), it can be considered a simple, fast and effective method to characterize the structure of AFS, with little sampling effort taking advantage of the planting design of AFS. However, implementing this methodology in other regions where AFS design might be different may require adjustments in the PCQM sampling and the consideration of local allometries which can cause changes in carbon storage (Albrecht and Kandji 2003).

Using thresholds identified from the Random Forest model predictions allowed us to accurately map AFS. Although field measurements of some cacao crops obtained an AGB greater 140 Mg ha^{-1}, the predictions for AGB in AFS plots did not show values over 138 Mg ha^{-1}, while forests did not have values below this limit. Therefore, we established 139 Mg ha^{-1} as the AGB threshold between forest and AFS. Nevertheless, when applying this threshold, some riparian forest that are in recovery after an avalanche caused by the *Las Cruces* river in 2011 are classified as AFS; this highlights that secondary forests that are in early stages of recovery may exhibit similar AGB, LAI and CC values as AFS, which may affect the detection of AFS as presented here and needs to be taken into account. A possible solution might be to incorporate textural features directly in the detection of AFS using a classification approach instead of using thresholds of CC, LAI and AGB.

Spectral information of Sentinel-2 has been shown to accurately estimate stand parameters for Pine monocultures (Hawryło and Węzyk 2018) and AGB in forests (Dang et al. 2019) whereby the reflectance bands B11 (SWIR-1) and B12 (SWIR-2) were important predictor variables. SWIR bands deliver information about crop conditions, as health and moisture and therefore have been useful for vegetation mapping, crop classification and forest monitoring (Zhang et al. 2017; Jadhav and Deshmukh 2019). Some studies in tropical forest also have shown the importance of the bands in the shortwave infrared region for classification and regression using machine learning methods (Zhang et al. 2019; Chen et al. 2019). Here we have confirmed that SWIR-1 and SWIR-2 provide important information for predicting LAI, CC and AGB in Andean AFS (see Supplementary Fig. 2 and Fig. 4). It is interesting that the LAI canopy water content estimated by SNAP (lai_cw; Weiss and Baret 2016) was also useful for the LAI and AGB prediction models. The influence of lai_cw in the AGB and LAI estimation may be due to the higher generation of water vapor by covers that contain higher amounts of photosynthetic material (AFS or forest), related with their biomass content. In general, textural features were of little importance for prediction models although some authors have reported their utility to discriminate land-use types (Laurin et al. 2016). Using

textural features, it has been possible to recognize coffee under shade in eastern Africa (Lelong and Thong-Chane 2003) and estimate AGB in forests when combined with other spectral indices (Lu 2005; Safari and Sohrabi 2016). The higher influence of spectral variables rather than the textural features when both are used as predictors of AGB could reflect a low order in plantation design of the AFS stand structure in the study area (Lu 2006). It can also be caused by a mismatch in scale between field plot area (30 by 30 m) and textural features calculated with a moving window resulting into a size of 50 by 50 m. In order to evaluate the importance of textural features, we rerun Random Forest models only with them. While models generated based on textural features only, resulted in satisfactory predictions of CC, LAI and ABG with R^2 greater than 0.8, these are less well suited for establishing thresholds to detect AFS (see Supplementary Fig. 5).

The high R^2 values obtained between observed and predicted values (Fig. 5), support the idea that a combination of spectral and textural features is an efficient way to estimate AGB using multispectral data (Gao et al. 2018). The excellent model adjustments are likely caused by the local scale of the analysis and the good representation of the management gradient. Probably, more ground sampling at the 'monoculture side' of the management gradient could improve the distinction of AFS with little CC (i.e. monocultures) from AFS and other vegetation types such as early stages of forest regeneration. Because there is no measurement of AGB in pastures or other kinds of low biomass natural vegetation, probably lower values in biomass are overestimated. In order to accurately estimate AGB in AFS and detect this land-use type in the landscape, AGB measurements should be performed in patches of forest regeneration as well to avoid confounding effects. Moreover, detailed information about the composition and wood density of the tree species will improve the AGB estimation through allometries.

