

Article

Does Sentinel-1A Backscatter Capture the Spatial Variability in Canopy Gaps of Tropical Agroforests? A Proof-of-Concept in Cocoa Landscapes in Cameroon

Frederick N. Numbisi ^{1,2,*}  and Frieke Van Coillie ¹ 

¹ Remote Sensing | Spatial Analysis Lab (REMOSA), Department of Environment, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Gent, Belgium; Frieke.Vancoillie@UGent.be

² West and Central Africa Regional Office, World Agroforestry Center (ICRAF), P.O. Box 16317 Yaoundé, Cameroon

* Correspondence: frederick.nkeumoenumbisi@ugent.be; Tel.: +49-178-310-7190

Received: 14 November 2020; Accepted: 15 December 2020; Published: 19 December 2020



Abstract: A reliable estimation and monitoring of tree canopy cover or shade distribution is essential for a sustainable cocoa production via agroforestry systems. Remote sensing (RS) data offer great potential in retrieving and monitoring vegetation status at landscape scales. However, parallel advancements in image processing and analysis are required to appropriately use such data for different targeted applications. This study assessed the potential of Sentinel-1A (S-1A) C-band synthetic aperture radar (SAR) backscatter in estimating canopy cover variability in cocoa agroforestry landscapes. We investigated two landscapes, in Center and South Cameroon, which differ in predominant vegetation: forest-savannah transition and forest landscape, respectively. We estimated canopy cover using in-situ digital hemispherical photographs (DHPs) measures of gap fraction, verified the relationship with SAR backscatter intensity and assessed predictions based on three machine learning approaches: multivariate bootstrap regression, neural networks regression, and random forest regression. Our results showed that about 30% of the variance in canopy gap fraction in the cocoa production landscapes was shared by the used SAR backscatter parameters: a combination of S-1A backscatter intensity, backscatter coefficients, difference, cross ratios, and normalized ratios. Based on the model predictions, the VV (co-polarization) backscatter showed high importance in estimating canopy gap fraction; the VH (cross-polarized) backscatter was less sensitive to the estimated canopy gap. We observed that a combination of different backscatter variables was more reliable at predicting the canopy gap variability in the considered type of vegetation in this study—agroforests. Semi-variogram analysis of canopy gap fraction at the landscape scale revealed higher spatial clustering of canopy gap, based on spatial correlation, at a distance range of 18.95 m in the vegetation transition landscape, compared to a 51.12 m spatial correlation range in the forest landscape. We provide new insight on the spatial variability of canopy gaps in the cocoa landscapes which may be essential for predicting impacts of changing and extreme (drought) weather conditions on farm management and productivity. Our results contribute a proof-of-concept in using current and future SAR images to support management tools or strategies on tree inventorying and decisions regarding incentives for shade tree retention and planting in cocoa landscapes.

Keywords: Sentinel-1; SAR; cocoa agroforests; digital hemispherical photographs (DHPs); canopy gap; spatial correlation; random forest; neural network; machine learning; forest structure

1. Introduction

The spatial variability in canopy cover over forest land is a vital information for monitoring and assessing the influence of forest management practices on ecosystem services (ESs). Ecosystem services are produced by different components or functions of forests and tree-based land uses as cocoa agroforests [1–3]: habitat, tree diversity, canopy cover, etc. Canopy cover refers to the proportion of the forest floor that is covered by the vertical (or map) projection of the tree crowns; it is different from canopy closure—the proportion of sky hemisphere obscured by vegetation when viewed from a single ground point [4]. Such distinction may be a necessary guide to field sampling and management of related ESs in different vegetation types and land uses. Cocoa agroforestry systems (CAFS) feature canopy cover and tree height that are similar to that in transition forest types [5]. The widely practiced canopy cover (or shade) typologies in CAFS include: (1) cocoa with production shade tree species, (2) cocoa with mixed shade species, (3) rustic cocoa or ‘cabruca’, and (4) successional agroforests. These agroforests constitute mostly canopy shade trees (CST) above 5 m tall, which provide a stratification of above 10% canopy cover [6]. The associated shade trees provide the different canopy strata in CAFS. The CSTs provide both shade and non-shade uses. The known range of ecosystem provision and supporting services of associated shade trees qualify CAFS for climate change mitigation mechanisms, such as REDD⁺ [3]. Cocoa agroforestry land use is, yet, excluded from the international definition of forests [7]. However, there are associated disservices to upper canopy tree species (UCT), as well: host to disease vectors, competition for soil water and nutrient, etc. Thus, shade management is still an issue of contention between farmers, and the shade densities often vary with site characteristics, such as landscape structure, land use history, vegetation type and distribution, etc. Besides the technical difficulties related to tree planting in CAFSs, farmers are unable to measure and maintain the, essentially theoretical, recommended shade proportion on farms [8,9].

Beside the challenges of shade management, climate change impact projections have not been readily established for agroforestry systems due to their complex structure [10]. Fluctuations in weather pattern often result in changes in micro-climate. For CAFS, the micro-climate is influenced by the density and distribution of associated shade trees [11]. However, current weather fluctuations and the projected extreme dry seasons herald modifications in hitherto standard farm management, for example, the type and density of shade tree species, spatial configurations for cocoa trees in CAFS, as well, changes in cocoa production sites [12,13]. Moreover, obstructions in forest cover due to deforestation and corresponding fluctuations in leaf area drive changes in micro-climates [14]. Thus, estimates of canopy cover, at landscape level, may provide supporting information to monitor spatial changes in micro-climate; such estimates may also be useful for predicting drought- or disease-related risks on cocoa farm productivity.

The assessment of canopy cover in forests has been widely done using spherical densimeters and digital hemispherical photos [15]. However, such methods are beyond the skills of the predominantly rural small-scale cocoa farmers; though they are meant to be used by farmers. Some scholars use tree stock density and parameters of upper canopy trees (UCTs) (such as crown length) as a measure of shade level [16,17]. However, such an approach requires large-scale field inventory, which is demanding in time and human resources, as well as difficult under inaccessible field conditions. Moreover, such measurements are impractical in certain farm conditions, for instance, in farms with overlapping canopy of the shade trees. Thus, point estimation methods are little reliable for application at both farm and landscape scales. Considering the contiguous effects of changes in micro-climate across forest or farms units, cocoa-producing farmers and countries need support for managing shade conditions beyond the farm scale. For instance, it may be logical to assess shade conditions at landscape scales since the amount and distribution of shade often vary within and between CAFS. Moreover, it is impractical to devise and implement a sustainable shade management plan for each farm. Thus, a management plan for shade trees may be suggested for each cocoa production landscapes.

Several tools are used for monitoring and managing canopy cover in forest and tree-based land cover: Light Detection and Ranging (LiDAR) approach [18,19], information from aerial photographs [20],

Unmanned Aerial Vehicle (UAV) images [21], and Digital Hemispherical Photographs (DHPs). DHPs have known potentials, amongst other applications, in providing timely, up-to-date, and large-scale monitoring of changes in vegetation structure. For instance, the leaf area index (LAI) for subtropical forests has been reliably estimated using DHP [22]. Different remote sensing methods are applicable to estimate LAI [23–25]. The effectiveness of DHP in estimating LAI has also been reported for deciduous forests [26] and for rubber plantations by comparing data from Worldview-2 imagery [27]. Although synthetic aperture radar (SAR) imagery has been widely used for canopy cover estimation and biomass monitoring, such remarkable applications have been extensive for temperate forests [28–33], and few studies also assess canopy cover in tropical forests from optical satellite data [34–36]. Such information, however, is currently non-existent for cocoa agroforestry systems.

Unlike remote sensing images that are acquired using optical sensors, the backscatter intensity of SAR images only provides indirect information about canopy structure. In the case of dual polarization SAR imagery, cross-polarization backscatter (Vertical transmitted (VH) and Horizontal received (HV)) can distinguish the structural parameters of different crops types [31]. Moreover, the biophysical attributes of forests, such as canopy height, can be reliably inferred from SAR imagery [37,38]. However, the SAR backscatter intensity over most forest types is complex due to the often non-distinctive structural attributes and complex composition of trees. The attribute of vegetation that relates with SAR backscatter intensity also depends on sensor properties, such as the wavelength, polarization, and incidence angle of the transmitted (and received) radar signal [39–41]. Besides sensor properties, backscatter intensity depends on structural properties of vegetation: moisture content of leaves and branches, the size and orientation of branches and leaves, the spatial configuration of trees in a resolution cells, and the shape and size of tree canopies (leaf-on versus leaf-off). These structural properties, assessed across time and season, can be used to map forest clear cut and differentiate crop fields [31,33]. However, the sensitivity and potentials of SAR backscatter have not been assessed in relation to tree-crop production systems with multi-strata canopy, such as cocoa agroforestry. In the context of measuring the structural parameters in high vegetation canopy, Reference [4] distinguished measures of canopy cover/shade. According to Reference [4], DHP are random point measures that integrate information from different heights of the canopy. Thus, it provides a good resolution and an efficient bias estimator of the canopy cover, actually canopy closure. In this experiment, we estimated canopy gap fraction from field sampling by DHP of overstorey canopy.

In the framework of large-scale shade management for cocoa tree-crops, and the prediction weather-induced impacts on farm productivity, this study evaluates the potential of S-1A SAR data for detecting canopy gaps in CAFS and cocoa production landscapes. We consider two cocoa production landscape in the Center and South regions of Cameroon. The landscapes differ in the prevailing weather pattern, predominant vegetation type, and cocoa agroforests structure and practices [42,43]. In both landscapes, we collected field reference data in CAFSs of a varying range of tree density, visually assessed canopy cover, farm ages, and observed management of canopy trees. In a first analysis, we assessed the relationship between canopy gap fraction and S-1A backscatter intensity from both VH (cross-polarization) and VV (co-polarization) polarization images; canopy gap fraction was estimated from in-situ DHPs. We assessed machine learning (ML) approaches for their potential to delineate different canopy gaps fractions for the sampled cocoa agroforests parcels. We hypothesized that C-band SAR backscatter intensity may reflect variability in the intensity and distribution of canopy gaps in cocoa agroforests, and this may be informative for making predictions at the landscape scale. We aimed to evaluate the spatial distribution in canopy gaps as a leverage for monitoring changes in shade cover and climate change impacts in cocoa production landscapes.

2. Materials and Methods

2.1. Study Sites

The first study site is found in the Bokito district of the Mbam & Inoubou Division, which is located in the northern part of the Center Region of Cameroon. This division covers a total surface area of about 6821 km², and it is located between latitudes 4°39'–4°49' N and longitudes 11°4'–11°19' E, within the Universal Transverse Mercator zone (UTM) Zone 32N. The Bokito district has a topography of mostly lowlands with altitudes ranging from 338 to 1109 m a.s.l (Figure 1c). The district is within the agro-ecological zone with humid forest and bi-modal rainfall pattern, and the climate is characterized by four seasons: with a small rainy season from March to June and a large rainy season from August to November. The vegetation is mainly a forest-savannah transition zone that is currently covered by both savannah, secondary forests, and gallery forest patches. The Bokito locality is situated about 150 km from Yaoundé. In the landscape, establishments of cocoa-based agroforests are widespread, prolonged, and still frequent [42,44,45]. Among the major cocoa production zones in Cameroon, this administrative division is characterized by relatively low annual precipitation levels (annual total rainfall of about 1400 mm in Bokito [46]), but it is recognized for producing good quality cocoa. In the center region of Cameroon, the natural vegetation is, however, heavily degraded by pressure on land tenure due to local high population density that can be greater than 100 inhab. km⁻². Bokito is located in the zone with a relatively low population density (29 inhab. km⁻²) compared to other parts of the region.

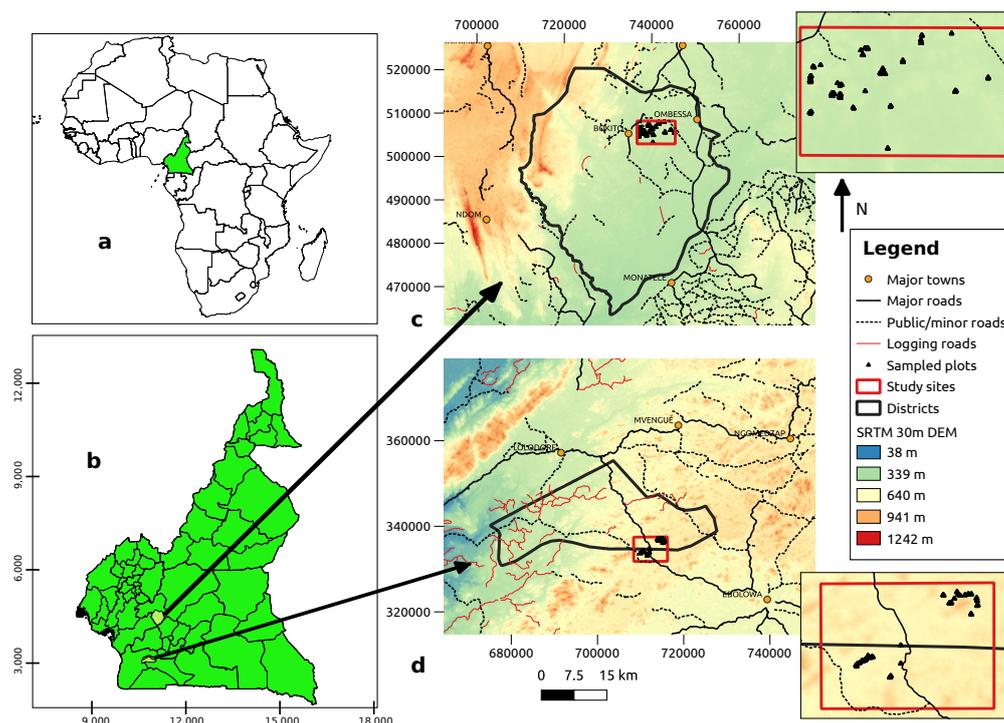


Figure 1. Geographic location of Cameroon and the study sites: (a) Cameroon in Africa. (b) Districts of the study sites in the Mbam & Inoubou Division of the Center Region and the Mvila Division of the South Region (Map Coordinate Reference System (CRS) is World Geodetic System (WGS) 84, Degree Decimal). (c) The forest-savannah transition landscape (study site 1) in the Bokito district (CRS: WGS 84/Universal Transverse Mercator zone (UTM)). (d) The forest landscape (study site 2) in the Efoulan district (CRS: WGS 84/UTM). The image inserts in (c,d) show sampling locations in each site. (Other data source: World Resources Institute (WRI) Interactive forestry atlas of Cameroon 2013, Shuttle Radar Topography Mission (SRTM) 1 Arc-Second 30 m Global Digital Elevation Model.)

