



Article

Identifying Coffee Agroforestry System Types Using Multitemporal Sentinel-2 Data and Auxiliary Information

Agustín Escobar-López¹, Miguel Ángel Castillo-Santiago^{1,*}, José Luis Hernández-Stefanoni²,
Jean François Mas³ and Jorge Omar López-Martínez^{4,5}

¹ El Colegio de la Frontera Sur, Departamento de Observación y Estudio de la Tierra, la Atmósfera y el Océano, Carretera Panamericana y Periférico Sur s/n, San Cristóbal de las Casas 29290, Chiapas, Mexico

² Centro de Investigación Científica de Yucatán A. C. Unidad de Recursos Naturales, Calle 43 #130, Mérida 97205, Yucatán, Mexico

³ Centro de Investigaciones en Geografía Ambiental, Universidad Nacional Autónoma de México, Morelia 58090, Michoacán, Mexico

⁴ Cátedra CONACYT, Avenida Insurgentes Sur 1582, Colonia Crédito Constructor, Benito Juárez, Mexico City 03940, Distrito Federal, Mexico

⁵ El Colegio de la Frontera Sur, Departamento de Agricultura, Sociedad y Ambiente, Avenida Centenario Km 5.5, Chetumal 77014, Quintana Roo, Mexico

* Correspondence: mcastill@ecosur.mx

Abstract: Coffee is one of the most important agricultural commodities of Mexico. Mapping coffee land cover is still a challenge because it is grown mainly on small areas in agroforestry systems (AFS), which are located in hard-to-access mountainous regions. The objective of this research was to map coffee AFS types in a mountainous region using the changing spectral response patterns over the dry season as well as supplementary data. We employed Sentinel-1, Sentinel-2 and ALOS-Palsar images, a digital elevation model, soil moisture layers, and 150 field plots. First, we defined three coffee AFS types based on their structural and spectral characteristics. Then, we performed a recursive feature elimination analysis to identify the most relevant predictor variables for each land use/cover class in the region. Next, we constructed a predictor variable dataset for each AFS type and one for the remaining land use/cover classes. Afterward, four maps were generated using a random forest (RF) classifier. Finally, we combined the four maps into a unique land-cover map through a maximum likelihood algorithm. Using a validation sample of 932 sites derived from Planet images (4.5 m pixel size), we estimated a 95% map overall accuracy. Two AFS types were classified as having low error; the third, with the highest tree density, had the lowest accuracy. The results obtained show that the infrared and near-infrared bands from the Sentinel-2 scenes are particularly useful for coffee AFS discrimination. However, supplementary data are required to improve the performance of the classifier. Our findings also highlight the importance of the multi-temporal and multi-dataset approach for identifying complex production systems in areas of high topographic heterogeneity.

Keywords: Sierra Madre; Chiapas; random forest; shade coffee; recursive feature elimination



Citation: Escobar-López, A.; Castillo-Santiago, M.Á.; Hernández-Stefanoni, J.L.; Mas, J.-F.; López-Martínez, J.O. Identifying Coffee Agroforestry System Types Using Multitemporal Sentinel-2 Data and Auxiliary Information. *Remote Sens.* **2022**, *14*, 3847. <https://doi.org/10.3390/rs14163847>

Academic Editor: Javier J Cancela

Received: 13 June 2022

Accepted: 4 August 2022

Published: 9 August 2022

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



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Coffee is one of the most important agroforestry crops in Latin America; 80% of the global production of Arabic coffee is grown in this region [1]. Given its importance, this crop has significantly modified the structure of rural landscapes in coffee-growing regions [2]. About 25 million rural farmers depend on coffee growing for their livelihoods; most of them are small farmers with crops ranging from 1 to 5 ha [3].

In Mexico, coffee is cultivated mainly in agroforestry systems (AFS), also known as shade coffee [4]. In AFS, coffee plants grow in the understory under the canopy of native or introduced tree species [5]. Shade trees, which can be either timber or fruit trees, regulate light conditions for optimal coffee growing, capture carbon, control soil erosion,

and provide shelter for biodiversity [6]. Additionally, these trees are a source of firewood, food, and additional economic income for farmers [7,8]. Coffee agroforestry systems are found in a wide variety of socio-environmental contexts with different anthropic alteration level, densities and compositions of shade trees, and involving several varieties of coffee plants adapted to various geographic regions [9]. These modifications have also been aimed at reducing the incidence of pests and diseases [10].

Coffee AFS can be characterized by a combination of attributes, such as the density of shade trees, the abundance of non-native species and the use of agrochemicals, which can be used to describe the levels of anthropogenic crop disturbance. In this regard, Toledo and Moguel [11] elaborated a proposal for a disturbance gradient with four classes. According to their proposal, the AFS with the lowest disturbance level, "rustic systems", are those where coffee cultivation is introduced into the mature native forest, either replacing or supplementing the vegetation in the understory. In contrast, the system with the higher levels of anthropic impact, or "shade monocultures", are those where coffee plants are established under trees of a single species and often require the application of agrochemicals; these systems provide fewer environmental services [12]. Finally, these authors also included a fifth and the highest disturbance class, corresponding to coffee crops without shade trees (i.e., not an AFS), also known as sun coffee. In Mexico, coffee production is concentrated in four states: Chiapas, Veracruz, Oaxaca, and Puebla. The first covers 35% of the total area cultivated in Mexico, or approximately 252,000 ha, with 90% involving some type of AFS [4,13]. However, the incidence of pests and diseases, particularly coffee rust (*Hemileia vastatrix*), has forced many farmers to replace their old coffee varieties with more resistant ones that require fewer shade trees, thus increasing the anthropization of coffee-growing systems [14].

Despite the importance of shade coffee production in Chiapas, the detailed spatial distribution of the different types of coffee agroforestry systems is still unknown. This lack of information is caused by the poor accessibility to coffee plantations since a significant number of coffee growers farmers are located in highly marginalized areas, with small coffee plot areas (about one hectare) scattered in hard-to-access mountain landscapes [15].

Mapping the area covered by coffee AFS using remote sensors has been unsuccessful, particularly regarding systems with a high density of shade trees, which have been identified with low accuracy [16,17]. Topographic heterogeneity, AFS structural complexity, and shade tree coverage are some of the factors that restrain correct identification of these AFS. As a result of the complexity of coffee-growing landscapes, their spectral patterns are often misidentified by other types of land cover, such as forests and secondary vegetation [17,18].

In the case of sun coffee plantations, which have few shade trees, if any, land cover classification has been performed well. Accuracy values above 90% have been reported in plantations with reduced vegetation cover, by combining spectral bands and vegetation indices from optical images [19] and including texture metrics as predictive variables [20].