While CC, LAI and AGB have been successfully estimated in AFS around the world using multispectral sensors (e.g. Hansen et al. 2013; Dube and Mutanga 2015; Korhonen et al. 2017) and high resolution images (see Taugourdeau et al. 2014), to our knowledge this is the first study doing so in the Andean region. Even more, AFS in Colombia are frequently classified as forests in land-use classifications and hence no quantification or monitoring of their extent exists. Considering the deviation of CC between this study and the global product by Hansen et al. (2013) using Landsat products (Fig. 7), we highlight the need for a refined mapping of AFS in order to better quantify their economic, social and environmental value, especially in developing countries (Garrity et al., 2006).

Conclusions

Above-ground biomass, leaf area index and canopy closure are essential ecosystem variables and associated to important ecosystem processes and services. Here, we have shown that these variables can be accurately estimated and mapped for species-rich Andean forests and AFS in rough terrain and under different management intensities. Moreover, they can be used to detect AFS, which so far, has been a challenge and has not been a major subject of remote sensing studies. Considering that AFS do provide important ecosystem services and are important for the conservation of biodiversity in human dominated tropical landscapes, the presented study opens new avenues for mapping and monitoring the dynamics of AFS using freely available remote sensing imagery. This can improve the planning and decision-making associated with this traditional tropical land use and strengthen its multifunctional role associated with ecosystem service provisioning, climate change mitigation, biodiversity conservation and human well-being.

References

Albrecht A, Kandji ST (2003) Carbon sequestration in tropical agroforestry systems. Agr Ecosyst Environ 99(1–3):15–27

Alvarez E, Duque A, Saldarriaga J, Cabrera K, De G, Lema A, Moreno F, Orrego S, Rodríguez L (2012) Forest Ecology and Management Tree above-ground biomass allometries for carbon stocks estimation in the natural forests of Colombia. For Ecol Manage 267:297–308

Bégué A, Arvor D, Lelong C, Vintrou E, Simoes M (2015) Agricultural systems studies using remote sensing. In: Thenkabail PS (ed) Remote sensing handbook. Land resources: monitoring, modeling, and mapping, vol II. CRC Press, Boca Raton, FL, USA. Taylor and Francis Group, London, UK; New York, NY, USA, pp 113–130

Bhagwat SA, Willis KJ, Birks HJB, Whittaker RJ (2008) Agroforestry: a refuge for tropical biodiversity? Trends Ecol Evol 23(5):261–267

Bivand R, Keitt T, Rowlingson B, Pebesma E, Sumner M, Hijmans R, Rouault E, Bivand MR (2015) Package 'rgdal'. Bindings for the geospatial data abstraction library. https://cran.r-project.org/web/packages/rgdal/index.html. Accessed 12 Sept 2019

Boyd DS, Danson FM (2005) Satellite remote sensing of forest resources: three decades of research development. Prog Phys Geogr 29(1):1–26

Brüning LZ, Krieger M, Meneses-Pelayo E, Eisenhauer N, Pinilla MPR, Reu B, Ernst R (2018) Land-use heterogeneity by small-scale agriculture promotes amphibian diversity in montane agroforestry systems of northeast Colombia. Agr Ecosyst Environ 264:15–23

Calders K, Newnham G, Burt A, Murphy S, Raumonen P, Herold M, Culvenor D, Avitabile V, Disney M, Armston J, Kaasalainen M (2015) Nondestructive estimates of aboveground biomass using terrestrial laser scanning. Methods Ecol Evol 6(2):198–208

Cardinael R, Umulisa V, Toudert A, Olivier A, Bockel L, Bernoux M (2018) Revisiting IPCC Tier 1 coefficients for soil organic and biomass carbon storage in agroforestry systems. Environ Res Lett 13(12):124020