The second site is found in Efoulan district that is located between 2°55' to 3°12' N and 10°35' to 11°2' E, within the UTM Zone 32N, in the Mvila Division of the South Region of Cameroon. The topography belongs to the continental plateau of southern Cameroon, with undulating and rolling hills that alternate with some small rivers and widely distributed swamps [47]. The district has an altitude range between 215 to 900 m a.s.l (Figure 1d). Located also within the aforementioned agro-ecological zone, the climate is characterized by two rainy seasons (March–June and September–November) and two dry seasons between these. The average annual rainfall is around 2000 mm. The vegetation cover is predominantly secondary forests, remnant primary forest areas, and agricultural lands. Farming practices are mostly extensive via shifting cultivation—slash and burn agriculture—and cocoa agroforestry expansion into forest areas [9,48]. The Efoulan locality, the headquarter of the district, is situated 32 km from Ebolowa and 43 km to Lolodolf (main towns) in the tropical humid forest zone. The administrative boundaries of Efoulan council cover about 830 km² [43]. According to the census data in 2005, the total population of about 28,000 inhabitants distributed in about 37 villages. Settlements are mainly distributed along the main roads within the district. The population is rural and depends on agricultural and forest exploitation activities.

2.2. Assessment of Canopy Cover

To estimate canopy cover (or conversely, gap fraction—a measure of the proportion of gaps within tree canopies) in cocoa agroforests, a total of 134 farm plots with size approximating 20 × 20 m² were sampled in 2016 and 2017 for collection of DHPs of the canopy (Table 1). We used a hand-held Garmin 64s Global Positioning System (GPS) tool to locate plot boundaries and DHP collection points. It is equipped with a space-based global navigation satellite system (GLONASS) receiver, an alternative to a Global Navigation Satellite System (GNSS) receiver, which provides reliable navigation and location data. Thus, based on the manufacture specifications, location accuracy is within a range of approximately 3 m. During sampling, we placed the GPS in a stable position for a necessary duration to get stable GPS coordinate readings. And, considering potential distortions in the accuracy of ground control points (GCPs), while extracting baskscatter data, we used a image masks (50 × 50 m²) that were large enough to include the boundaries of representative sample plots (20 × 20 m²) and the hemispherical camera viewpoints. Upward-looking DHP were taken at a height of 1 m using a Nikon D7000 DSLR camera with a fisheye objective lens (Figure 2a). The photos were taken in a sequence of thirteen points in each farm, and this was later reduced to five sampling locations as latter had no significant RMSE of gap fraction compared to the thirteen points. The adapted DHP sampling scheme (Figure S1) and camera settings [49–51] are detailed in the supplementary materials. DHPs were processed using the free and open software, ImageJ, in three steps [22]. First, pre-processing entailed a batch process for cropping the circular canopy portion of each image and extracting the different RGB image channels: red, green, and blue. The latter were used for further analysis as these offer the highest contrast between sky and vegetation [51]. Secondly, the cropped blue channel images were further subject to binarization into black/white image of two pixel classes: sky and non-sky pixels. The binary image was created using a threshold level determined by analyzing the histogram of the blue channel image. We applied an automated segmentation procedure: the vegetation and sky pixels were separated by applying the “Minimum” thresholding algorithm to the cropped circular images [52]. Lastly, canopy gap fraction (a proxy for canopy cover) was estimated as the proportion of sky or white pixel in the circular cropped image [4,52] (Figure 2b and Table 2).



Figure 2. Protocol of canopy gap fraction estimation from DHPs: (a) DSLR camera and fisheye lens setup in sample point. (b) DHP processing chain.

Table 1. DHP inventory dates, the number of sampled plots (cocoa agroforestry systems (CAFS) and transition forests), and used Sentinel-1A (S-1A) synthetic aperture radar (SAR) images to test correlation with canopy cover. B: Bokito site, E: Efoulan site.

Site	Dates of DHPs Inventory	No. Sampled Plots (20 m × 20 m)	Acquisition Dates of Matched S-1A SAR Images
B	28 May 2016, 29 May 2016, 01 June 2016, 03 June 2016, 05 June 2016	17	31 May 2016
B	07 June 2016, 10 June 2016	5	12 June 2016
B	16 July 2016, 17 July 2016, 19 July 2016, 20 July 2016	6	18 July 2016
B	28 October 2017, 29 October 2017, 01 November 2017, 02 November 2017, 03 November 2017, 04 November 2017	34	29 October 2017
B	05 November 2017, 06 November 2017, 07 November 2017, 08 November 2017, 09 November 2017	30	10 November 2017
E	19 August 2016, 20 August 2016, 22 August 2016	6	16 August 2016
E	23 August 2016, 25 August 2016, 27 August 2016	7	28 August 2016
E	22 November 2017, 24 November 2017, 25 November 2017, 27 November 2017, 28 November 2017, 1 December 2017, 02 December 2017	29	27 November 2017

Table 2. Sample estimate of gap fraction at 5 points in a plot using a cropped circular area $\approx 3,378,360$ pixels.

Index	Photo	No of Gaps	Gap Area (Pixels)	Gap Fraction (%)
1	FRD_0856.jpg	3566	387,927	11.48
4	FRD_0859.jpg	4373	414,596	12.27
7	FRD_0862.jpg	4896	450,799	13.34
10	FRD_0865.jpg	2481	267,422	7.91
13	FRD_0868.jpg	2902	337,304	9.98

2.3. Relating SAR Backscatter to Canopy Gap Fraction

To match SAR backscatter with the in-situ field measurements of canopy cover, we identified multi-seasonal S-1A SAR images in the range of ± 0 to 6 days from the date of DHPs inventory. We used images that were captured using the Interferometric Wide-swath (IW) mode, which have been geo-referenced and provided by the European Space Agency (ESA) as Level-1 Ground Range Detected (GRD) image products with a spatial resolution of 10 m. The different SAR data, DHP collection dates, and corresponding number of sampled plots are shown in Table 1. We use the term “sample plots” to refer to different mapped areas of about 20 m \times 20 m; in each plot, DHPs of the over-storey canopy were taken and processed for estimating in-situ canopy gap fraction (Figure 2). Thus, the number of sampled plots for each row constituted a total of plots that were sampled in a particular farm and also across several different locations during each of the DHPs inventory dates (Figure 1).

We conducted a time-series analysis of available S-1A SAR images over the sampled locations in the study landscapes. Time-series analysis was conducted in Google Earth Engine (GEE), and the backscatter intensity (sigma nought) were aggregated by the sampled locations (merged pixel masks) in each landscape and year. The distribution of mean VV and VH backscatter did not show a strong temporal variability over the sampled locations, especially during the period of field sampling and canopy gap measurements (Figures S2 and S3). For available SAR images during the window period of field sampling, we compared the mean backscatter between image date based on a non-parametric Kruskal-Wallis sum rank test at alpha 0.05; the Bonferroni correction was applied to adjust alpha value against type I error rate (Table S1). We observed significant differences in mean backscatter between a few image dates though the differences between dates are not consistent for VV and VH backscatter. This mild temporal differences in backscatter intensity over the sampled field locations informed our selection of SAR images, matching with in-situ DHP sampling dates, and extraction of backscatter intensity for subsequent analyses.

Each of the used SAR images subset of the study areas (Table 1) were pre-processed using a workflow described in Reference [53]. In this study, though, we applied a Radiometric Terrain Correction (RTC) of the amplitude bands to Gamma0 intensity [32,54] with a spatial resolution of 10 m. Image processing was conducted in the the Sentinel Application Platform (SNAP) toolbox [55,56]. According to Reference [54], the radiometric terrain flattening to Gamma0 makes use of the local illumination area of the sensor and reduces image acquisition bias due to incidence angle. Figure 3 illustrates SAR backscatter intensity, by RGB image composites, for the sample landscapes in the wet and dry seasons.

Thus, compared to geometric terrain correction (GTC) based on incidence angle, the RTC corrects for local topography and provides easier interpretation of backscatter statistics [32,54]. The Gamma0 VV and VH polarization intensity bands of each image were calibrated to logarithmic scale (decibel, dB). The logarithm is applied [57] to normalize the backscatter intensity and improve contrast between different ground features.

Several studies exhibit the relationship between SAR backscatter and vegetation canopy parameters. To complement the assessment of vegetation structure, such as the normalized vegetation index (NDVI), radar vegetation index (RVI) and other backscatter ratios have been applied, with different objectives, to classify L- and C-band SAR images [31,57–60]. RVI from L-band SAR backscatter has a coupled correlation with both vegetation water content and LAI (plant structure) [60]. According to Reference [31], the cross ratio (VH/VV) of Sentinel-1 (S-1) backscatter is more sensitive to vegetation water content. In addition, for crop fields, the S-1 backscatter strength decreases with increasing NDVI of cereals, and, unlike the VV polarization, vegetation scattering (stronger in VH polarization) increases with increasing LAI [57]. Ref. [57] used different co- and cross-polarization bands to assess canopy attributes of cereals from S-1 SAR. Ref. [56] also observed that the sensitivity of S-1 backscatter parameters to crop height and canopy cover depend on crop type. The authors used several backscatter ratios, arithmetic differences of backscatter band polarizations, and normalized difference. However, these cited studies focused on assessing vegetation structure and canopy cover of temperate and low

canopy crops, such as cereals. Seasonal fluctuations in S-1A SAR backscatter were observed for tropical forests and agroforests [53]. Thus, though not as remarkable as in temperate forest conditions, one may expect backscatter contributions from fluctuations in canopy gaps sizes.

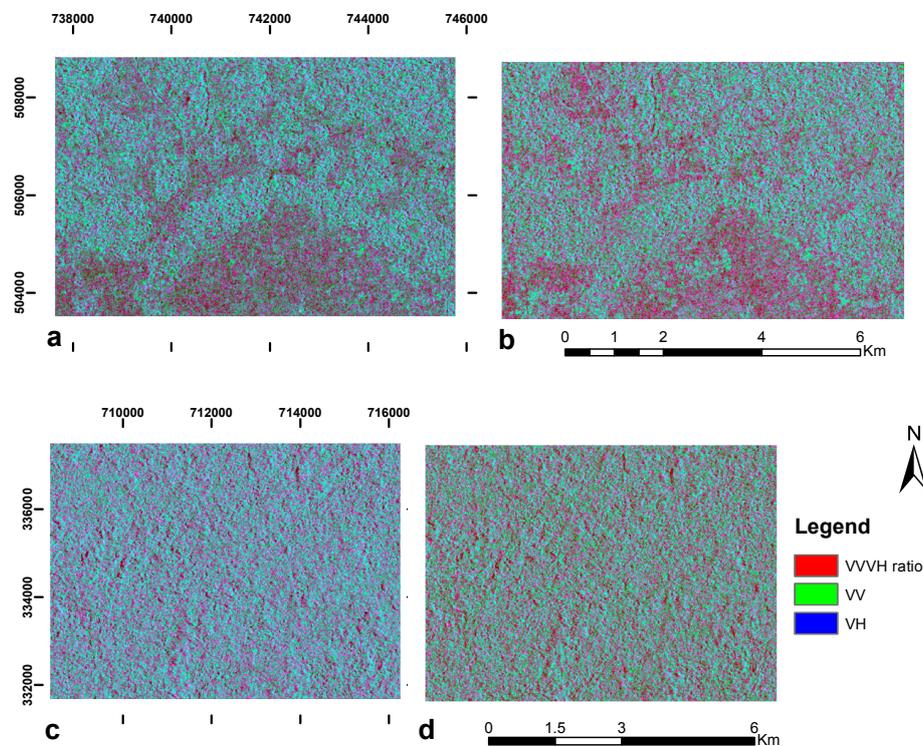


Figure 3. Illustration of SAR backscatter in the study sites: (a,b) wet and dry season SAR backscatter of the landscape in Bokito site, respective images acquired on 12 June 2016 and 10 November 2017. (c,d) wet and dry season SAR backscatter of the landscape in Efulan site, respective images acquired on 16 August 2016 and 27 November 2017. Map CRS: WGS 84/UTM.

In this study, we used ten (10) Gamma0 backscatter features to assess backscatter sensitivity to canopy cover in cocoa agroforests and transition forest stands (Table 3). The backscatter parameters comprized the following intensity coefficients, backscatter ratios, and backscatter differences and normalized differences.

Table 3. Gamma0 backscatter features used to assess backscatter sensitivity to canopy cover.

Index	Description of Feature	Used Acronym
1	VV backscatter intensity	VV Gamma0
2	VH backscatter intensity	VH Gamma0
3	VV Gamma0 in dB(decibels)	VVdb
4	VH Gamma0 in dB	VHdb
5	VVdb – VHdb	VV-VH
6	VHdb – VVdb	VH-VV
7	VVdb/VHdb	VVVHratio
8	VHdb/VVdb	VHVVratio
9	$(VVdb - VHdb)/(VVdb + VHdb)$	VV-VHnorm
10	$(VHdb - VVdb)/(VVdb + VHdb)$	VH-VVnorm

To analyze the relationship between canopy gap fraction and SAR backscatter intensity, and the backscatter behavior of the different canopy structures of cocoa agroforests and transition forests,

we created masks of either the corresponding pixel location of DHP sample points or outline polygons of the sampled agroforest areas or forest stands. Two scenarios were evaluated for the relationship between SAR backscatter coefficient and the proxy for canopy cover, canopy gap fraction: Model (1) pixel-based extraction of SAR backscatter intensity: representative 5 pixels from the overlaid five DHP collection points in each sample plot; Model (2) polygon-based extraction of SAR backscatter intensity: representative 25 pixels covering each sample plot area (Figure 4). The extracted backscatter intensity by both scenarios are, respectively, summarized in Tables S2 and S3 in the supplementary materials.