In situations of intermediate complexity, using high-resolution images in areas with homogeneous topography, Hebbar et al. [21] identified commercial poly- and monocultures with low error. In more complex AFS, the use of supplementary information, including slope, temperature, precipitation, and soil fertility, improved the accuracy of crop identification [22]. Separately, Kelley et al. [23] used spectral indices and land surface temperature derived from multi-seasonal Landsat 8 imagery to detect coffee AFS with a percentage of shade trees above 30%.

On the other hand, the modification of AFS spectral patterns associated with phenological changes has been little explored. Few studies have attempted to discriminate the types of AFS using temporal variations of their spectral response [23–26].

Considering the importance of the shade coffee plantations for Chiapas México and the difficulties in identifying them, the aim of this study was twofold. First, we sought to develop a method for the identification of coffee AFS types with different densities of shade trees, using variations in spectral response throughout the dry season as well as spectral

indices and auxiliary data. We also expected to improve the accuracy of land cover maps of coffee of agroforestry systems of Chiapas using this approach.

2. Materials and Methods

2.1. Study Area

The study area stretches across 2381 km² located in the central part of the mountain range called Sierra Madre de Chiapas (Figure 1). This region harbors a high biodiversity and has the highest coffee production in the state of Chiapas [27]. The study area was limited to potential coffee growing areas, i.e., those between 700 and 2800 m above sea level using a digital elevation model downloaded from the Digital Library of Maps of the National Institute of Statistics, Geography, and Informatics [28].

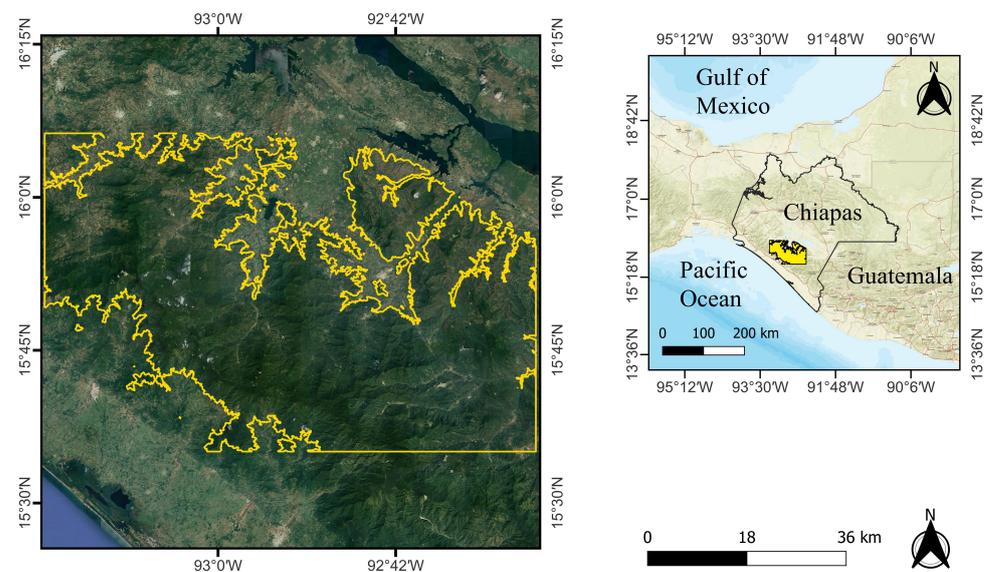


Figure 1. Location of the study area. Source: World Imagery, ESRI, Copyright: © 2022 Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community.

Due to its mountainous relief, the area covers a wide altitudinal gradient, from 700 to 2700 m above sea level. In the dry season, the minimum temperature fluctuates between 9 and 15 °C and the maximum temperature between 21 and 33 °C, with precipitation between 25 and 300 mm [29]. The region encompasses a significant diversity of forest types, including coniferous forest, mountain cloud forest, tropical dry forest and tropical wet forest. In some cases, these show degradation and fragmentation of the forest cover as a result of illegal logging, livestock ranching, and rain-fed agriculture. As a result, fragments of secondary forests in different successional stages are also common [30].

2.2. Field Data and Characterization of Agroforestry Systems

The characterization of coffee AFS types and the calibration of the models were conducted using an inventory of 263 shade coffee plots collected in 2019 by the *Café de la Concordia* (CAFECO, for its acronym in Spanish) cooperative organization of coffee growers. In each plot, CAFECO recorded information on coffee and the geographic coordinates of the plot center, the number and varieties of coffee plants, and the number and botanical names of the shade trees. We also collected similar data in an additional 15 plots; we also registered the mean height, mean diameter at the breast height, crown diameter and the number of strata of coffee plants and shade trees.

According to an analysis of field data (abundance and species composition of shade trees, and density of coffee plants) and a visual interpretation of satellite images, we generated a dataset of 150 field plots and grouped into three classes of AFS (Figure 2).

The definition of these classes partially matches that proposed by Moguel and Toledo [31], adapted to the particular characteristics of the local production systems in the study area. The definitions of coffee AFS types used in this study are the following:

1. *Reduced-shade coffee polyculture.* This system has one or two tree strata and an understory layer with coffee plants. The highest stratum shows some trees of natural vegetation, usually the tallest trees (>7 m). When there is an intermediate stratum, it usually includes fruit and timber tree species; the most common species are *naranja* (*Citrus × sinensis* (L.) Osb.), *aguacate* (*Persea americana* Mill.), *plátano* (*Musa paradisiaca* L.), *zapote* (*Pouteria sapota* (Jacq.) H.E. Moore and Stearn), and *mango* (*Mangifera indica* L.). The percent shade is generally less than or equal to 30%, the tree density varies from 16 to 30 trees per hectare, and the density of coffee plants ranges from 2500 to 4400 plants per hectare. Although trees are not evenly distributed, the distance between shade trees is usually wide, so there are open areas, and coffee plants are frequently apparent in high-resolution images (Figure 2b).
2. *Rustic coffee polyculture.* This AFS has two tree strata and an understory of coffee plants. The highest stratum includes trees of natural vegetation, in some cases alternating with introduced timber trees, mainly *cedro* (*Cedrela odorata* L.) and *roble* (*Quercus robur* L.); the average height of this stratum is 10 m. The second stratum generally comprises introduced species with a mean height of 6 m, commonly *chalum* (*Inga vera* Willd.), *caspirol* (*Inga laurina* (Sw.) Willd.), *paterna* (*Inga spuria* H and B. Ex Willd.), and fruit trees such as *naranja* (*Citrus × sinensis* (L.) Osb.), among others. The percent shade ranges from 30% to 60%. The density of shade trees varies between 24 and 38 trees/ha, and the density of coffee plants, between 2500 and 4800 plants/ha. In satellite images, these systems appear more homogeneous in color compared to forests and tropical forests and are less fragmented (Figure 2c) and less intensely colored than reduced-shade polycultures.
3. *Rustic coffee.* This system also has one or two strata of tree vegetation; the highest stratum is dominated by species of natural vegetation, with occasional introduced trees. The intermediate stratum consists mainly of timber trees, including *chalum* (*Inga vera* Willd.), *caspirol* (*Inga laurina* (Sw.) Willd.) and *paterna* (*Inga spuria* H and B. Ex Willd.) of lower height. The percent shade is greater than 60%. Compared to the other AFS classes, this class has a higher density of shade trees (30–44 trees/ha), with a similar density of coffee plants (2500–3300 plants/ha). This system is the one leading to greater spectral confusion with forests and tropical forests because the three show similar tonalities and texture patterns (Figure 2d).