Chen L, Wang Y, Ren C, Zhang B, Wang Z (2019) Optimal combination of predictors and algorithms for forest aboveground biomass mapping from sentinel and SRTM data. Remote Sens 11(4):414

DaMatta FM (2004) Ecophysiological constraints on the production of shaded and unshaded coffee: a review. Field Crops Research 86(2–3):99–114

Dang ATN, Nandy S, Srinet R, Luong NV, Ghosh S, Kumar AS (2019) Forest aboveground biomass estimation using machine learning regression algorithm in Yok Don National Park. Vietnam Ecological Informatics 50:24–32

Dhyani SK, Ram A, Newaj R, Handa AK, Dev I (2020) Agroforestry for carbon sequestration in tropical India. In: Ghosh PK, Mahanta SK, Mandal D, Mandal B, Ramakrishnan S (eds) Carbon management in tropical and subtropical terrestrial systems. Springer, Singapore

Dossa EL, Fernandes ECM, Reid WS, Ezui K (2008) Above-and belowground biomass, nutrient and carbon stocks contrasting an open-grown and a shaded coffee plantation. Agrofor Syst 72(2):103–115

Dube T, Mutanga O (2015) Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in uMgeni catchment, South Africa. ISPRS J Photogramm Remote Sens 101:36–46

Erinjery JJ, Singh M, Kent R (2018) Mapping and assessment of vegetation types in the tropical rainforests of the Western Ghats using multispectral Sentinel-2 and SAR Sentinel-1 satellite imagery. Remote Sens Environ 216:345–354

Fournier RA, Hall RJ (eds) (2017) Hemispherical photography in forest science: theory, methods, applications. Springer, Dordrecht, Netherlands. https://doi.org/10.1007/978-94-024-1098-3

Frazer GW, Canham CD, Lertzman KP (1999) Gap Light Analyzer (GLA), version 2.0: imaging software to extract canopy structure and gap light transmission indices from true colour fisheye photographs, users manual and program documentation. Simon Fraser University, Burnaby, British Columbia, and the Institute of Ecosystem Studies, Millbrook

Garrigues S, Shabanov NV, Swanson K, Morisette JT, Baret F, Myneni RB (2008) Intercomparison and sensitivity analysis of Leaf Area Index retrievals from LAI-2000, Accu-PAR, and digital hemispherical photography over croplands. Agric For Meteorol 148(8–9):1193–1209

Garrity DP (2004) Agroforestry and the achievement of the Millennium Development Goals. Agrofor Syst 61(1–3):5–17

Garrity D, Okono A, Grayson M, Parrot S (2006) World agroforestry into the future. World Agroforestry Centre, Nairobi

Gao Y, Lu D, Li G, Wang G, Chen Q, Liu L, Li D (2018) Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. Remote Sens 10(4):627

Gomez C, Mangeas M, Petit M, Corbane C, Hamon P, Hamon S, De Kochko A, Le Pierres D, Despinoy M (2010) Use of high-resolution satellite imagery in an integrated model to predict the distribution of shade coffee tree hybrid zones. Remote Sens Environ 114(11):2731–2744

Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Kommareddy A (2013) High-resolution global maps of 21stcentury forest cover change. Science 342(6160):850–853

Haralick RM, Shanmugam K (1973) Textural features for image classification. IEEE Transact Syst man Cybern 6:610–621

Hawryło P, Wężyk P (2018) Predicting growing stock volume of scots pine stands using Sentinel-2 satellite imagery and airborne image-derived point clouds. Forests 9(5):274

Hijmans RJ, Van Etten J (2012) raster: Geographic analysis and modeling with raster data. R package version 2.0–05. http://CRAN.R-project.org/package=raster

IDEAM (2010) Leyenda Nacional de Coberturas de la Tierra. Metodología CORINE Land Cover adaptada para Colombia Escala 1:100.000. Instituto de Hidrología, Meteorología y Estudios Ambientales. Bogotá, D. C., p 72