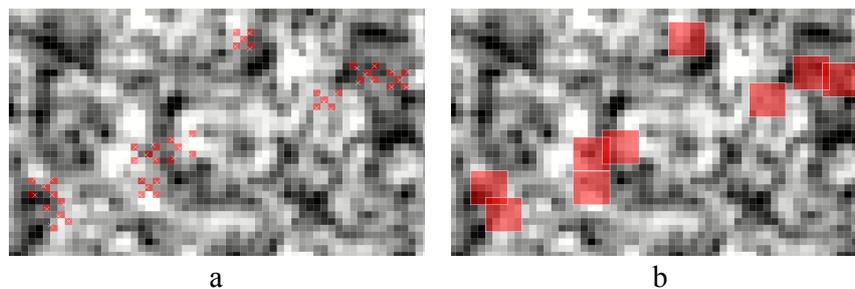


Figure 4. SNAP Toolbox snippets, of some sampling locations, illustrating the extraction of backscatter intensity from the 10 band combinations, example of VV backscatter in dB: (a) for 5 pixels masks at the DHP collection points. (b) for 25 pixels (50 m × 50 m) polygon masks encompassing sampled plots and corresponding DHP points.

2.4. Predicting Canopy Gap from SAR Backscatter

Machine learning (ML) algorithms have been applied for retrieving and modeling vegetation canopy parameters from SAR backscatter. Reference [31] used the Random Forest (RF) algorithm to model the sensitivity of S-1A backscatter to vegetation dynamics. In addition, the use of neural networks (NN) is not new; its usage encompass retrieval of different vegetation canopy parameters from SAR backscatter [61–63]. We explored, iteratively, three different ML methods for the prediction of canopy gap fraction from SAR backscatter intensity: Stepwise Multivariate Regression, the back-propagation Multi-layer Perceptron (MLP) Neural Network Regression, and the Random Forest (RF) regression algorithm. For data exploration and model assessment, a combination of Python and R programming software was used.

2.4.1. Bootstrap Regression

First, the stepwise regression analysis was used to test for significant backscatter predictors of canopy gap fraction in sampled cocoa agroforests and secondary forests. As model validation criteria, a 10-fold cross validation and the akaike information criterion (AIC) was used to select the significant regressors. Following an evaluation of the regression model performance and conditions for a parametric regression analysis, the assumptions for normality of model residuals, homogeneity of variance, and independence of variables were not met. Hence, we ran a stepwise regression with bootstrap resampling; we used 1000 bootstrap samples (or replications) for the non-parametric stepwise regression. Separate regression models were run for the ten (10) Gamma0 SAR backscatter predictor variables (or features): backscatter intensity extracted using 5 pixels masks for each reference plot (model A) and for 25 pixels and those extracted using polygon masks of 25 pixels (model B). The pixels used for model A were, thus, a subset of those used in model B.

2.4.2. Neural Network (NN) Prediction

Secondly, Multi-Layer Perceptron (MLP) Neural Networks were tested for predicting canopy gap fraction in cocoa agroforests from C-band SAR backscatter. The development of a NN that learns from data is a very empirical procedure involving the following cycle: an idea for a NN architecture

based on available data (number of neurons, number of layers in each neuron, and activation function), implementing the idea in code, experiment how the NN work for your application, and tuning the hyper-parameter to improve learning or updating the NN architecture. Considering that no NN architecture has been developed for CAFS as such, we tested and iterated different NN architectures and parameters. The backscatter features, extracted using polygon masks of 25 pixels for each reference plot, were used as predictor variables in all the models. The following three NN provided the best estimates of canopy gap fraction based on the prediction error (RMSE).

(1) NN1 comprised 10 input features (backscatter parameters), two hidden layers of 100 and 5 nodes, respectively, and the output layer. No regularization parameter was used, and the rectified linear unit (ReLU) activation function was used for the hidden and output layers (Figure S6a in Supplementary Materials). The asymptote regions in the Sigmoid and TanH function provide almost zero gradients or slope (Figure S6b), and the values in this region slow learning. The ReLU function is a non-linear function, and the term “rectified” implies taking a maximum of zero: the derivative or slope of the function is 1 for positive z (activation) values and zero for negative z . Thus, with the ReLU function, learning can be very fast since $z > 0$ in most cases. For this NN, a dropout ratio of 0.2% was used for both hidden layers, and the learning rate set at 0.001. For all the NN, we used the RMSprop (root mean squared prop) optimization algorithm to speed up gradient descent or learning. It updates the model parameters, with the square root of the squared derivatives, at each iteration to enhance oscillations in the direction that speeds-up the NN learning algorithm.

(2) NN2 had 3 successive hidden layers of 128, 64, and 10 nodes, respectively, and 1 output layer. The “Tanh” (hyperbolic tangent) activation function (Figure S6b) was set for the hidden and output layers, L2 regularization was set to 0.01. The Tanh function, like the ReLU, is non-linear and is a modification of the sigmoid function. The sigmoid function has values that range between 0 and 1 (with highest slope at 0.5), and it centers the derivatives or slopes of the activation values (z) around zero. This is handy for learning data with binary output. The tanh function centers the activation values (z) at zero (with highest derivative a zero) and a range between -1 and $+1$, which, compared to the sigmoid function, makes learning easier. To optimize learning for the NN2, dropout ratios of 0.4%, 0.3%, and 0.2% were set for the successive hidden layers, as well as a learning rate of 0.01.

(3) NN3 was different from Tuned1NN in the activation function (ReLU), L2 regularization value (0.05), and dropout rate of 0.8%, 0.4%, 0.2%, and a learning rate of 0.001. For all NN models, a random initial weighting and bias was used for the hidden layer1, and a back-propagation algorithm, based on a RMSE optimizer, was used to train the NNs and estimate the loss or gradient descent. The NN assessment was carried out using the Tensorflow library (and Keras distribution) in R.

The geo-referenced SAR images and sampled field data were used for extracting SAR backscatter data (Figure 4). The extracted SAR backscatter and estimated in-situ canopy gap fraction constituted the data set $[X_i, Y_i]$ for training the NNs, where

$$[X_i] = [\gamma_{ij}^0] \text{ and } [Y_i] = [\text{gapfraction}_i] \text{ } i = 1, 2, 3 \dots, N \text{ and } j = 1, 2, \dots, 10,$$

with N the total number of sample locations, and j the 10 extracted SAR Gamma0 backscatter intensity or ratios for location i . Thus, each X_i represents an input in training the NN, and Y_i is mean gap fraction for i sample of agroforest or transition forest: $[Y_i]$ is the output array for testing the NN prediction error.

Adapted from Reference [61] used NN training procedure we used is summarized as follows:

1. Initializing the weights and bias.
2. Layout input 2D array $[X_i]$ and desired output 1D array $[Y_i]$.
3. Calculate the NN output 1D array $[\hat{Y}_1]$.
4. Update the weights.
5. Iterating back to step 2 as needed, and depending on set learning rate, or stopping the training when a set error tolerance level is reached—early stopping.

For the input data X_i , the squared error of canopy cover prediction by the NN is:

$$SE_i(\%) = (Y_i - \hat{Y}_i)^2, \quad (1)$$

where \hat{Y}_i is the NN output for the input X_i .

Thus, RMSE, for N samples:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N SE_i}. \quad (2)$$

2.4.3. Random Forest Ensemble Regression

Thirdly, a Random Forest (RF) decision tree algorithm was applied to predict gap fraction from the 10 SAR backscatter predictors. As in the NN, we used the backscatter variables that were extracted using polygon masks of 25 pixels, per sample plot, as features for the RF model. The extracted (3350) data points for both study sites were separated into a ratio of 70%:30% training and test subsets. The decision tree hyper-parameter was tuned to an upper bound of 500 decision trees (*ntree*)—the default sufficient number of trees for running RF within the mlr package of R statistical software. Moreover, the RF important variable analysis, based on the Gini statistics [64,65], was used to identify the most influential backscatter variables in the decision trees. The RF model was applied to predict canopy gap maps for the study sites.

2.5. Semi-Variogram Analysis of Spatial Correlation in Canopy Gap Fraction

In final analysis, we evaluated the spatial continuity in canopy gap, at landscape scale, using an experimental semi-variogram model. A semi-variogram model is calculated using several different approaches in literature, and we applied the a widely used function (Equation (3)).

$$\gamma(h) = \frac{1}{2|N(h)|} \sum_{N(h)} (z_i - z_j)^2, \quad (3)$$

where:

1. $\gamma(h)$: the semi-variogram function.
2. $N(h)$: the set of all pairwise Euclidean distance $i - j = h$.
3. $|N(h)|$: the number of distinct pairs in $N(h)$.
4. z_i and z_j : data values at spatial location i and j , respectively.
5. h : distance, in magnitude only.

In geostatistics and semi-variogram theory, the experimental semivariogram is defined as the half of the average squared difference between objects or values separated by a given lag distance (h). The experimental semivariogram is computed by increasingly changing the lag distance (h) to obtain an ordered set of semi-variances from the spatial feature or data. The assumption is that closer objects are related than distant ones. In this sense, spatial features or objects located at shorter distances (h) apart are likely similar or have low variance. No variance is expected for two objects or features at zero h . Spatial variance increases with distance between spatial points or features; near observation tend to share more characteristics than distant ones. However, beyond a certain distance threshold, referred to as the range, the target features or observations are no longer spatially correlated [66]. The range indicates the distance of spatial heterogeneity, and below this distance, the data are statistically dependent. At the lag distance equal to the range, the maximum variance is attained; known as sill, this estimates the true variance in the target spatial objects or features. After the sill, the semi-variogram curve flattens off as there is no further change in the variance of the ground feature. The nugget effect or y -intercept explains a combination of errors due to data sampling and variations in the data at smaller spatial scales relative to the distance between spatial variable.

The semi-variogram model for a random spatial variable depends on parameters, such as the lag distance (h) at which spatial correlation is estimated, the choice of model, nugget or y intercept, and sill (variance in empirical data). In our analysis, the parameters, such as the variogram model, the nugget, and sill, were not set beforehand; we computed an experimental semi-variogram for the estimated canopy gap fraction and fitted a set of three common variogram models to the experimental variogram [66]: Exponential, Spherical, and Gaussian models. We assumed that forest vegetation, as well agricultural lands and grasslands, have an isotropic distribution. The best fit model was automatically selected, and its estimated parameters used to model the semi-variogram. Further, we interpret the spatial correlation in canopy gap fraction based on the reported semi-variogram curves.

A graphical abstract of the used methods and analyses is provided in Figure 5.

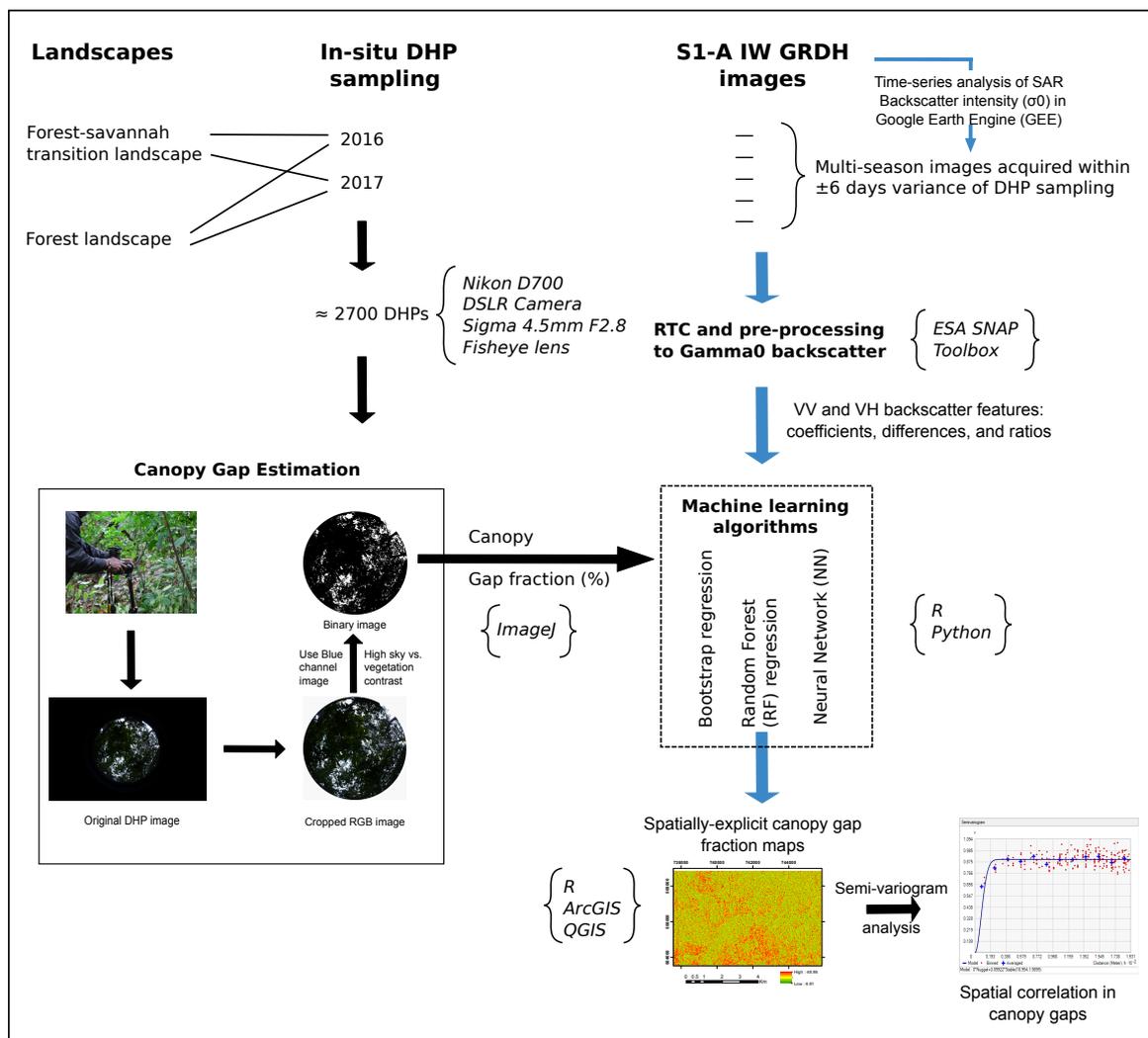


Figure 5. Framework for the used methods and estimation of canopy gap distribution. The tools or software, used at each analysis step, are indicated within braces. The black arrows = all acquired information from preceding steps in the analysis; the blue arrows = informed selection of data or algorithm from available options.