The study area has no sun coffee plantations but includes other types of land cover, such as mature forests, disturbed forests, shrub and herbaceous secondary vegetation, human settlements, pastures, oil palm (*Elaeis guineensis* Jacq.), mango plantations and bare soil. These land covers were classified into three groups. The final classes used for a land cover map with coffee plantations are shown in Table 1. The sixth group, *other classes*, includes human settlements, agriculture and bare soil.

Table 1. AFS and land-cover classes in the study area.

ID	Classes
1	Reduced-shade coffee polyculture
2	Rustic coffee polyculture
3	Rustic coffee
4	Mature forests
5	Disturbed forests
6	Other classes

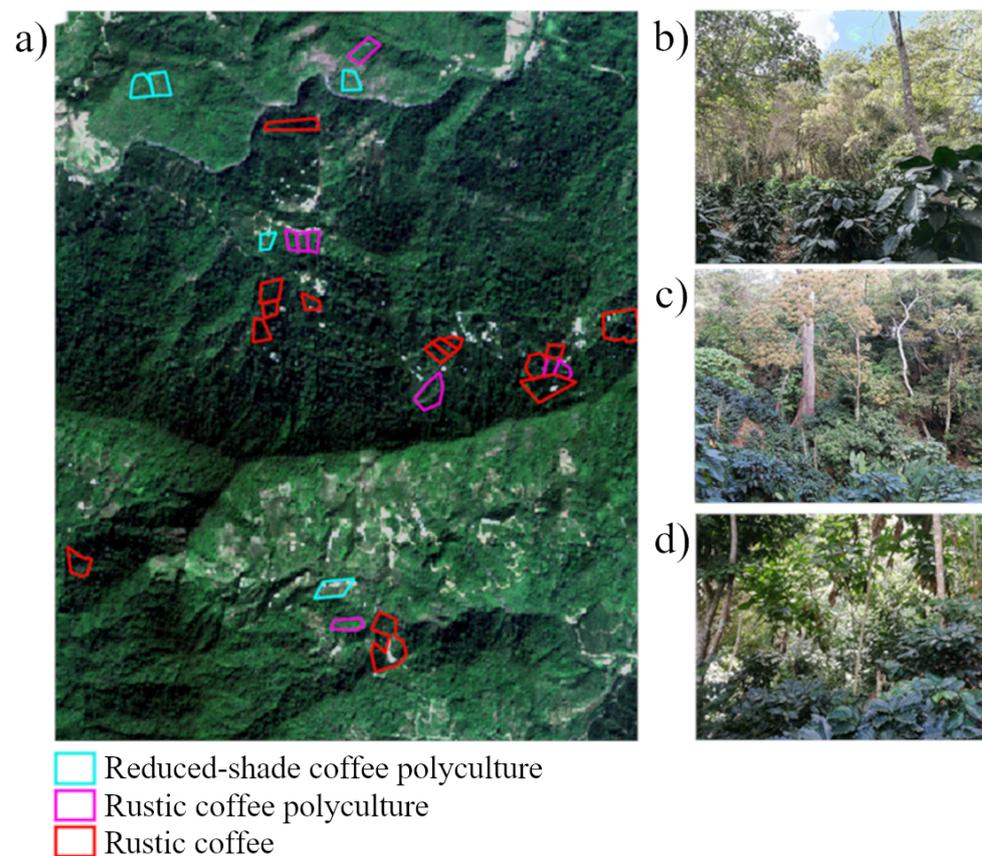


Figure 2. Coffee agroforestry systems defined for the study area. (a) Sample polygons of each type of coffee agroforestry system, (b) reduced-shade coffee polyculture, (c) rustic coffee polyculture, (d) rustic coffee.

2.3. Imagery and Auxiliary Data

To reduce the presence of clouds and haze, we selected optical satellite imagery from January to May 2019 only, corresponding to the dry season. We used a monthly time series of five Sentinel-2 images (Level 1C); two Sentinel-2 scenes covered the whole study area. We also analyzed five Sentinel-1 (Interferometric Wide Swath Level 1) radar images with VV and VH polarization acquired on similar dates as the optical data and one Alos Palsar with HH and HV polarization from the JAXA Earth Observation Research Center (<https://www.eorc.jaxa.jp> (accessed on: 7 January 2019)). Satellite imagery were downloaded from the USGS Earth Explorer website (<https://earthexplorer.usgs.gov/> (accessed on: 10 January 2019)). The auxiliary data employed consist of a digital elevation model (DEM) and a climatic data set. A 30 m pixel size DEM was downloaded from the Digital Library of Maps of the National Institute of Statistics, Geography, and Informatics (Instituto Nacional de Estadística, Geografía e Informática) [28]. We used the following climatic data: monthly temperature (minimum and maximum), monthly precipitation, and soil moisture. These variables were interpolated from meteorological data and prepared in raster layers with a pixel size of 120 m (for further details, refer to Hernández-Stefanoni et al. [32]).

2.4. Image Processing

Using the software SNAP [33], atmospheric corrections were applied to all optical images to reduce the potential effect of water vapor and obtain bottom-of-atmosphere (BOA) reflectance values. Bands were resampled using the Sen2Res algorithm to match the pixel size to 10 m.

To highlight any changes in the spectral response of AFS during the dry season, the following six vegetation indices were calculated for each of the five months analyzed.

The *chlorophyll vegetation index* (CVI) and *modified simple ratio for vegetation* (MSR) highlight information associated with chlorophyll content [34]. The combination of the *modified chlorophyll absorption in reflection index* (MCARI) and the *optimized soil adjusted vegetation index* (OSAVI) reduce background reflectance and improve sensitivity to variability in the leaf chlorophyll content [35]; this combination is especially useful for reducing the reflectance of non-photosynthetic components and soil [34]. In addition, we calculated the *Beison Datt vegetation index* (DATT), the *RGB intensity*, and the *soil background line* (SBL) to evaluate more efficient alternatives for estimating canopy attributes and color saturation in RGB composites, and to discriminate between soil and vegetation cover (Table 2).