Isaac ME, Timmer VR, Quashie-Sam SJ (2007) Shade tree effects in an 8-year-old cocoa agroforestry system: biomass and nutrient diagnosis of Theobroma cocoa by vector analysis. Nutr Cycl Agroecosyst 78(2):155–165

Ishii HT, Tanabe SI, Hiura T (2004) Exploring the relationships among canopy structure, stand productivity, and biodiversity of temperate forest ecosystems. Forest Science 50(3):342–355

Jadhav PP, Deshmukh VB (2019) Optimum band selection in sentinel-2A satellite for crop classification using machine learning technique. Int Res J Eng Technol 6(4):1619–1625

Jafari SM, Zarre S, Alavipanah SK (2013) Woody species diversity and forest structure from lowland to montane forest in Hyrcanian forest ecoregion. J Mt Sci 10(4):609–620

Jose S (2012) Agroforestry for conserving and enhancing biodiversity. Agrofor Syst 85(1):1–8

Karlson M, Ostwald M, Bayala J, Bazié HR, Ouedraogo AS, Soro B, Sanou J, Reese H (2020) The potential of Sentinel-2 for crop production estimation in a smallholder agroforestry landscape. Burkina Faso Front Environ Sci 8:85

Klein AM, Steffan-Dewenter I, Buchori D, Tscharntke T (2002) Effects of land-use intensity in tropical agroforestry systems on coffee flower-visiting and trap-nesting bees and wasps. Conserv Biol 16(4):1003–1014

Klein AM, Steffan-Dewenter I, Tscharntke T (2006) Rain forest promotes trophic interactions and diversity of trap-nesting Hymenoptera in adjacent agroforestry. J Anim Ecol 75(2):315–323

Korhonen L, Ali-Sisto D, Tokola T (2015) Tropical forest canopy cover estimation using satellite imagery and airborne lidar reference data. Silva Fennica 49(5):1–18

Korhonen L, Packalen P, Rautiainen M (2017) Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. Remote Sens Environ 195:259–274

Kuyah S, Öborn I, Jonsson M (2017) Regulating ecosystem services delivered in agroforestry systems. In: Dagar JC, Tewari VP (eds) Agroforestry. Springer, Singapore

Laurin GV, Puletti N, Hawthorne W, Liesenberg V, Corona P, Papale D, Chen Q, Valentini R (2016) Discrimination of tropical forest types, dominant species, and mapping of functional guilds by hyperspectral and simulated multispectral Sentinel-2 data. Remote Sens Environ 176:163–176

Leimona B, Noordwijk MV (2017) Smallholder agroforestry for sustainable development goals: ecosystem services and food security. Palawija Newsletter 34(1):1–6

Lelong C, Dupuy S, Alexandre C (2014) Discrimination of tropical agroforestry systems in very high resolution satellite imagery using object-based hierarchical classification: a case-study on cocoa in Cameroon. South-East Eur J Earth Obs Geom 3:255–258

Lelong C, Thong-Chane A (2003) Application of textural analysis on very high resolution panchromatic images to map coffee orchards in Uganda. In IGARSS 2003. Paper presented at IEEE international geoscience and remote sensing symposium. IEEE Proceedings, IEEE Cat. No. 03CH37477. Vol. 2, pp. 1007–1009

Liaw A, Wiener M (2002) Classification and regression by random forest. R news 2(3):18–22

Lu D (2005) Aboveground biomass estimation using landsat TM data in the Brazilian Amazon. Int J Remote Sens 26(12):2509–2525

Manduell KL, Harrison ME, Thorpe SK (2012) Forest structure and support availability influence orangutan locomotion in Sumatra and Borneo. Am J Primatol 74(12):1128–1142