3. Results

3.1. Canopy Gap Estimation from DHPs

A boxplot analysis shows differences in canopy gap distribution for cocoa agroforests versus transition forests (Figure 6). Generally, the estimated gap fraction ranged between 2.58 to 58.23%, and, for the sample cocoa agroforests, canopy gaps were more variable in the Efoulan landscape compared to that of Bokito. The canopy gap fraction in cocoa agroforests ranged between 4.22 to 58.22%, and the range for transition forest was 2.57 to 4.46%. The differences in gap fraction between the two sites were, actually, not unexpected considering the disparities in the predominant vegetation structure both sites, and generally, more so due to the cocoa management practices. Cocoa tree density is higher in CAFS in the Center region (1217 to 1644 trees ha⁻¹) than in the south region (1048 trees ha⁻¹), and the reverse for the density of non-cocoa trees [46,67,68]. However, the range and distribution of canopy gap fraction is constraint by tree density and structure in the sampled plots.

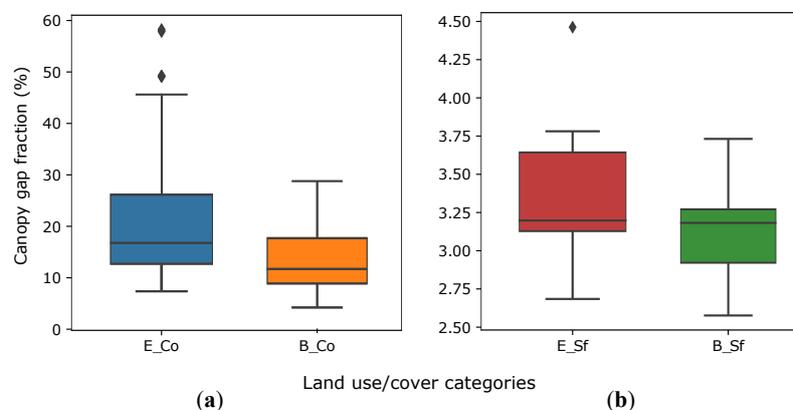


Figure 6. Boxplot of estimated gap fraction aggregated by land use categories in the 2 study sites: (a) Cocoa agroforests. (b) Transition forests. B-Co: Cocoa agroforests in Bokito, B-Sf: Secondary/Transition forests in Bokito, E-Co: Cocoa agroforests in Efoulan, and E-Sf: Secondary/Transition forests in Efoulan.

3.2. Relationship between Canopy Gap Fraction and SAR Backscatter

The statistics of stepwise regression analysis are summarized in Table 4. The stepwise regression model A indicated a combination of four backscatter features as important predictors of gap fraction, and the predictors explained 30.4% of the variance in canopy gap fraction ($R^2 = 0.304$, $F(4,129) = 14.1$, $p < 0.001$). The VV intensity contributed significantly in predicting canopy gap fraction ($\beta = 1.23$, $p < 0.001$), as did the log-transformed VV intensity ($\beta = -2.51$, $p < 0.001$) and ratio of VV on VH ($\beta = -1.97$, $p < 0.001$). For the model B, the log-transformed VV and VH backscatter, as well as their ratio, were significant predictors of canopy gap fraction.

Figures 7 and 8 illustrate the distribution of significant predictors of canopy gap fraction based on the stepwise regression model. The SAR backscatter of the co-polarization band, VV, varies between cocoa agroforests and transition forests. The VH backscatter was a significant predictor only in model B, and its distribution did not vary between the aggregated land use types (Figure 8). The cross-polarized (VH or HV) backscatter is particularly sensitive to volume scattering over forest canopy [69]. Thus, the 25 pixels (model B) cover an area of $50 \times 50 \text{ m}^2$ which, compared to 5 pixels (model A), was large enough to provide more backscatter information that relates to vegetation volume, which may be similar in the considered land cover and land use categories. Cocoa afroforests have a similar canopy structure to tropical secondary forests [67], and it is difficult to differentiate them on RS data collected from optical sensors [70]. The cross-polarization backscatter may, therefore, be regarded as less sensitive to the canopy gaps in cocoa agroforests for the sampled landscapes.

Table 4. Summary statistics for the stepwise regression models of canopy gap fraction.

Model Parameter	Model A Statistics			Model B Statistics		
	β	<i>t</i> -Test	<i>p</i> -Value	β	<i>t</i> -Test	<i>p</i> -Value
Intercept	0.00	−2.85	0.005	0.00	2.09	0.038
VH				0.32	1.92	0.057
VH (db)				1.00	2.51	0.013
VV	1.23	3.96	0.000			
VV (db)	−1.51	−6.25	0.000	−2.51	−4.56	0.000
VVHratio (db)	−0.49	−3.49	0.000	−1.97	4.15	0.000
VHVratio (db)	−0.56	−3.03	0.002			
<i>p</i> -Value		0.000			0.000	
F (n/df)		13.72 (4/19)			14.1 (4/129)	
R ²		0.29			0.30	
Adjusted R ²		0.27			0.28	
RMSE (%)		8.64			8.60	

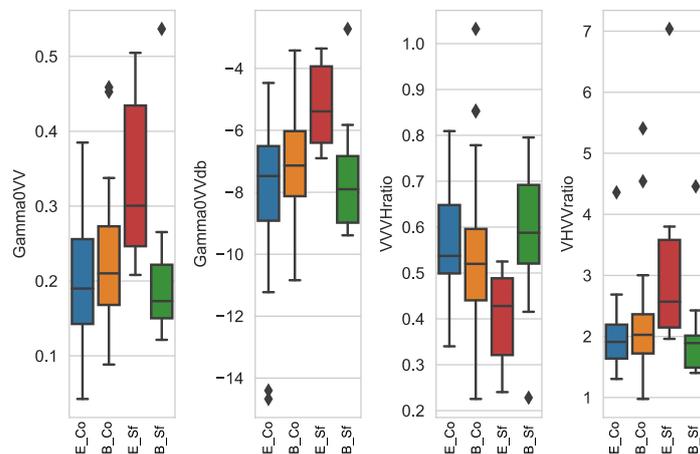


Figure 7. Boxplot distribution of significant predictors in model A, for the 4 land use categories (a combination of site and land use) (Figure 4a).

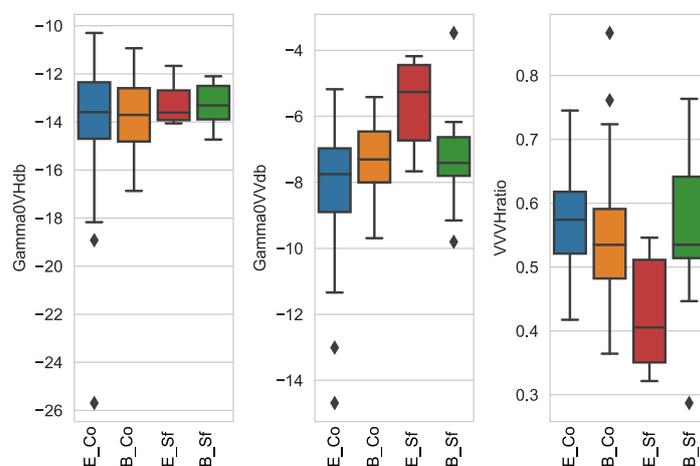


Figure 8. Boxplot distribution of backscatter the significant stepwise regression predictor variables, aggregated at four land use categories (a combination of site and land use), based on model B (Figure 4b).

3.3. Spatial Prediction of Canopy Gap Distribution

3.3.1. Bootstrap Regression Prediction

The non-parametric regression models, based on bootstrap resampling, showed that the modeled gap fractions from the sample data can provide reliable inference for the variability in canopy gaps in other locations in the population (landscapes). For both bootstrap models, the 95% Confidence Interval (CI) of both the coefficient of determination (r^2) and RMSE did not include zero (Figure 9), indicating that the model's prediction performance was reliable, without an instance of predicting a gap fraction of 0%. As well, a similar observation is showed for the 95% CI of the intercept and coefficients in model A (Table 5). On average, about 30.0% of the variance in canopy gap fraction was explained or modeled by the used backscatter variables in Model B with a 95% CI of 10.7 to 55.04% (Figures 9a,b). The average prediction error was 8.4% (Figures 9c,d). Thus, about 70% of the variance in gap fraction may be explained by other variables that were either not measured or captured by the variables used in models A and B. However, predicting canopy gap fraction from VV backscatter (in decibels), compared to the other variables, had a low bias (-0.13 , S.E of 1.95) and a narrow CI range (Table 5). This is indicative of the sensitivity of VV backscatter to the estimated canopy gap fraction. Moreover, the 95% CI of VV coefficient did not cross zero, which, once more, justifies that the estimates may be generalized for predictions about the population (the landscapes). The VH backscatter, again, was sensitive to canopy gap fraction only in model B. However, for model B, the 95% CI of intercept (-7.07 , 157.69) and VH (-357.2 , 480.4) included zero (Table 5); an indication that their parameters may not provide reliable information for making inference on canopy gap for the landscapes. In addition, the 95% CI for the intercept of model B included zero, which reveals a likely violation of the underlying assumption in the parametric regression model—the 95% CI (3.727, 132.58) did not cross zero.

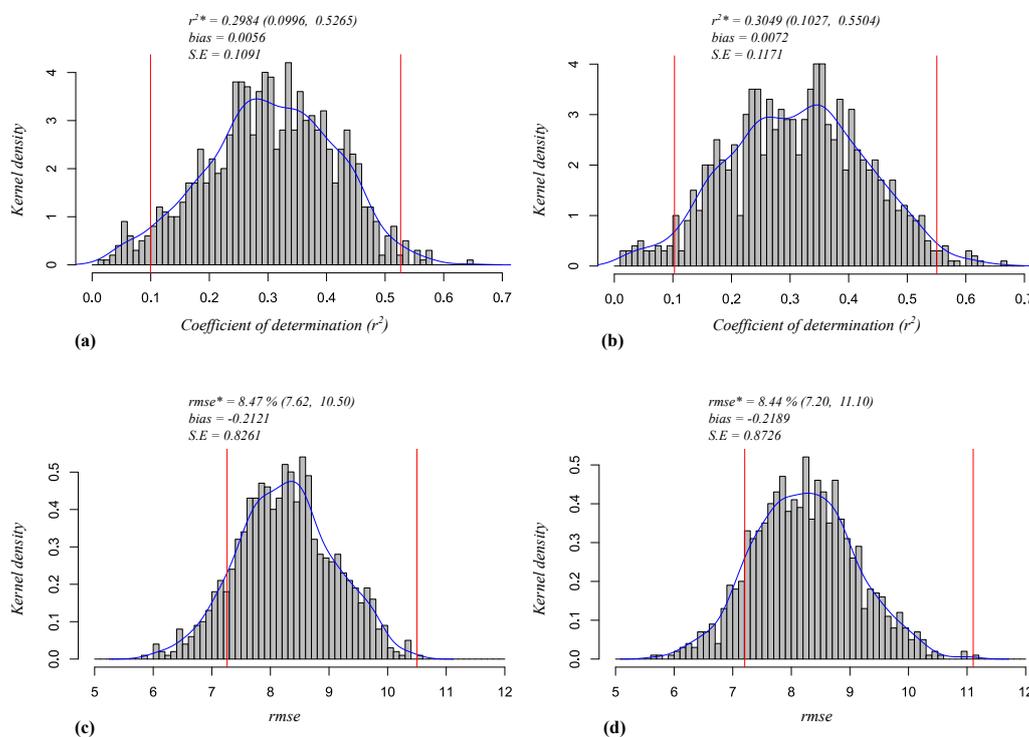


Figure 9. Bootstrap regression model performance based on 1000 bootstrap samples: (a,b) Histogram of the bootstrap coefficient of determination for model A and B, respectively. (c,d) Histogram of the bootstrap RMSE for model A and B, respectively. The blue lines represent the density plot by Gaussian Kernel smoothing, and the red lines indicate the 95% Confidence Interval (CI) of the mean r^2 *(95% CI) and mean RMSE *(95% CI).

Table 5. Summary statistics for the bootstrap regression models of canopy gap fraction (* the estimate of coefficient based on 1000 bootstrap samples, and ^a the 95 % confidence intervals based on the bias corrected and accelerated calculations).

Bootstrap Model	Model Parameters	Bootstrap Regression Statistics				Parametric Regression Statistic
		t *	Bias (β)	S.E	bca 95 % CI ^a	95 % CI
Model A (pixel)	Intercept	-42.13	-0.096 (0.00)	19.15	-89.05, -8.87	-71.3, -12.92
	VV Gamma0	145.13	7.431 (1.236)	45.48	50.80, 228.40	73.03, 218.23
	VVdb	-8.40	-0.132 (-1.511)	1.95	-13.29, -5.08	-11.06, -5.7413
	VVVH ratio	-37.69	-1.166 (-0.499)	12.71	-63.04, -13.36	-59.04, -16.34
	VHV ratio	-7	-0.875 (-0.568)	2.94	-13.04, -2.95	-12.25, -2.57
Model B (polygon)	Intercept	68.15	-1.111 (0.0)	39.21	-7.07, 157.69	3.73, 132.58
	VH Gamma0	209.41	0.621 (0.328)	174.98	-357.2, 480.4	-6.33, 425.15
	VHdb	5.08	-0.108 (1.006)	2.74	0.33, 11.95	1.24, 10.37
	VVdb	-17.05	0.314 (-2.517)	4.47	-26.00, -7.22	-24.44, -9.67
	VVVH ratio	-202.61	3.401 (-1.971)	56.46	-312.4, -79.30	-299.15, -106.08

3.3.2. Neural Network (NN) Prediction

The trained NN are compared in terms of their parameters and in prediction error of canopy gap fraction (Table 6). The RMSE of the predicted gap fraction is 7.18% for NN1, 8.02% for NN2, and 8.00% for NN3. These values are lower than the RMSE for the stepwise and RF models. The relative errors (MAE) shows high training bias for the NN with lowest error (NN3) (Figure 10). Although its RMSE was lower than for the other NN, NN1 did not generalize better than NN3, in terms of gradient descent (loss) and MAE. However, a relatively high training bias is observable for NN3, which, if addressed, can reduce the prediction error.