We preprocessed the Sentinel-1 images using the standard generic workflow available in SNAP. This workflow applies a precise orbit of acquisition, removes thermal and edge noise, and performs radiometric calibration and geometric terrain correction [36]. The layers used are shown in Table 3.

Table 2. Equations of the vegetation indices used in this study. R_i = Reflectance in range i .

Vegetation Index	Equation	Reference
CVI	$R_{842} \frac{R_{665}}{R_{560}^2}$	(1) [37]
MSR	$\frac{R_{800} - R_{445}}{R_{680} - R_{445}}$	(2) [38]
MCARI/OSAVI	$\frac{[(R_{700} - R_{670}) - 0.2(R_{700} - R_{550})] (\frac{R_{700}}{R_{670}})}{(1 + 0.16) \frac{(R_{800} - R_{670})}{R_{800} + R_{670} + 0.16}}$	(3) [35]
DATT	$\frac{R_{850}}{R_{550} \times R_{708}}$	(4) [38]
RGB Intensity	$(\frac{1}{30.5})(R_{0.490} + R_{0.560} + R_{0.665})$	(5) [39]
SBL	$R_{945} - 2.4 \times R_{0.665}$	(6) [40]

2.5. Data Analysis and Land Cover Classification

To aim in the selection of explanatory variables, we plotted the monthly values of vegetation indices and spectral bands for each type of AFS. This preliminary analysis revealed that some AFS showed changes in spectral patterns in certain months, i.e., bigger separability between index values/spectral bands, so we tried to keep these variables/months in the following stages. Then, a correlation analysis was performed with all possible explanatory variables, including spectral bands, vegetation indices and auxiliary data; the Spearman index was used to detect possible cases of non-linear correlation. At this stage, all highly correlated variables were eliminated, i.e., those with a Spearman index greater than 0.85. With the resulted set of predictor variables, four subsets were constructed—one for each AFS type and one for the rest of land-cover types in the study area. The recursive feature elimination (RFE) algorithm was used to select the best subset of predictors for each model [41] under the criterion of mean decrease in accuracy (MDA). The MDA was calculated through random permutation of the input variables, and the decrease in the accuracy of the resulting prediction was assessed [42]. Each tested model included reflectance and

backscatter data from optical and radar images, vegetation indices, and supplementary data (soil moisture, altitude, temperature and precipitation).

Table 3. Data used for the identification of coffee agroforestry systems.

Data Type	Sources	Date	Input Variables
Optical	Sentinel-2 (Dry Season)	01/23/2019	Reflectance bands:
		02/20/2019	Band 2—Blue
		03/24/2019	Band 3—Green
		04/23/2019	Band 4—Red
		05/24/2019	Band 5—Red edge
			Band 7—Red edge
			Band 8—NIR
			Band 8A—Red edge
			Band 9—Water vapour Band 11—SWIR
			Vegetation indices:
			CVI
			MSR
			MCARI/OSAVI
			DATT
			SBL
			RGB Intensity
Auxiliary data	DEM		Altitude
	Climatic data		Mean monthly soil moisture (January, February, March, April, May)
			Mean monthly temperature (January, February, March, April, May)
			Mean monthly precipitation (January, February, March, April, May)
Radar	Sentinel-1A (Dry Season)	02/12/2019	Beam mode: IW
		03/24/2019	Polarization: VV + VH
		05/08/2019	Band: C-Band
			Spatial resolution: 20 m
			Ascending
	Sentinel-1B (Dry Season)	01/25/2019	Beam mode: IW
04/23/2019		Polarization: VV + VH	
			Band: C-Band
			Spatial resolution: 20 m
			Ascending
	Alos PALSAR (Dry Season)	2019	Beam mode: FBD
			Polarization: HH + HV
			Band: L-Band
			Spatial resolution: 25 m

The parameters of the best sets of predictive variables defined for each model were adjusted using the random forest algorithm. Through the cross-validation of ten interactions, the *ntree* (number of trees to grow) and *mtry* (number of variables randomly sampled as candidates in each division) parameters were chosen for the selection of the best classification model.

Once the models were calibrated, the four respective classifications were obtained, which were spatially overlaid to identify conflicting pixels, i.e., those assigned to more than one type of AFS or land cover class. The class to which each conflicting pixel should belong was assigned using the maximum likelihood algorithm; Sentinel-2 scenes and the training areas previously were used as input data. The resulting classifications were combined to generate a single land cover map with the different types of AFS identified. The overall methodological outline is shown in Figure 3.

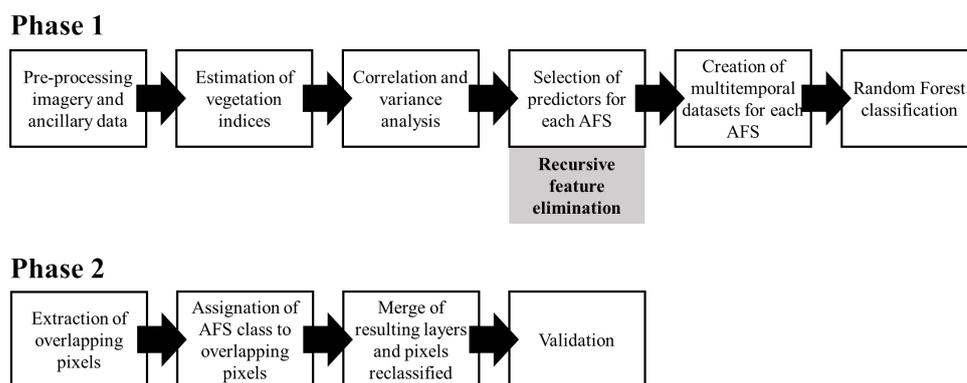


Figure 3. Methodology for the identification of AFS classes.

2.6. Map Validation

We followed the recommendations of Oloffson et al. [43] to validate the final map. A stratified random sampling design was used to estimate the size and distribution of the reference sample. The number of validation sites in each land-cover class was allocated proportionally to their size; small classes were assigned at least 50 sites. Thematic accuracy statistics were derived according to the equations described in Oloffson et al. [44].

The ground truth of the reference data was obtained by visual interpretation of high spatial resolution imagery; Planet image mosaics (4.7 m pixel size) downloaded from the NICFI Satellite Data Program (<https://www.planet.com/nicfi/> (accessed on: 25 January 2019)) and satellite data from ESRI's World Imagery platform were used. This platform provides 1 m resolution images worldwide (accessed on: 11 January 2019) [45].