Marín MP, Andrade H, Sandoval A (2016) Fijación de carbono atmosférico en la biomasa total de sistemas de producción de cacao en el departamento del Tolima, Colombia. Revista UDCA Actualidad y Divulgación Científica 19(2):351–360

Martone M, Rizzoli P, Wecklich C, González C, Bueso-Bello JL, Valdo P, Schulze D, Zink M, Krieger G, Moreira A (2018) The global forest/non-forest map from TanDEM-X interferometric SAR data. Remote Sens Environ 205:352–373

Mbow HOP, Reisinger A, Canadell J, O'Brien P (2017) Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (SR2). Ginevra, IPCC

Mitchell K (2015) Quantitative analysis by the point-centered quarter method. ArXiv Preprint, arXiv 1010:1–56

Morin D, Planells M, Guyon D, Villard L, Mermoz S, Bouvet A, Thevenon H, Dejoux J-F, Le Toan T, Dedieu G (2019) Estimation and mapping of forest structure parameters from open access satellite images: development of a generic method with a study case on coniferous plantation. Remote Sens 11(11):1275

Nair PKR, Nair VD, Kumar BM, Showalter JM (2010) Carbon sequestration in agroforestry systems. Adv Agron 108:237–307

Nair PR (1985) Classification of agroforestry systems. Agrofor Syst 3(2):97–128

Orozco GV, Espinosa CMO, Salazar JCS, Pantoja CFL (2015) Almacenamiento de carbono en arreglos agroforestales asociados con café (Coffea arabica) en el sur de Colombia. Revista de Investigación Agraria y Ambiental (RIAA) 5(1):213–221

Pinilla MC, Rueda AJ, Pinzón CA (2018) Métodos para el monitoreo agroclimático alrededor de embalses: estudio de caso para la hidroeléctrica Sogamoso, Santander. Colombia, Fundación Natura, p 76

Porras INA, Vorley B, Amrein A, Douma W, Clemens H (2015) Payments for ecosystem services in smallholder agriculture: lessons from the Hivos-IIED learning trajectory. IIED and Hivos

QGIS Development Team., 2015. QGIS geographic information system. Open Source Geospatial Foundation Project, Versão. Vol. 2. No. 7

R Development Core Team, R. F. F. S. C., 2011. R: A language and environment for statistical computing

Safari A, Sohrabi H (2016) Ability of Landsat-8 OLI derived texture metrics in estimating aboveground carbon stocks of coppice oak forests. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, p 41

Satyanarayana B, Mohamad KA, Idris IF, Husain ML, Dahdouh-Guebas F (2011) Assessment of mangrove vegetation based on remote sensing and ground-truth measurements at Tumpat, Kelantan Delta, East Coast of Peninsular Malaysia. Int J Remote Sens 32(6):1635–1650

Satyanarayana B, Muslim AM, Horsali NAI, Zauki NAM, Otero V, Nadzri MI, Ibrahim S, Husain M-L, Dahdouh-Guebas F

(2018) Status of the undisturbed mangroves at Brunei Bay, East Malaysia: a preliminary assessment based on remote sensing and ground-truth observations. PeerJ 6:e4397

Segura M, Kanninen M, Suárez D (2006) Allometric models for estimating aboveground biomass of shade trees and coffee bushes grown together. Agrofor Syst 68(2):143–150

Seidel D, Fleck S, Leuschner C (2012) Analyzing forest canopies with ground-based laser scanning: a comparison with hemispherical photography. Agric For Meteorol 154:1–8

Sharma G, Hunsdorfer B, Singh KK (2016) Comparative analysis on the socio-ecological and economic potentials of traditional agroforestry systems in the Sikkim Himalaya. Tropical Ecology 57(4):751–764

Shukla PR, Skea J, Calvo Buendia E, Masson-Delmotte V, Pörtner H-O, Roberts DC, Zhai P, Slade R, Connors S, van Diemen R, Ferrat M, Haughey E, Luz S, Neogi S, Pathak M, Petzold J, Portugal Pereira J, Vyas P, Huntley E, Kissick K, Belkacemi M, Malley J (2019) Climate change and land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. IPCC, Geneva, Switzerland

Scheper AC (2019) The potential of coffee agroforestry systems to enhance crop productivity, pest control, carbon sequestration and biodiversity: Evidence from theEje Cafetero, Colombia. Geosciences, Utrecht University Repository. Masters thesis. https://dspace.library.uu.nl/handle/1874/384704.