Table 6. Summary of parameters and statistics of the three Multi-Layer Perceptron (MLP) neural networks (NNs) (mse: Mean Squared Error, RMSE: Root Mean Squared Error, RMSprop: root mean square optimization algorithm, ReLU: Rectified Linear Unit, Tanh: hyperbolic tangent).

Layers	Parameters	Neural Networks		
		NN1	NN2	NN3
Input layer	No. input nodes	10	10	10
Hidden Layer1	No. of nodes	100	128	128
	Activation Function	ReLU	Tanh	ReLU
	Bias Regularizer		L2 (0.01)	L2 (0.05)
	Dropout (%)	0.2	0.4	0.8
Hidden Layer2	No. of nodes	5	64	64
	Activation Function	ReLU	Tanh	ReLU
	Bias Regularizer		L2 (0.01)	L2 (0.05)
	Dropout (%)	0.2	0.3	0.4
Hidden Layer3	No. of nodes		10	10
	Activation Function		Tanh	ReLU
	Bias Regularizer		L2 (0.01)	L2 (0.05)
	Dropout (%)		0.2	0.2
Output Layer	No. of nodes	1	1	1
	Loss	mse	mse	mse
	Optimizer	RMSprop	RMSprop	RMSprop
	Learning rate	0.001	0.01	0.001
Prediction	RMSE (%)	7.18	8.02	8.00

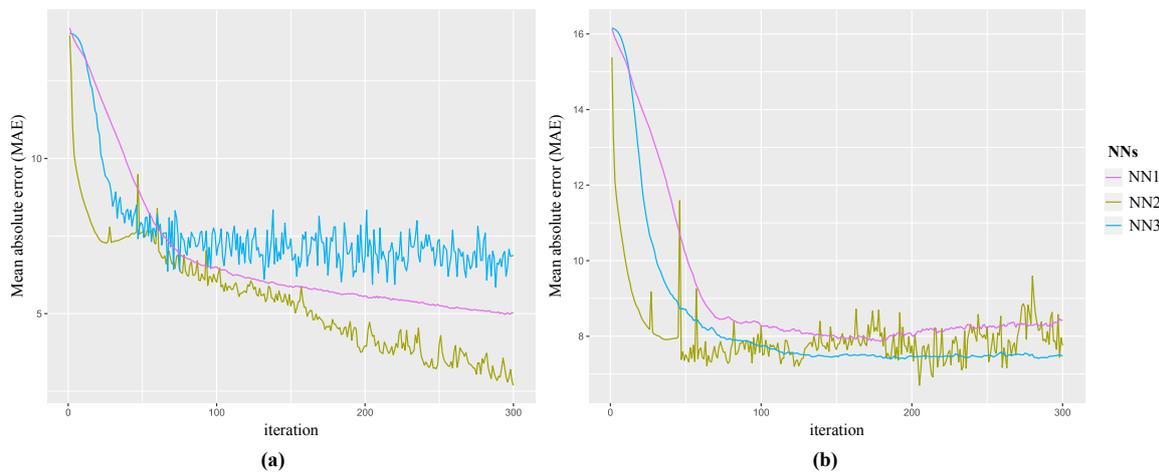


Figure 10. Mean absolute errors (MAE) of NNs for retrieving the canopy gap fraction (%) from SAR backscatter: (a) Training errors. (b) Validation errors.

3.3.3. Random Forest (RF) Prediction

The RF regression prediction error (RMSE of 8.48%) for gap fraction was lower than for the stepwise regression model which considered only a subset of backscatter parameters. The model performance accuracy was 40.94%, and, from the RF importance variable plot, the co-polarization backscatter intensity (Gamma0VV and Gamma0VH in decibel) contributed most to this (Figure 11). Thus, the co-polarization backscatter showed higher potential in predicting canopy cover. Based on the RF model results, prediction maps of the study sites are illustrated in Figure 12. The canopy gap fraction range was similar for both study sites (6.92% to 48.96%). The predicted mean canopy gap was 19.81% (SD: 10.09%) and 17.49% (SD: 8.70%), respectively, for Bokito and Efulan sites. A semi-variogram analysis revealed a low spatial correlation in canopy gap for the Bokito landscape with a distance range of 18.9 m, compared to the 51.1 m range in the Efulan landscape (Figure 13c,d). This reflects the difference in vegetation type and distribution between the landscapes. A view of some cocoa agroforestry plantation areas shows the level of mosaic and discontinuity in canopy cover within and across farms (Figures 12e–h).

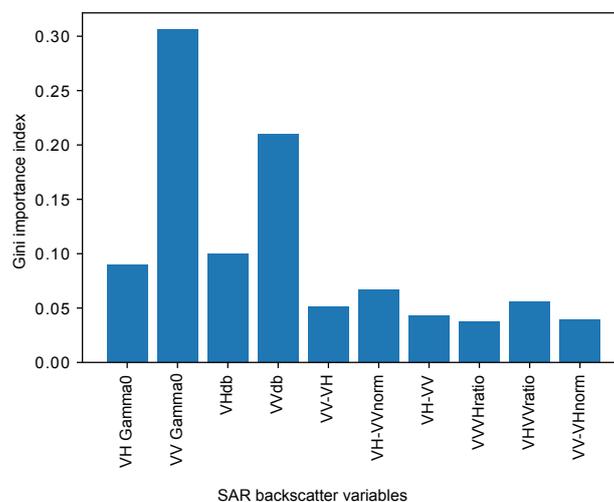


Figure 11. Barplot of SAR backscatter importance based on the Random Forest (RF) regression model prediction of canopy gap fraction from the 10 backscatter parameters.

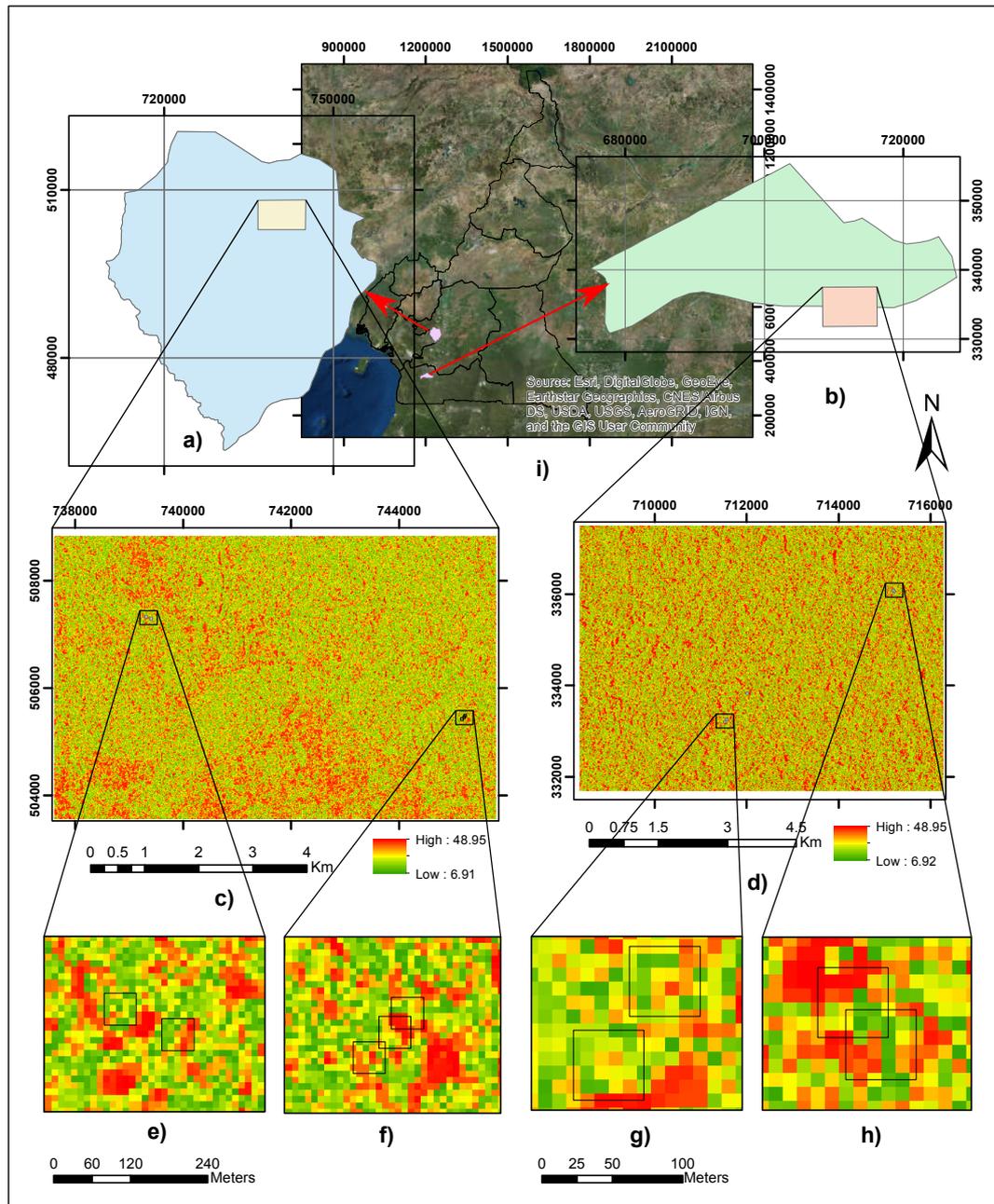


Figure 12. Canopy cover maps based on predictions from the RF regression algorithm (contrast improved by histogram equalization): (a) Study site in Bokito District. (b) Study site in Efulan District. (c) Canopy gap fraction prediction in the Bakoia-Guefige landscape based on S-1A SAR image of 10 November 2017 (mean: 19.81%, SD: 10.09%). (d) Canopy gap fraction prediction in the Minkane-Minto landscape based on S-1A SAR image of 27 November 2017 (mean: 17.49%, SD: 8.70%). (e–h) Snips of some field sample areas. (i) Location of the study sites in Cameroon (source: WRI database and World imagery base map).

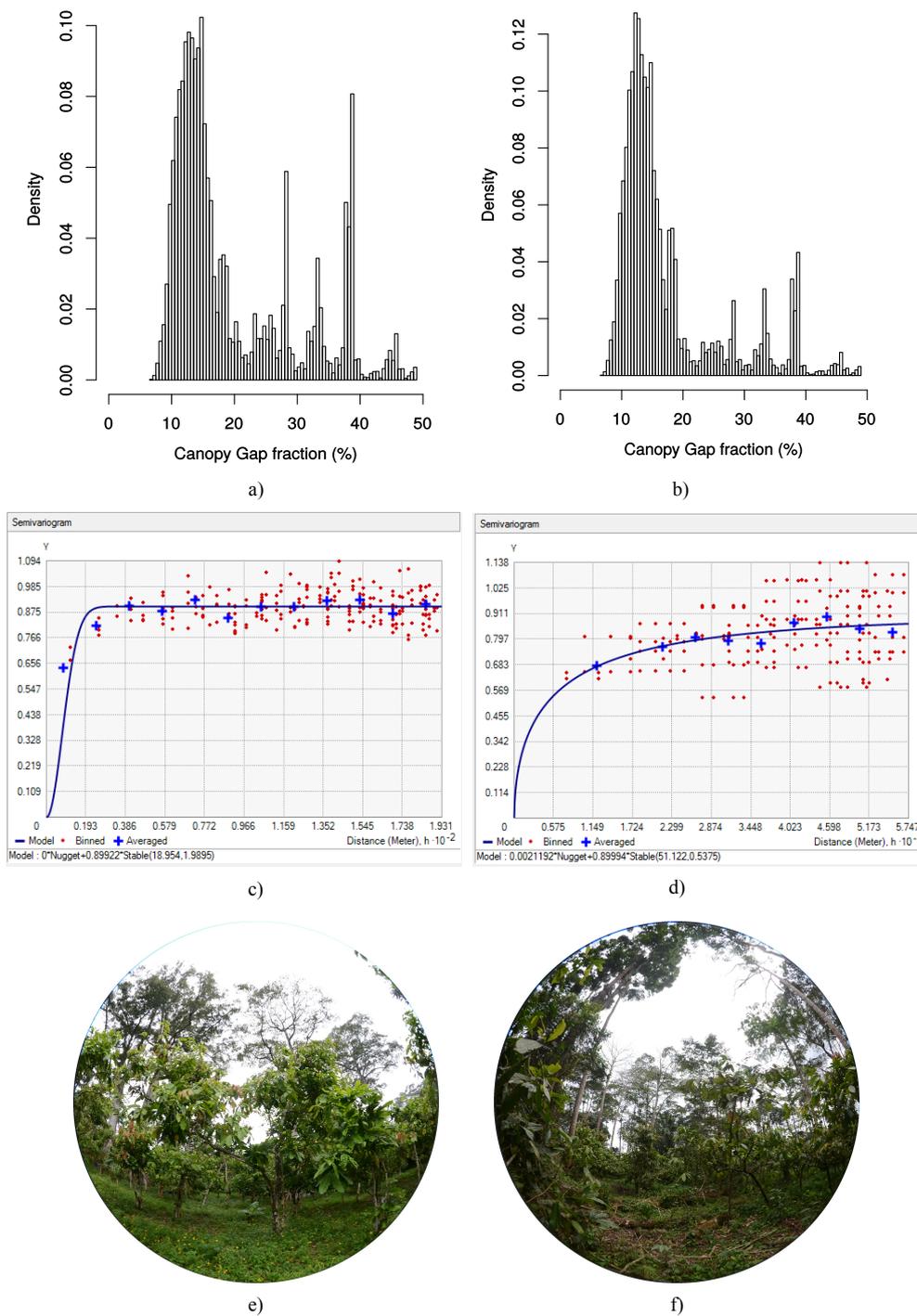


Figure 13. Experimental semi-variogram of the predicted canopy gap fraction by RF: (a) Histogram of the predicted canopy gap in the Bokito site. (b) Histogram of the predicted canopy gap in the Efoulan site. (c) Semi-variogram of canopy gap distribution in Bakoia-Guefige landscape (Bokito site). (d) Semi-variogram of canopy cover distribution in Minkane-Minto landscape (Efoulan site). Field view, from a fisheye wide angle lens (180° view angle), of the density and distribution of shade (or canopy) trees typical of mature CAFS in (e) the Bokito site and (f) the Efoulan site.