3. Results

3.1. Selecting Predictors and Applying the Classification Model

The subsets defined for each model from the implementation of the RFE algorithm are shown in Table 4. The two January NIR bands functioned well as predictive variables for the three AFS types but played no significant role in the model for other land-cover types. Green and red-edge bands are important for the two AFS with the highest tree density, but their role is taken up by the SWIR band in the model with the lowest tree density (reduced-shade polyculture). In all models, only the first four or five variables have high predictive importance. However, eliminating any variable with low importance in the models increases the error in the resulting classified map. In Figure 4, we show the predictors' importance for each model. Note that none of the models contain variables derived from radar imagery.

Table 4. Predictors selected for each classification process. NIR = Near infrared, SWIR = Short-wave infrared, Bi = band i.

Class	Predictors Selected Using RFE
Reduced-shade coffee polyculture	NIR B8 (January), NIR B8A (January), SWIR B11 (April, May), CVI (January), MSR (January)
Rustic coffee polyculture	Red edge B7 (January), NIR B8 (January), NIR B8A (January), Green B3 (January), MCARI/OSAVI (January), Soil humidity (January)
Rustic coffee	Red edge B7 (January), NIR B8 (January), NIR B8A (January), Green B3 (January), Red edge B5 (January), RGB Intensity (January)
Mature forests, disturbed forests and other classes	Blue B2 (January, February, March), SWIR B11 (May), RGB Intensity (February), DATT (January), SBL (January)

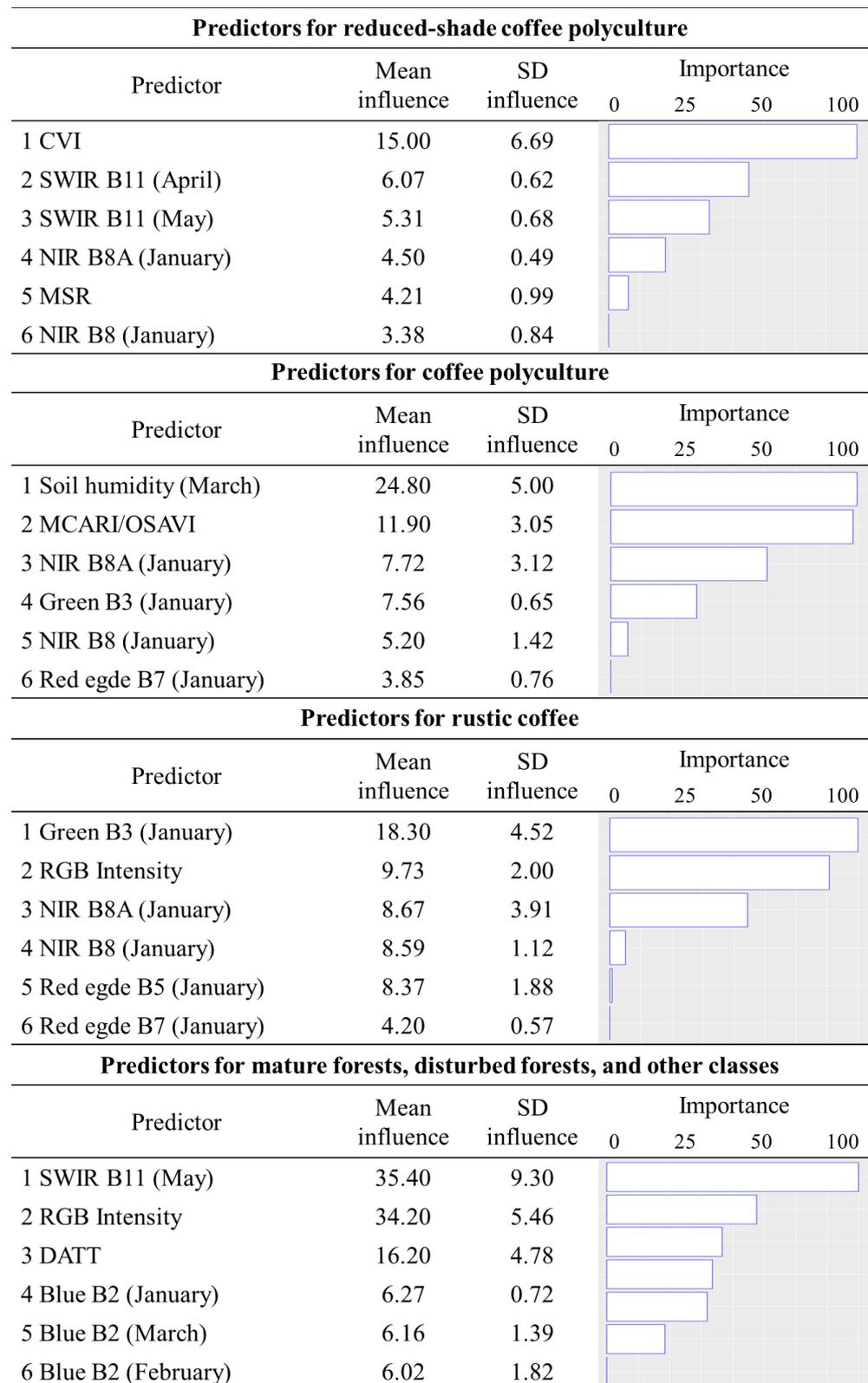


Figure 4. Predictors and their influence for each class in the RF classification. SD = standard deviation.

The optimal parameters for the random forest classification models ($mty = 3$ and $n\text{tree} = 500$) achieved accuracy values above 98% for the three AFS models and above 99% for the other land-cover classes.

The individual maps for reduced-shade polycultures, rustic polycultures, rustic coffee, and other classes are shown in Figure 5.