SNAP - ESA Sentinel Application Platform v6.0 [Computer software]. (2019). Retrieved from http://step.esa.int. Accessed 2019

Somarriba E, Beer J, Orihuela JA, Andrade HJ, Cerda R, Declerck F, Detlefsen G, Escalante M, Giraldo LA, Ibrahim MA, Krishnamurthy L, Mena VE, Mora JR, Orozco L, Scheelje M, Campos JJ (2012) Mainstreaming agroforestry in Latin America. In: Nair PKR, Garrity D (eds) Agroforestry-the future of global land use. Springer, Berlin

Somarriba E, Cerda R, Orozco L, Cifuentes M, Dávila H, Espin T, Mavisoy H, Avila G, Alvarado E, Poveda V, Astorga C, Say E, Deheuvels O (2013) Carbon stocks and cocoa yields in agroforestry systems of Central America. Agr Ecosyst Environ 173:46–57

Taugourdeau S, Le Maire G, Avelino J, Jones JR, Ramirez LG, Quesada MJ, Charbonnier F, Gómez-Delgado F, Harmand J, Rapidel B, Vaast P, Roupsard O (2014) Leaf area index as an indicator of ecosystem services and management

practices: An application for coffee agroforestry. Agr Ecosyst Environ 192:19–37

Tropek R, Sedláček O, Beck J, Keil P, Musilová Z, Šímová I, Storch D (2014) Comment on "high-resolution global maps of 21st-century forest cover change." Science 344(6187):981–981

TSITSI, B. (2016) Remote sensing of aboveground forest biomass: a review. Tropical Ecology 57(2):125–132

Waldron A, Garrity D, Malhi Y, Girardin C, Miller DC, Seddon N (2017) Agroforestry can enhance food security while meeting other sustainable development goals. Trop Conserv Sci 10:1–6

Warde W, Petranka JW (1981) A correction factor table for missing point-center quarter data. Ecology 62(2):491–494

Watch GF (2002) Global forest watch. World Resources Institute, Washington, DC. http://www.globalforestwatch.org. Accessed March 2019

Weiss M, Baret F (2016) S2ToolBox Level 2 products: LAI, FAPAR. FCOVER, Institut National de la Recherche Agronomique (INRA), Avignon

Yapo T (2019) How implementing agroforestry in plantations can help côte d'ivoire achieve its sustainable development goals. https://www.un-redd.org/single-post/2019/05/17/How-Implementing-Agroforestry-in-Plantations-Can-Help-Côte-dIvoire-Achieve-its-Sustainable-Development-Goals. Accesed 23 Nov 2019

Zapata Arango PC (2019) Composición y estructura del dosel de sombra en sistemas agroforestales con café de tres municipios de Cundinamarca. Colombia Ciência Florestal 29(2):685–697

Zhang T, Su J, Liu C, Chen WH, Liu H, Liu G (2017) Band selection in Sentinel-2 satellite for agriculture applications. Paper presented at 23rd International Conference on Automation and Computing (ICAC). IEEE. pp 1–6

Zhang TX, Su JY, Liu CJ, Chen WH (2019) Potential bands of sentinel-2A satellite for classification problems in precision agriculture. Int J Autom Comput 16(1):16–26

Zomer RJ, Bossio DA, Trabucco A, Yuanjie L, Gupta DC, Singh VP (2007) Trees and water: smallholder agroforestry on irrigated lands in Northern India. IWMI. Vol. 122