4. Discussion

We estimated canopy gap fraction as the proportion of gap within the vegetation canopy based on DHP taken below the cocoa trees or at the forest floor (1.3 m above ground). The estimates from DHP do not refer to canopy gaps above the cocoa trees since the DHP do not represent the entire vertical canopy profile. The transmitted energy from C-band wavelength is limited to top parts of vegetation canopy [71], and, in the case of CAFS, the energy can penetrate to the canopy stratum constituted by cocoa trees. Thus, the used backscatter values, by large, represent attenuated energy from the canopy gaps above cocoa trees. Using the SAR backscatter variable, we explained up to 30% of the variance in the estimated gap fraction with 8.4% error. However, including measurements of canopy gap in a vertical profile may improve the predictions from C-band SAR. In addition, the use of backscatter from longer wavelengths, L-band, for example, may better explain variability in estimated gap fraction.

Our results had RMSE and r^2 values that comparable to other published literature. For instance, Meyer et al. [72], in estimation of LAI in temperate forests by comparing several indices from Landsat-8 and Sentinel-2 images, observed r^2 values in ranging between 0.328 to 0.454. However, our analysis is based on SAR backscatter intensity which, unlike information from optical images as in the aforementioned reference, is not sensitive to vegetation spectral information or leaf phenological status. In predicting LAI of vineyards from remote sensing optical (Sentinel-2) and SAR (S-1) images, Beeri et al. [73] observed r^2 values in the range 28.3% to 31.0%. In this study, we did not apply the machine learning models to predict canopy gap fraction; rather, we explored state-of-the-art algorithms to show-case the potentials of SAR images in supporting canopy cover inventory and informing management interventions in cocoa production landscapes.

Canopy gap fraction represents the proportion of sky hemisphere as viewed from a specific point [4]. Our field measurements of canopy cover were based on a DHP with a non-zero angle of view (AOV) perspective of the canopy. The non-zero AOV (180° for the used fisheye lens) had a wide canopy area and, mainly, side view of tree crowns at each sampling point. Thus, the canopy cover maps may well be interpreted as angular perspective of canopy cover and not a vertical (or map) projection as such. These maps reflect the spatial distribution of vegetation types and canopy trees in both landscapes, and it should be interpreted with consideration of a prediction error of 8.4%.

Unlike the low frequency microwaves (L- and P-bands) that can reach the understorey canopy and underlying soil, the high frequency C-band microwave penetration is limited to the upper canopy strata [41]. Moreover, the sensitivity of either co- or cross-polarization S-1A backscatter intensity to vegetation canopy vary with the crop type [56,57]. In our experiment, however, the VH polarization was less performant in predicting canopy cover based on the bootstrap regression results. The cross-polarized backscatter (VH) significantly contributed in explaining variability in gap fraction only in model B. Ideally, the dominant canopy components, often timber trees species, in cocoa agroforests and transition forests comprise similar tree species and canopy structure. Most of the tree species are light-tolerant and deciduous, which results to a less dense canopy structure. Thus, since the C-band microwave penetrate forest canopy only to a limited extent, CAFS and transition forests may experience similar volume scattering of microwaves as such. This may explain the similar distribution of the VH backscatter across the aggregated land use categories (Figure 8), and its poor performance in modeling canopy cover. Obvious differences between the two land use/cover types relate to the density of canopy trees, which may also be due to spatial fluctuations in the vertical canopy profile. For this, the distribution of VV backscatter intensity varies between land use aggregates (Figures 7 and 8).

Making use of all the backscatter intensity and ratios, the MLP NN showed a high potential in retrieving canopy gap (or cover) status over cocoa agroforests and adjacent transition forests. Though the canopy gap prediction errors (MAE) for the NN were in a range of 5 to 7%, the RMSE values were lower than those from bootstrap regression models that used a few backscatter predictors. The NN1 showed reduction in training error to about 2% with increase in model iteration, but the model was poor in generalizing new data. The high number of predictor bands for the low sample data may explain the high variance or model over-fitting beyond 75 model runs or iterations. So far,

further tuning of NN hyper-parameter, NN3, reduced over-fitting but did not improve the prediction error; meanwhile, the MAE was by far higher than that of the former NN.

4.1. Sensitivity of S-1A Backscatter (VV vs. VH) to Canopy Gap Fraction

The normalized backscatter coefficients (sigma nought) is a measure of the cross-sectional area of the, spatially random, scatterers within each resolution cell or pixel. Thus, the radar backscatter energy depends on the shape, dielectric constant, density, and spatial configuration of the ground targets in each pixel; there are also contributions from sensor properties, such as wavelength, incidence angle, and polarization. We used the gamma backscatter which corrects for effects of different incident angle on backscatter cross-section [32,54,69].

According to our model predictions by the bootstrap regression and RF, the VV backscatter was more sensitive or important in predicting canopy gap fraction. Literature has widely reported a high sensitivity of cross-polarized backscatter to forest canopy or volume. The gap fraction is related to LAI (or plant area index (PAI)) and a commonly used proxy for canopy volume; it refers to a different parameter of vegetation canopy. LAI has been reported to be correlated with backscatter volumes over forest and crop fields [57,60], and it is commonly estimated using DHP [74,75]. Our gap fraction estimates, based on DHP measurements, is the inverse of LAI [22]. Thus, we hypothesize a different relationship between the former and microwave backscatter: high sensitivity to the co-polarized (VV) backscatter to surface scattering in canopy gaps. A further explanation is that the C-band wavelength mildly penetrates forest canopies and, thus, relates to LAI at a lesser degree than to canopy gap. However, our models are subject of further fine-tuning and improvement considering the high variance in the predictions.

Several factors may, thus, influence information retrieval from SAR images. For this study, temporal changes in SAR backscatter intensity may have contributed to a reduced prediction potentials. However, for the period of field data collection, a time-series analysis showed differences in backscatter only between a few images (Figures S2 and S3, and Table S1 in the Supplementary Materials). Moreover, our analysis was conducted for perennial tropical agricultural land use, and consistent changes in vegetation structure may be expected only between periods of major climatic and rainfall patterns: the dry and rainy seasons. Amongst other factors, SAR backscatter intensity is dependent on the geometry of ground features, as well as their dielectric properties. SAR backscatter intensity is also largely influenced by spatial distribution of ground features in each resolution cell; this intensity may vary between representative pixels for similar ground features—the salt and pepper or speckle effect in SAR images. Speckle effect in SAR images may contribute to degrading the sensitivity to measured parameters on ground features [76]; thus, canopy gap fraction prediction from the used SAR images may be negatively influenced by variability in backscatter intensity values for a similar range of gap fractions estimated by hemispherical photos. Nonetheless, an important source of uncertainty in our prediction is the hemispherical photos we took below the canopy (1.3 m above ground) which could result in an underestimation of actual gaps at the upper canopy strata. Here, the short wavelengths from Sentinel-1 C-band sensor penetrate at most (and interact with features in) the upper canopy strata.

Forests are often considered as heterogeneous volumes when in contact with microwave. The main scatterers from forest canopies comprise a combination of leaves, twigs, and branches that are randomly distributed and have different geometry. Over forest, cross-polarized backscatter is dominated by the contribution from canopy elements than by surface components. Thus, the cross-polarized backscatter often correlates best with biomass, and it is reported to have a high sensitivity to volume scattering [57,77]. However, the depolarization effect of vegetation reduces with increasing vegetation. According to Reference [69], there is a high contrast between H and V polarization over bare ground, and the contrast reduces with increasing amount of vegetation. This may provide additional explanation for the observed distribution of VH between land cover types and its prediction performance for canopy gap fraction.

In general, the small canopy elements are numerous and located higher in the canopy and the larger elements tend to be near the ground. The reduced penetration depth of C-band (and X-band) microwave into forest canopies is, therefore, due to the relatively large sizes of lower canopy elements compared to the short wavelengths. In dense forest canopies, therefore, the backscatter from such microwaves easily saturates and represents only a cross-section of the top canopy. However, the edge of forests or canopy trees is particularly significant for radar system as such edges can provide conditions that result in “double-bounce” backscatter [69,78]. The co-polarized backscatter is sensitive to conditions of double bounce which is provided by surface components of vegetation [69]. For the sparse and multi-strata canopy as in CAFS, canopy elements are to some extent arranged in vertical directions considering the gaps within and between different canopy strata. Thus, even for C-band microwave (Sentinel 1A in this case), we can expect double-bounce backscatter within such canopy conditions. For example, radar scattering between the elements of the canopy and/or trunk of a tree in the higher canopy stratum relative to canopy elements in a lower stratum. These conditions for double-bounce, and the vertical elements within canopies and canopy gaps, may explain the observed high importance of the VV backscatter in our prediction of canopy gap fraction.

4.2. Spatial Variability in Canopy Gap Fraction

Our experimental semi-variogram models showed that canopy gap distribution in the Bokito landscape is spatially correlated up to a distance of ≈ 18 m, unlike in the Efoulan landscape where spatial correlation in canopy gap is observable up to a distance range of ≈ 51 m. The semi-variogram for the canopy gap fraction in the forest-savannah transition landscape showed a combination of Gaussian and Spherical spatial distribution (Figure 13c). This is indicative of a short-scale continuity in canopy gaps in the landscape. The characteristic canopy tree and vegetation structure in the landscape may explain the variogram results. In most CAFS in the landscape, the top canopy stratum of shade trees comprise few, and often isolated, trees of the following predominant species: *Ceiba pentandra* (Figure 13e), *Milicia excelsa*; *Canarium schweinfurthii*, *Ricinodendron heudelottii* (Non-Timber Forest Product (NTFP) tree), *Pycnanthus angolensis*, *Triplochiton scleroxylon*, etc. The average crown diameter for mature specimens of these tree species is usually in the range of 15 to 25 m. For instance, based on forest inventory data in Nigeria, Reference [79] observed that the crown length of *Milicia excelsa* is in the range 3.6–27 m with a mean of 17 m. Thus, the spatial correlation in gap fraction at a range distance of ≈ 18 m may be contributed by the isolated large shade trees in CAFS.

For the forest landscape, the spatial distribution of canopy gap fraction follow an exponential semi-variogram model (Figure 13d). There was spatial correlation in canopy gaps at long scales up to a distance of 51 m. The dominant canopy stratum in most CAFS, compared to the Bokito site, comprised a higher density of different timber and NTFP tree species (Figure 13f). Some common species identified on CAFS include: *Pycnanthus angolensis*, *Iringia gabonensis* (NTFP), *Milicia excelsa*, *Piptadeniastrum africanum*, *Pentaclethra macrophylla*, *Albizia glaberrima*, *Tetrapleura tetraptera* (NTFP), etc. A similar interpretation, therefore, stands that the spatial correlation in gap fraction up to distance of 51 m in the forest landscape is as a result of the typical spatial distribution of large shade trees in CAFS. The spatial correlation in canopy gap, however, applies more to cocoa agroforests with multi-storey shade trees and secondary forest areas as these were the main sampling areas. Field data on canopy gap fraction did not include other features of the landscapes, such as agricultural fields of other crop, bare land, and savannah vegetation (grass and shrub). Following the results from an experimental semi-variogram, variogram models may be further explored to ascertain the spatial distribution of canopy gaps at different spatial distances and directions.

4.3. Management Considerations for Canopy Gap Distribution

Observations in different landscapes have widely suggested that about 30–50% shade may suffice to sustain cocoa production in agroforestry systems [8,80,81]. Yet, the estimation of shade cover in multi-strata CAFS is daunting and challenging; often, farmers lack the appropriate technical skills to

operate different forestry equipment for canopy cover assessment and are bound to rely on personal subjective observations of shade percentage in farms. Thus, for the large proportion of cocoa farmers, small-scale rural producers, such equipment widely do not serve the intended purpose. To circumvent such limitations towards appropriate shade management, an assessment of canopy gap distribution at a landscape scale is suggested in this study, which may guide management of shade trees and cocoa yield. We did not measure cocoa productivity in sampled farms; consequently, we do not directly relate farm productivity with canopy gap fraction estimates. However, over a one-year assessment, Reference [82] estimated the average potential yield of cocoa in the Bokito landscape to be 819 kg ha^{-1} , and the mean accessible yield of cocoa in mature CAFS of Central Cameroon (including Bokito, Ngomedzap, and Talba villages) was estimated to be $869 \text{ kg ha}^{-1} \text{ yr}^{-1}$ [45]. Thus, managers may relate, though not in a direct causal sense, the spatial distribution of canopy gaps with reported cocoa productivity in a said landscape.

Cocoa productivity is influenced by several ecological, economic, and environmental factors, which may vary from one production landscape to another. In ongoing joint stakeholders commitments towards zero deforestation cocoa, therefore, improving management strategies in relation to farm intensification, associated tree diversity, or farmers' technical knowledge [9,45,68] alone would be ineffective to sustain and enhance cocoa production. Climate parameters are currently widely at play in influencing farm productivity in cocoa landscapes. For instance, according to Reference [83], cocoa productivity in Bokito is more susceptible to climatic parameters compared to a production zone in the forest landscape. However, due to a high prevalence of the pest that causes cocoa black pod disease, *Pythophthora megakarya*, and related farm management challenges, CAFS in the forest landscapes, South of Cameroon, have reported low productivity [84]. Thus, managers and stakeholders may require a "solution basket" to sustain and improve cocoa production, and, top in the basket, there is need to map productivity and identify areas likely vulnerable to erratic climate parameters. We suggest an assessment tool for potential changes in canopy structure in cocoa landscapes, which can guide sampling and managing shade tree density and canopy groves in landscapes as such.

5. Conclusions

In this study, we assessed the relationship between the canopy cover, with focus on cocoa agroforests, and the C-band SAR backscatter intensity. Field inventory was conducted in two cocoa production landscapes in Cameroon that differ in vegetation structure, agricultural practice, and land tenure, amongst other livelihood and institutional difference.

Our results show that the co-polarization backscatter, VV, was more sensitive to canopy gap compared to VH polarization. Moreover, of the significant bootstrap regression predictors of canopy gap fraction, the distribution of VV backscatter varied between the aggregated land use/cover categories. Still, the VV backscatter intensity contributed most to the explained variance in canopy gap fraction by RF regression. Thus, we suggest the VV backscatter to be more sensitive to canopy gaps in cocoa agforests, and this co-polarization backscatter coefficients can be subject of focus to improve predictions from C-band SAR data. Based on the RF model prediction, we provide canopy gap fraction maps for the study sites. Noteworthy, the canopy gap prediction errors (RMSE and MAE) for the NN models were lower than the RF and stepwise regression; this may be a potential basis for further landscape analyses. The results show potential of S1 for canopy gap fraction prediction in CAFS, although further fine-tuning is needed to obtain a better prediction accuracy.