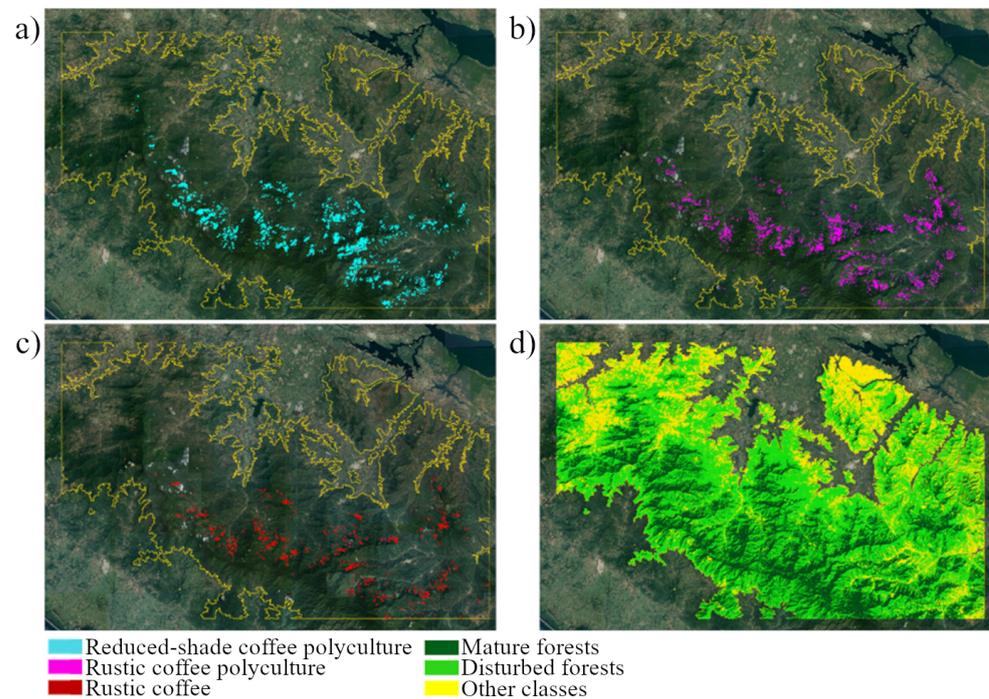


Figure 5. Individual maps for (a) reduced-shade coffee polyculture, (b) rustic coffee polyculture, (c) rustic coffee, and (d) mature forest, disturbed forest and other classes.

The conflicting pixels identified from class overlaying accounted for 0.71% of the total study area. The disturbed forest was the class that showed the greatest misidentification with other classes, mainly with rustic polycultures and rustic coffee.

Reduced-shade polycultures shade was frequently located in areas adjacent to human settlements and areas with agricultural or livestock activities, between 700 and 1500 m above sea level. Rustic polycultures and introduced coffee were identified mainly in disturbed forests areas near mature forests and tropical forests at altitudes above 1400 m a.s.l. However, rustic coffee plantations were located near disturbed forests with a higher tree density. Of the total area identified as an agroforestry coffee system, reduced-shade polycultures were the systems that comprised the largest area (40%), while rustic coffee plantations encompassed the smallest area (23%); the extent occupied by AFS was considerably smaller compared with mature and disturbed forests.

The information in the infrared range of the electromagnetic spectrum contributed significantly to differentiation between disturbed forest and coffee AFS because the coffee flowering phase of the phenological cycle produces an apparent change in the red edge and SWIR bands in the spectral signature of coffee AFS (Figure 6).

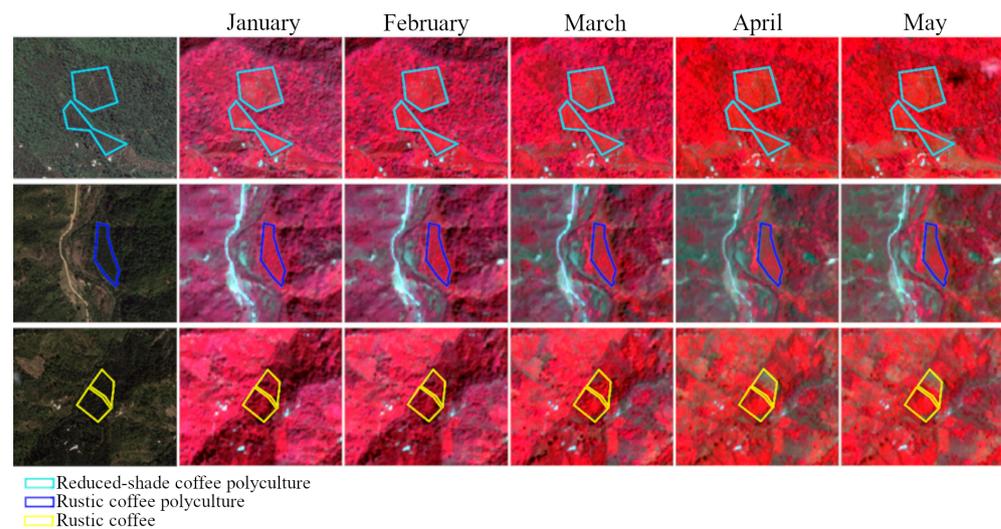


Figure 6. Spectral profile variations for coffee agroforestry systems in color infrared (B8, B4, B3) from Sentinel-2.

3.2. Model Validation

The map resulting from the overlay of individual classifications is shown in Figure 7. Table 5 shows the confusion matrix obtained from the reference sample, and Table 6 shows their respective accuracy statistics. The global accuracy of the map was 95.04%. The AFS identified with less error are the reduced-shade coffee polyculture and rustic coffee polycultures, which involve a more significant anthropic influence within the proposed AFS classification. Reduced-shade coffee polycultures showed only two instance of a validation site misclassified as rustic coffee polyculture. Rustic coffee polycultures are misclassified either as reduced-shade coffee polyculture or rustic coffee, mainly at sites that transition from one AFS to another, whereas the rustic coffee class was confused a couple of times with disturbed forest and rustic coffee polyculture. Two of the three AFS were well identified, with high accuracy (>92%), but the third, rustic coffee, was the class with the most significant error. This error is mainly due to their spectral similarity with disturbed forests.

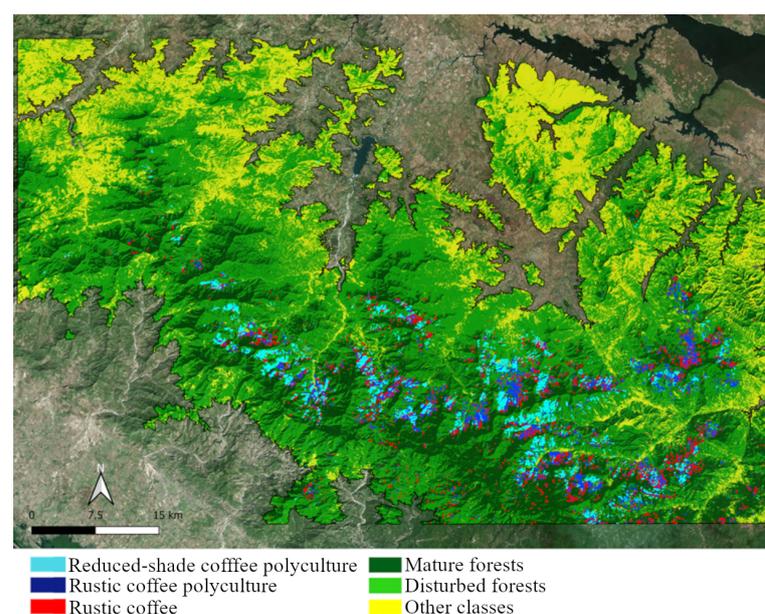


Figure 7. Final map after to overlap all models classifications.

Table 5. Confusion matrix generated using 932 validation sites.