The NN model results can be explored further in either of the following ways, or both, depending on available time or resources: first, further tuning of NN layers, number of neurons in each layer, and model regularization parameters can reduce the high bias and prediction error of NN Tuned1; a second and less feasible procedure is collection of additional and sufficiently large sample data and tuning of model parameters may reduce the over-fitting of NN Tuned1 and hopefully retain its potential low prediction error. For the second approach, the needed training data may also be retrieved from multi-temporal and multi-polarization SAR backscatter [61].

Until now, literature mainly reports shade cover percentage or levels as a proxy for shade tree management in cocoa agroforests. For example, Reference [85] suggested that a shade level of above 30% may not provide expected outcome of sustained cocoa crop production and biodiversity. While such values provide useful guide for shade management at the scale of individual CAFS, shade tree management plans at landscape levels may require information on the spatial distribution of canopy cover. Such information is yet unknown for cocoa production landscapes in Cameroon. The distance range of spatial correlation in gap fraction provides new information for the landscapes. We envisage that this is a valuable basis for inventorying and managing tree shade cover in CAFS at the landscape scale. As example, a practical application is in monitoring canopy cover status in CAFS landscapes and designing sustainable management practices, especially for areas likely vulnerable to changing and extreme weather conditions as prolonging dry seasons.

We have conducted an analysis based on field data and the availability of Sentinel-1 SAR in the context of tropical agroforestry landscapes. Our analysis and results are not meant to provide a “black box” or generic application for all of such landscapes. We used state-of-the-art machine learning algorithms to explore the potential of free S1 data in a landscape with typical vegetation cover, type of anthropogenic action and land use, and poor data availability. Worthy of note, this study supports a proof-of-concept, from applying C-band SAR backscatter information, and the findings are comparable to others reported in literature [76,86]. Nonetheless, they are based on backscatter intensity from C-band microwave, which only penetrates forest canopy to a certain degree. We think that improved results can be achieved using data from SAR images acquired using longer wavelengths as L- or P-band. For instance, data from up-coming satellite missions, such as NISAR (NASA-ISRO Synthetic Aperture Radar), which will be launched, hopefully, in 2022, may be very handy for such applications. However, due to overlapping satellite flight orbits, S-1 data acquisition has a high frequency of SAR images collected over areas at higher latitudes than at the equatorial regions. Likewise, up-coming SAR mission may also be launched with predetermined flight orbits and specific duty cycles for prioritized locations (e.g., India for NISAR) and lower image acquisition frequency over certain parts of the globe. Hence, currently operational S-1 C-band SAR image archives will remain a valuable source of complementary remote sensing data to be explored for different applications.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-4292/12/24/4163/s1>, Figure S1: DHP Sampling; Figures S2 and S3: Time-series analysis of S1-A SAR backscatter; Figures S4 and S5: Summary statistics of extracted SAR backscatter; Figure S6: Neural network activation functions; Table S1: Summary of mean difference in SAR backscatter (σ_0) across the DHP field sampling window period in the study study landscapes; Tables S2 and S3: Summary statistics of extracted SAR backscatter.

Author Contributions: Conceptualization, F.N.N.; methodology, F.N.N., F.V.C.; formal analysis, F.N.N.; investigation, F.N.N.; data curation, F.N.N.; writing—original draft preparation, F.N.N.; writing—review and editing, F.N.N., F.V.C.; visualization, F.N.N.; supervision, F.V.C.; project administration, F.N.N., F.V.C.; funding acquisition, F.N.N. All authors have read and agreed to the published version of the manuscript.

Funding: Special Research Fund, Ghent University, contract number BOF DOS01W03314.

Acknowledgments: This study was conducted under the Special Research Fund of Ghent University for students from developing countries. The research was conducted under the guidance of Robert De Wulf. Field work was supported by the World Agroforestry Center in Cameroon. Field inventory was assisted by local resource persons, administrative personnel, and the cocoa farmers who granted access to their plantations. We thank the European Space Agency for the freely accessible archive of Sentinel-1 SAR images. We also thank to the developers (and contributors) of the R and Python data mining and programming software packages. We thank the four anonymous reviewers for their constructive comments and suggestions that helped to improve the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Costanza, R.; Arge, R.; Groot, R.D.; Farber, S.; Grasso, M.; Hannon, B.; Limburg, K.; Naeem, S.; Neill, R.V.O.; Paruelo, J.; et al. The value of the world's ecosystem services and natural capital. *Nature* **1997**, *387*, 253–260. [[CrossRef](#)]
2. Mortimer, R.; Saj, S.; David, C. Supporting and regulating ecosystem services in cacao agroforestry systems. *Agrofor. Syst.* **2018**, *92*, 1639–1657. [[CrossRef](#)]
3. Bernard, F.; Minang, P.A. Community forestry and REDD + in Cameroon: What future ? *Ecol. Soc.* **2019**, *24*, 1–12. [[CrossRef](#)]
4. Korhonen, L.; Heikkinen, J. Automated analysis of in situ canopy images for the estimation of forest canopy cover. *For. Sci.* **2009**, *55*, 323–334.
5. Somarriba, E.; Orozco-Aguilar, L.; Cerda, R.; López-Sampson, A. Analysis and design of the shade canopy of cocoa-based agroforestry systems. *Achiev. Sustain. Cultiv. Cocoa* **2018**, 469–499. [[CrossRef](#)]
6. Somarriba, E.; Cerda, R.; Orozco, L.; Cifuentes, M.; Dávila, H.; Espin, T.; Mavisoy, H.; Ávila, G.; Alvarado, E.; Poveda, V.; et al. Carbon stocks and cocoa yields in agroforestry systems of Central America. *Agric. Ecosyst. Environ.* **2013**, *173*, 46–57. [[CrossRef](#)]
7. FAO. *FRA 2015 Terms and Definitions—The Forest Resources Assessment Programme*; FAO: Roma, Italy, 2015.
8. Mbile, P.; Ngaunkam, P.; Besingi, M.; Nfoumou, C.; Degrande, A.; Tsobeng, A.; Sado, T.; Menimo, T. Farmer management of cocoa agroforests in Cameroon: Impacts of decision scenarios on structure and biodiversity of indigenous tree species. *Biodiversity* **2009**, *10*, 12–19. [[CrossRef](#)]
9. Alemagi, D.; Duguma, L.; Minang, P.A.; Nkeumoe, F.; Feudjio, M.; Tchoundjeu, Z. Intensification of cocoa agroforestry systems as a REDD+ strategy in Cameroon: Hurdles, motivations, and challenges. *Int. J. Agric. Sustain.* **2015**, *13*, 187–203. [[CrossRef](#)]
10. Luedeling, E.; Kindt, R.; Huth, N.I.; Koenig, K. Agroforestry systems in a changing climate—Challenges in projecting future performance. *Curr. Opin. Environ. Sustain.* **2014**, *6*, 1–7. [[CrossRef](#)]
11. Sonwa, D.J.; Weise, S.F.; Janssens, M.J.J.; Schroth, G.; Shapiro, H.Y. Structure of cocoa farming systems in West and Central Africa: A review. *Agrofor. Syst.* **2018**. [[CrossRef](#)]
12. Schroth, G.; Läderach, P.; Martinez-valle, A.I.; Bunn, C.; Jassogne, L. Vulnerability to climate change of cocoa in West Africa: Patterns , opportunities and limits to adaptation. *Sci. Total Environ. J.* **2016**, *556*, 231–241. [[CrossRef](#)] [[PubMed](#)]
13. Gateau-rey, L.; Tanner, E.V.J.; Rapidel, B.; Marelli, P.; Royaert, S. Climate change could threaten cocoa production: Effects of 2015-16 El Niño-related drought on cocoa agroforests in Bahia , Brazil. *PLoS ONE* **2018**, *13*, e0200454. [[CrossRef](#)] [[PubMed](#)]
14. Hardwick, S.R.; Toumi, R.; Pfeifer, M.; Turner, E.C.; Nilus, R.; Ewers, R.M. The relationship between leaf area index and microclimate in tropical forest and oil palm plantation: Forest disturbance drives changes in microclimate. *Agric. For. Meteorol.* **2015**, *201*, 187–195. [[CrossRef](#)] [[PubMed](#)]
15. Korhonen, L.; Korhonen, K.T.; Rautiainen, M.; Stenberg, P. Estimation of forest canopy cover: A comparison of field measurement techniques. *Silva Fennica Res. Artic.* **2006**, *40*, 577–588. Available online: <https://jukuri.luke.fi/handle/10024/532615> (accessed on 18 December 2020)
16. Fiala, A.C.S.; Garman, S.L.; Gray, A.N. Comparison of five canopy cover estimation techniques in the western Oregon Cascades. *For. Ecol. Manag.* **2006**, *232*, 188–197. [[CrossRef](#)]
17. Riemann, R.; Liknes, G.; O'Neil-Dunne, J.; Toney, C.; Lister, T. Comparative assessment of methods for estimating tree canopy cover across a rural-to-urban gradient in the mid-Atlantic region of the USA. *Environ. Monit. Assess.* **2016**, *188*. [[CrossRef](#)]
18. Pope, G.; Treitz, P. Leaf Area Index (LAI) estimation in boreal mixedwood forest of Ontario, Canada using Light detection and ranging (LiDAR) and worldview-2 imagery. *Remote Sens.* **2013**, *5*, 5040–5063. [[CrossRef](#)]
19. Qu, Y.; Shaker, A.; Silva, C.A.; Klauber, C.; Pinagé, E.R. Remote sensing of leaf area index from LiDAR height percentile metrics and comparison with MODIS product in a selectively logged tropical forest area in Eastern Amazonia. *Remote Sens.* **2018**, *10*, 970. [[CrossRef](#)]
20. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A review of supervised object-based land-cover image classification. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 277–293. [[CrossRef](#)]
21. Shin, P.; Sankey, T.; Moore, M.M.; Thode, A.E. Evaluating unmanned aerial vehicle images for estimating forest canopy fuels in a ponderosa pine stand. *Remote Sens.* **2018**, *10*, 3–5. [[CrossRef](#)]

22. Beckschäfer, P.; Fehrmann, L.; Harrison, R.D.; Xu, J.; Kleinn, C. Mapping leaf area index in subtropical upland ecosystems using rapideye imagery and the randomforest algorithm. *IForest* **2014**, *7*, 1. [CrossRef]
23. Kalácska, M.; Sánchez-Azofeifa, G.A.; Rivard, B.; Calvo-Alvarado, J.C.; Journet, A.R.P.; Arroyo-Mora, J.P.; Ortiz-Ortiz, D. Leaf area index measurements in a tropical moist forest: A case study from Costa Rica. *Remote Sens. Environ.* **2004**, *91*, 134–152. [CrossRef]
24. Zheng, G.; Moskal, L.M. Retrieving Leaf Area Index (LAI) Using Remote Sensing: Theories, Methods and Sensors. *Sensors* **2009**, *9*, 2719–2745. [CrossRef] [PubMed]
25. Ganguly, S.; Nemani, R.R.; Zhang, G.; Hashimoto, H.; Milesi, C.; Michaelis, A.; Wang, W.; Votava, P.; Samanta, A.; Melton, F.; et al. Generating global Leaf Area Index from Landsat: Algorithm formulation and demonstration. *Remote Sens. Environ.* **2012**, *122*, 185–202. [CrossRef]
26. Chianucci, F.; Cutini, A. Estimation of canopy properties in deciduous forests with digital hemispherical and cover photography. *Agric. For. Meteorol.* **2013**, *168*, 130–139. [CrossRef]
27. Koedsin, W.; Yasen, K. Estimating Leaf Area Index of Rubber Tree Plantation Using Worldview-2 Imagery. *J. Life Sci. Technol.* **2016**, *4*, 1–6. [CrossRef]
28. Kersten, F. Radar Polarimetry- Potential for Geosciences Microwave remote sensing and SAR. *Microwaves* **2006**, *1*–10. Available online: http://www.geo.tu-freiberg.de/oberseminar/os05_06/Franziska_Kersten.pdf (accessed on 18 December 2020).
29. Tanase, M.A.; Ismail, I.; Lowell, K.; Karyanto, O.; Santoro, M. Detecting and quantifying forest change: The potential of existing C- and X-band radar datasets. *PLoS ONE* **2015**, *10*, e0131079. [CrossRef]
30. Ningthoujam, R.K.; Tansey, K.; Balzter, H.; Morrison, K.; Johnson, S.C.; Gerard, F.; George, C.; Burbidge, G.; Doody, S.; Veck, N.; et al. Mapping forest cover and forest cover change with airborne S-band radar. *Remote Sens.* **2016**, *8*, 577. [CrossRef]
31. Vreugdenhil, M.; Wagner, W.; Bauer-Marschallinger, B.; Pfeil, I.; Teubner, I.; Rüdiger, C.; Strauss, P. Sensitivity of Sentinel-1 Backscatter to Vegetation Dynamics: An Austrian Case Study. *Remote Sens.* **2018**, *10*, 1396. [CrossRef]
32. Rüetschi, M.; Schaepman, M.E.; Small, D. Using multitemporal Sentinel-1 C-band backscatter to monitor phenology and classify deciduous and coniferous forests in Northern Switzerland. *Remote Sens.* **2018**, *10*, 55. [CrossRef]
33. Rüetschi, M.; Small, D.; Waser, L.T. Rapid detection of windthrows using Sentinel-1 C-band SAR data. *Remote Sens.* **2019**, *11*, 115. [CrossRef]
34. Ko, D.; Bristow, N.; Greenwood, D.; Weisberg, P. Canopy cover estimation in semiarid woodlands: Comparison of field-based and remote sensing methods. *For. Sci.* **2009**, *55*, 132–141.
35. Asrat, Z.; Taddese, H.; Ørka, H.O.; Gobakken, T.; Burud, I.; Næsset, E. Estimation of Forest Area and Canopy Cover Based on Visual Interpretation of Satellite Images in Ethiopia. *Land* **2018**, *7*, 92. [CrossRef]
36. Hirschmugl, M.; Sobe, C.; Deutscher, J.; Schardt, M. Combined Use of Optical and Synthetic Aperture Radar Data for REDD+ Applications in Malawi. *Land* **2018**, *7*, 116. [CrossRef]
37. Dobson, M.C.; Ulaby, T.F.; Leland, P.E.; Sharik, T.L.; Bergen, K.M.; Kellndorfer, J.; Kendra, J.R.; Li, E.; Lin, Y.C.; Nashashibi, A.; et al. Estimation of Forest Biophysical Characteristics in Northern Michigan with SIR-C_X-SAR. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 877–895. [CrossRef]
38. Vastaranta, M.; Niemi, M.; Karjalainen, M.; Peuhkurinen, J.; Kankare, V.; Hyypä, J.; Holopainen, M. Prediction of forest stand attributes using TerraSAR-X stereo imagery. *Remote Sens.* **2014**, *6*, 3227–3246. [CrossRef]
39. Moreira, A.; Prats-iraola, P.; Younis, M.; Krieger, G.; Hajnsek, I.; Papathanassiou, K.P. A Tutorial on Synthetic Aperture Radar. *IEEE Geosci. Remote Sens. Mag.* **2013**. [CrossRef]
40. Steele-Dunne, S.; McNairn, H.; Monsivais-Huertero, A.; Judge, J.; Liu, P.W.; Papathanassiou, K. Radar Remote Sensing of Agricultural Canopies. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2249–2273. [CrossRef]
41. Sivasankar, T.; Kumar, D.; Srivastava, H.S.; Patel, P. Advances in Radar Remote Sensing of Agricultural Crops: A Review. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2018**, *8*, 1126. [CrossRef]
42. Jagoret, P.; Michel-Dounias, I.; Snoeck, D.; Ngnogué, H.T.; Malézieux, E. Afforestation of savannah with cocoa agroforestry systems: A small-farmer innovation in central Cameroon. *Agrofor. Syst.* **2012**, *86*, 493–504. [CrossRef]

43. Yemefack, M.; Ngendakumana, S.; Robiglio, V.; Assah, E.; Feudjio, T.P.M.; Ewane, N.N.; Magne, A.; Anne, M.; Minang, P.A.; Gyau, A.; et al. A Feasibility Study for Emission Reduction in the Efoulan Council, South Cameroon: A Project Design Document (PDD) 2013. Available online: http://humidtropics.cgiar.org/wp-content/uploads/downloads/2014/05/PDD-for-Efoulan-Council_Yemefack_final.pdf (accessed on 18 December 2020).
44. Gockowski, J.; Weise, S.; Sonwa, D.; Tchtat, M.; Ngobo, M. Conservation Because It Pays: Shaded Cocoa Agroforests in West Africa. *Habitat* **2004**, *29*. Available online: <https://hdl.handle.net/10568/103294> (accessed on 18 December 2020).
45. Saj, S.; Durot, C.; Mvondo Sakouma, K.; Tayo Gamou, K.; Avana-Tientcheu, M.L. Contribution of associated trees to long-term species conservation, carbon storage and sustainability: A functional analysis of tree communities in cacao plantations of Central Cameroon. *Int. J. Agric. Sustain.* **2017**, *15*, 282–302. [CrossRef]
46. Jagoret, P.; Kwesseu, J.; Messie, C.; Michel-Dounias, I.; Malézieux, E. Farmers' assessment of the use value of agrobiodiversity in complex cocoa agroforestry systems in central Cameroon. *Agrofor. Syst.* **2014**, *88*, 983–1000. [CrossRef]
47. Yemefack, M.; Rossiter, D.G.; Njomgang, R. Multi-scale characterization of soil variability within an agricultural landscape mosaic system in southern Cameroon. *Geoderma* **2005**, *125*, 117–143. [CrossRef]
48. Yemefack, M.; Njomgang, R.; Nounamo, L.; Rossiter, D.G. Quantified soil dynamics and spatial fragmentation within the shifting agricultural landscape in southern Cameroon. In Proceedings of the 19th World Congress of Soil Science, Soil Solutions for a Changing World, Brisbane, Australia, 1–6 August 2010; pp. 138–141.
49. Zhang, Y.; Chen, J.M.; Miller, J.R. Determining digital hemispherical photograph exposure for leaf area index estimation. *Agric. For. Meteorol.* **2005**, *133*, 166–181. [CrossRef]
50. Pfeifer, M.; Gonsamo, A.; Disney, M.; Pellikka, P.; Marchant, R. Leaf area index for biomes of the Eastern Arc Mountains: Landsat and SPOT observations along precipitation and altitude gradients. *Remote Sens. Environ.* **2012**, *118*, 103–115. [CrossRef]
51. Pfeifer, M.; Gonsamo, A. Manual to Measure and Model Leaf Area Index and Its Spatial Variability on Local and Landscape Scale; 2014; pp. 1–12. Available online: https://figshare.com/articles/journal_contribution/Manual_to_measure_and_model_leaf_area_index_and_its_spatial_variability_on_local_and_landscape_scale/928254 (accessed on 18 December 2020).
52. Beckschäfer, P. *Hemispherical_2.0-Batch Processing Hemispherical and Canopy Photographs with ImageJ-User Manual*; Georg-August-Universität Göttingen: Göttingen, Germany, 2015; pp. 1–6. [CrossRef]
53. Numbisi, F.N.; Van Coillie, F.M.B.; De Wulf, R. Delineation of Cocoa Agroforests Using Multiseason Sentinel-1 SAR Images: A Low Grey Level Range Reduces Uncertainties in GLCM Texture-Based Mapping. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 179. [CrossRef]
54. Small, D. Flattening Gamma: Radiometric Terrain Correction for SAR Imagery. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3081–3093. [CrossRef]
55. SNAP-ESA. SNAP-ESA Sentinel Application Platform. 2019. Available online: https://docplayer.net/55833385-Hemispherical_2-0-batch-processing-hemispherical-and-canopy-photographs-with-imagej-user-manual-by-philip-beckschafer-january-2015.html (accessed on 18 December 2020).
56. Nasirzadehdizaji, R.; Balik Sanli, F.; Abdikan, S.; Cakir, Z.; Sekertekin, A.; Ustuner, M. Sensitivity Analysis of Multi-Temporal Sentinel-1 SAR Parameters to Crop Height and Canopy Coverage. *Appl. Sci.* **2019**, *9*, 655. [CrossRef]
57. Bousbih, S.; Zribi, M.; Lili-Chabaane, Z.; Baghdadi, N.; El Hajj, M.; Gao, Q.; Mougenot, B. Potential of sentinel-1 radar data for the assessment of soil and cereal cover parameters. *Sensors* **2017**, *17*, 2617. [CrossRef] [PubMed]
58. Kumar, D.; Rao, S.; Sharma, J.R. Radar Vegetation Index as an Alternative to NDVI for Monitoring of Soyabean and Cotton. *Indian Cartogr.* **2013**, *33*, 91–96.
59. Haldar, D.; Dave, V.; Misra, A.; Bhattacharya, B. Radar Vegetation Index for assessing cotton crop condition using RISAT-1 data. *Geocarto Int.* **2020**, *35*, 364–375. [CrossRef]
60. Szigarski, C.; Jagdhuber, T.; Baur, M.; Thiel, C.; Parrens, M.; Wigneron, J.P.; Piles, M.; Entekhabi, D. Analysis of the Radar Vegetation Index and Potential Improvements. *Remote Sens.* **2018**, *10*, 1776. [CrossRef]
61. Wang, Y.; Dong, D. Retrieving forest stand parameters from SAR backscatter data using a neural network trained by a canopy backscatter model. *Int. J. Remote Sens.* **1997**, *18*, 981–989. [CrossRef]

62. Linderman, M.; Liu, J.; Qi, J.; An, L.; Ouyang, Z.; Yang, J.; Tan, Y. Using artificial neural networks to map the spatial distribution of understory bamboo from remote sensing data. *Int. J. Remote Sens.* **2004**, *25*, 1685–1700. [[CrossRef](#)]
63. Upreti, D.; Huang, W.; Kong, W.; Pascucci, S.; Pignatti, S.; Zhou, X.; Ye, H.; Casa, R. A Comparison of Hybrid Machine Learning Algorithms for the Retrieval of Wheat Biophysical Variables from Sentinel-2. *Remote Sens.* **2019**, *11*, 481. [[CrossRef](#)]
64. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
65. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer: New York, NY, USA, 2009; pp. 587–603. [[CrossRef](#)]
66. Zas, R.; Solla, A.; Sampedro, L. Variography and kriging allow screening *Pinus pinaster* resistant to *Armillaria ostoyae* in field conditions. *For. Int. J. For. Res.* **2007**, *80*, 201–209. Available online: <https://academic.oup.com/forestry/article-pdf/80/2/201/1376637/cpl050.pdf> (accessed on 18 December 2020). [[CrossRef](#)]
67. Sonwa, D.J.; Weise, S.F.; Nkongmeneck, B.A.; Tchatat, M.; Janssens, M.J.J. Structure and composition of cocoa agroforests in the humid forest zone of Southern Cameroon. *Agrofor. Syst.* **2016**, *91*, 451–470. [[CrossRef](#)]
68. Jagoret, P.; Snoeck, D.; Bouambi, E.; Ngnogue, H.T.; Nyassé, S.; Saj, S. Rehabilitation practices that shape cocoa agroforestry systems in Central Cameroon: Key management strategies for long-term exploitation. *Agrofor. Syst.* **2018**, *92*, 1185–1199. [[CrossRef](#)]
69. Woodhouse, I.H. *Introduction to Microwave Remote Sensing*; Taylor & Francis Group: Abingdon, UK, 2006; p. 370.
70. Ordway, E.M.; Asner, G.P.; Lambin, E.F. Deforestation risk due to commodity crop expansion in sub-Saharan Africa. *Environ. Res. Lett.* **2017**, *12*. [[CrossRef](#)]
71. Omar, H.; Misman, M.A.; Kassim, A.R. Synergetic of PALSAR-2 and Sentinel-1A SAR Polarimetry for Retrieving Aboveground Biomass in Dipterocarp Forest of Malaysia. *Appl. Sci.* **2017**, *7*, 675. [[CrossRef](#)]
72. Meyer, L.H.; Heurich, M.; Beudert, B.; Premier, J.; Pflugmacher, D. Comparison of Landsat-8 and Sentinel-2 Data for Estimation of Leaf Area Index in Temperate Forests. *Remote Sens.* **2019**, *11*, 1160. [[CrossRef](#)]
73. Beerli, O.; Netzer, Y.; Munitz, S.; Mintz, D.F.; Pelta, R.; Shilo, T.; Horesh, A.; Mey-tal, S. Kc and LAI Estimations Using Optical and SAR Remote Sensing Imagery for Vineyards Plots. *Remote Sens.* **2020**, *12*, 3478. [[CrossRef](#)]
74. Korhonen, L. Estimation of Boreal Forest Canopy Cover with Ground Measurements, Statistical Models and Remote Sensing. *Diss. For.* **2011**, *115*, 1–56. [[CrossRef](#)]
75. Korhonen, L.; Ali-Sisto, D.; Tokola, T. Tropical forest canopy cover estimation using satellite imagery and airborne lidar reference data. *Silva Fenn.* **2015**, *49*, 1405. [[CrossRef](#)]
76. Saatchi, S.; Marlier, M.; Chazdon, R.L.; Clark, D.B.; Russell, A.E. Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. *Remote Sens. Environ.* **2011**, *115*, 2836–2849. [[CrossRef](#)]
77. Schmullius, C.; Thiel, C.; Pathe, C.; Santoro, M. Radar Time Series for Land Cover and Forest Mapping. In *Remote Sensing and Digital Image Processing*; Kuenzer, C., Dech, S., Wagner, W., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 22, Chapter 16, pp. 323–356. [[CrossRef](#)]
78. Van Emmerik, T.H.M. Water Stress Detection Using Radar. Ph.D. Thesis, Technische Universiteit Delft, Delft, The Netherlands, 2017.
79. Borokini, T.I.; Onefeli, A.O.; Babalola, F.D. Inventory Analysis of *Milicia excelsa* (Welw C. C. Berg.) in Ibadan (Ibadan Metropolis and University of Ibadan), Nigeria. *J. Plant Stud.* **2012**, *2*. [[CrossRef](#)]
80. Tschardtke, T.; Clough, Y.; Bhagwat, S.A.; Buchori, D.; Faust, H.; Hertel, D.; Ho, D.; Jührbandt, J.; Kessler, M.; Perfecto, I.; et al. Multifunctional shade-tree management in tropical agroforestry landscapes—A review. *J. Appl. Ecol.* **2011**, *48*, 619–629. [[CrossRef](#)]
81. Vaast, P.; Somarriba, E. Trade-offs between crop intensification and ecosystem services: The role of agroforestry in cocoa cultivation. *Agrofor. Syst.* **2014**, *88*, 947–956. [[CrossRef](#)]
82. Fonkeng, E.E. *Cocoa Yield Evaluation and Some Important Yield Factors in Small Holder Theobroma cacao Agroforests in Bokito-Centre Cameroon*; Université De Dschang: Dschang, Cameroon, 2014.
83. Armathé, A.J.; Mesmin, T.; Unusa, H.; Soleil, B.R.A. A comparative study of the influence of climatic elements on cocoa production in two agrosystems of bimodal rainfall: Case of Ngomedzap forest zone and the contact area of forest-savanna of Bokito. *J. Cameroon Acad. Sci.* **2013**, *11*, 28–37.

84. Sonwa, D.J.; Coulibaly, O.; Weise, S.F.; Adesina, A.A.; Janssens, M.J. Management of cocoa: Constraints during acquisition and application of pesticides in the humid forest zones of southern Cameroon. *Crop Prot.* **2008**, *27*, 1159–1164. [[CrossRef](#)]
85. Blaser, W.J.; Oppong, J.; Hart, S.P.; Landolt, J.; Yeboah, E.; Six, J. Climate-smart sustainable agriculture in low-to-intermediate shade agroforests. *Nat. Sustain.* **2018**, *1*, 234–239. [[CrossRef](#)]
86. Korhonen, L.; Hadi.; Packalen, P.; Rautiainen, M. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. *Remote Sens. Environ.* **2017**, *195*, 259–274. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).