Reference Prediction	Reduced-Shade Coffee Polyculture	Rustic Coffee Polyculture	Rustic Coffee	Mature Forest	Disturbed Forest	Other Classes	Total
Reduced-shade coffee polyculture	48	2	0	0	2	0	52
Rustic coffee polyculture	2	47	2	0	0	0	51
Rustic coffee	0	1	46	0	2	0	49
Mature forest	0	0	0	50	2	0	52
Disturbed forest	0	0	2	1	525	7	535
Other classes	0	0	0	0	29	164	193
Total	50	50	50	51	560	171	932

Table 6. Accuracy statistics, where CI = confidence interval, PA = producer accuracy, UA = user accuracy, OA = Overall accuracy, \pm = variance.

	Area Per Class (km ²)	Area Estimated Per Class (km ²)	CI of Estimated Area (km ² \pm)	PA (%)	UA (%)	OA (%)
Reduced-shade coffee polyculture	26.82	26.47	3.06	93.55	92.31	95.04
Rustic coffee polyculture	43.56	42.09	3.96	95.39	92.16	
Rustic coffee	44.56	52.03	12.37	80.39	93.88	
Mature forest	208.80	205.02	13.81	97.93	96.15	
Disturbed forest	2271.90	2344.68	45.20	95.08	98.13	
Other classes	694.56	619.93	41.38	95.20	84.97	

4. Discussion

To improve the accuracy of coffee agroforestry maps, we used a strategy of adjusting separate models for each AFS type allowed, using a small and specific number of predictor variables for each land cover class, without affecting their performance. The resulting map showed a global accuracy of 95.04%. The incorporation of different vegetation indices for each AFS type increased the accuracy of each model; in other words, the use of different sets of predictor variables was more efficient in discriminating agroforestry systems with different shade densities. Although a potential disadvantage of this approach is the presence of conflicting pixels, particularly in land-cover classes that were not clearly differentiated spectrally, in the present study, those pixels represented only 0.7% of the study area and mostly corresponded to AFS with high tree density and disturbed forests.

In addition, although some variables were repeated in several models, the explanatory importance of each varied across the four models. Vegetation indices were key for discriminating AFS with lower tree density; high values of the CVI and MSR vegetation indices in the coffee flowering phases contrasted with the response of the disturbed forest. On the other hand, topographic or climatic data did not play a significant role in the models (except for one), probably due to their low spatial resolution.

Although it is difficult to compare the level of success of this work with that of other studies (due to the diversity of AFS types, landscape complexity or data sources), broadly speaking, the accuracies obtained in this study are moderately higher than those reported under similar conditions. In highly heterogeneous landscapes, the present study improved the accuracy of open-canopy coffee AFS (canopy opening greater than 60%) and closed-canopy coffee AFS (canopy opening between 20 to 60%) identification by approximately 30% [22,25,46]. Compared with studies using radar images, this research achieved greater accuracy in identifying reduced-shade polycultures versus AFS with similar characteristics, such as commercial polycultures or coffee plantations with a low tree coverage [47]. Our methodological approach also was efficient for differentiating AFS within forest landscapes, reported as a common issue when using conventional satellite images [48,49].

It should be noted that in the present study, in addition to optical images, we also tested Sentinel-1 and an ALOS-PALSAR images. However, the REF algorithm consistently eliminated them from the set of predictive variables since they provided little information. The topographic complexity of the study area appeared to be an obstacle for the use of radar data.

Usually, the flowering of coffee plants occurs at the end of the dry season [50]; this phenomenon causes changes in the concentrations of chlorophyll *a* and *b* and the leaf area index [51,52]. These phenological changes represent the leading cause of changes in spectral patterns in coffee-growing areas, consistent with the results of the studies by Bernardes et al. [50] and Júnior et al. [53], that addressed the relationship of coffee production at different stages of the cycle and the variation in vegetation indices over several years. Regarding our results, one model uses spectral data from the beginning and the end of the study period (January and April–May). This model corresponds to the AFS with few shade trees, so it is probably the one that is capturing the phenological changes of the coffee plants.

Accurate mapping of coffee agroforestry systems is essential for understanding the level of anthropic disturbance and change in the plant cover of coffee production areas.

According to our results, at least 50% of the AFS areas is heavily anthropized. This finding is consistent with the trends in other studies that reported changes in management practices to fight coffee rust, which generally involved replacing coffee varieties and reducing shade-tree density [54].

5. Conclusions

This study used a differentiated analysis approach by type of AFS to map coffee production in areas of high heterogeneity. The results reported herein are significant given the limitations highlighted in previous studies of the conventional use of remote sensing data for mapping coffee agroforestry systems and the accuracy levels reported for similar implementation contexts. The accurate identification of AFS contributes to the knowledge of the anthropic disturbance dynamics associated with coffee production by highlighting three different levels of landscape alteration for agroforestry practices. In this sense, subsequent studies may use this same approach in other coffee-growing areas in the state of Chiapas to explore the existence of systems with different characteristics than the AFS described herein and to evaluate the replicability of the method under different landscape characteristics. On the other hand, future studies may use specific subsets for other types of plant cover or land use, such as mature and disturbed forests, to further improve the accuracy of the resulting classification.

Author Contributions: Conceptualization: M.Á.C.-S., A.E.-L.; methodology: A.E.-L., M.Á.C.-S.; software: A.E.-L.; validation: M.Á.C.-S., A.E.-L., J.L.H.-S., J.F.M., J.O.L.-M.; formal analysis: A.E.-L.; investigation: A.E.-L., M.Á.C.-S.; Resources: M.Á.C.-S., A.E.-L., J.L.H.-S., J.F.M., J.O.L.-M.; writing—original draft preparation: A.E.-L.; writing—review and editing: M.Á.C.-S., J.L.H.-S., J.F.M., J.O.L.-M.; supervision: M.Á.C.-S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by CONACYT, grant number 2019-000002-01NACF-07726.

Data Availability Statement: The Sentinel-2 data used in this study were obtained from <https://earthexplorer.usgs.gov/> (accessed on: 10 January 2019). The Digital Elevation Model used can be downloaded from the Digital Library of Maps of INEGI (<https://www.inegi.org.mx/app/mapas/?tg=1015>) (accessed on: 7 January 2019). The Sentinel-1 data can be downloaded from <https://search.asf.alaska.edu/>. Alos Palsar data can be found in https://www.eorc.jaxa.jp/ALOS/en/palsar_fnf/data (accessed on: 7 January 2019). Climatic data were shared by Departamento de Observación y Estudio de la Tierra, la Atmósfera y el Océano in ECOSUR.

Acknowledgments: We want to thank the organization CAFECO for the assistance in the field trips in collecting data from the coffee plots and for sharing their information and knowledge about coffee cultivation in the study area. We also want to thank our colleagues in the Departamento de Observación y Estudio de la Tierra, la Atmósfera y el Océano in ECOSUR for sharing geospatial data and their comments on an earlier version of the manuscript. We also wish to thank two anonymous reviewers whose comments enriched this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. ITC. *The Coffee Guide*; Technical Report; The International Trade Centre (ITC): Geneva, Switzerland, 2021.
2. Perfecto, I.; Jiménez-Soto, M.E.; Vandermeer, J. Coffee Landscapes Shaping the Anthropocene. *Curr. Anthropol.* **2019**, *60*, S236–S250. [[CrossRef](#)]
3. Harvey, C.A.; Martínez-Rodríguez, M.R.; Cárdenas, J.M.; Avelino, J.; Rapidel, B.; Vignola, R.; Donatti, C.I.; Vilchez-Mendoza, S. The Use of Ecosystem-based Adaptation Practices by Smallholder Farmers in Central America. *Agric. Ecosyst. Environ.* **2017**, *246*, 279–290. [[CrossRef](#)]
4. CEDRSSA. *Reporte del café en México: Diagnóstico y perspectiva*; Technical Reports; Cámara de Diputados: Ciudad de México, Mexico, 2018.
5. Farfán-Valencia, F. *Descripción de la Estructura del Dosel Arbóreo al Interior de un Sistema Agroforestal con Café*; Technical Report; Centro Nacional de Investigaciones de Café, Manizales: Caldas, Colombia, 2019.
6. Jose, S. Environmental Impacts and Benefits of Agroforestry. In *Oxford Research Encyclopedia of Environmental Science*; Oxford University Press: Oxford, UK, 2019. [[CrossRef](#)]
7. Ospina, C. *Climate and Economic Benefits of Agroforestry Systems*; Climate Institute: Washington, DC, USA, 2017; p. 11.
8. DaMatta, F.; Nelson, R. Sustainable Production of Coffee In Agroforestry Systems in the Neotropics: An Agronomic and Ecophysiological Approach. *Agron. Colomb.* **2007**, *1*, 113–123.
9. Montagnini, F.; Somarriba, E.; Murgueitio, E.; Fassola, H.; Eibl, B. Función de los sistemas agroforestales en la adaptación y mitigación del cambio climático. In *Sistemas agroforestales. Funciones productivas, socioeconómicas y ambientales*; CATIE: Turrialba, Costa Rica, 2015; pp. 269–298.
10. Harvey, C.A.; Pritts, A.A.; Zwetsloot, M.J.; Jansen, K.; Pulleman, M.M.; Armbrrecht, I.; Avelino, J.; Barrera, J.F.; Bunn, C.; García, J.H.; et al. Transformation of Coffee-growing Landscapes Across Latin America. A Review. *Agron. Sustain. Dev.* **2021**, *41*, 62. [[CrossRef](#)] [[PubMed](#)]
11. Toledo, V.M.; Moguel, P. Coffee and Sustainability: The Multiple Values of Traditional Shaded Coffee. *J. Sustain. Agric.* **2012**, *36*, 353–377. [[CrossRef](#)]
12. Jha, S.; Bacon, C.M.; Philpott, S.M.; Rice, R.A.; Méndez, V.E.; Läderach, P. *A Review of Ecosystem Services, Farmer Livelihoods, and Value Chains in Shade Coffee Agroecosystems*; Springer: Berlin, Germany, 2011; pp. 141–208. [[CrossRef](#)]
13. INCAFECH. Datos importantes del café.
14. Valencia, V.; García-Barrios, L.; Sterling, E.J.; West, P.; Meza-Jiménez, A.; Naeem, S. Smallholder response to environmental change: Impacts of coffee leaf rust in a forest frontier in Mexico. *Land Use Policy* **2018**, *79*, 463–474. [[CrossRef](#)]
15. Higuera-Ciapara, I.; Rivera-Ramírez, J. *Chiapas: Problemáticas del Sector Cafetalero*; Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco: Guadalajara, Mexico, 2018.
16. Boell, M.; Ramos Alves, H.; Volpato, M.; Ferreira, D.; Lacerda, W. Exploiting Feature Extraction Techniques for Remote Sensing Image Classification. *IEEE Lat. Am. Trans.* **2018**, *16*, 2657–2664. [[CrossRef](#)]
17. Hunt, D.A.; Tabor, K.; Hewson, J.H.; Wood, M.A.; Reymondin, L.; Koening, K.; Schmitt-Harsh, M.; Follett, F. Review of Remote Sensing Methods to Map Coffee Production Systems. *Remote Sens.* **2020**, *12*, 2041. [[CrossRef](#)]
18. Souza, C.G.; Carvalho, L.; Aguiar, P.; Arantes, T.B. Algoritmos de Aprendizagem de Máquina e Variáveis de Sensoramento Remoto para o Mapeamento da Cafeicultura. *Bol. De Ciências Geodésicas* **2016**, *22*, 751–773. [[CrossRef](#)]
19. Schmitt-Harsh, M.; Sweeney, S.P.; Evans, T.P. Classification of Coffee-Forest Landscapes Using Landsat TM Imagery and Spectral Mixture Analysis. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 457–468. [[CrossRef](#)]
20. Mosomtai, G.; Odindi, J.; Abdel-Rahman, E.M.; Babin, R.; Fabrice, P.; Mutanga, O.; Tonnang, H.E.Z.; David, G.; Landmann, T. Landscape Fragmentation in Coffee Agroecological Subzones in Central Kenya: A Multiscale Remote Sensing Approach. *J. Appl. Remote Sens.* **2020**, *14*, 044513. [[CrossRef](#)]
21. Hebbar, R.; Ravishankar, H.M.; Trivedi, S.; Manjula, V.B.; Kumar, N.M.; Mukharib, D.S.; Mote, J.K.; Sudeesh, S.; Raj, U.; Raghuramulu, Y.; et al. National Level Inventory of Coffee Plantations Using High Resolution Satellite Data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-3/W6*, 293–298. [[CrossRef](#)]
22. Bautista Calderon, E.A.; Ordaz Chaparro, V.M.; Gutiérrez Castorena, M.d.C.; Gutiérrez Castorena, E.V.; Cajuste Bontemps, L. Sistemas agroforestales de café en Veracruz, México: Identificación y cuantificación espacial usando SIG, percepción remota y conocimiento local. *Rev. Terra Latinoam.* **2018**, *36*. [[CrossRef](#)]
23. Kelley, L.C.; Pitcher, L.; Bacon, C. Using Google Earth Engine to Map Complex Shade-Grown Coffee Landscapes in Northern Nicaragua. *Remote Sens.* **2018**, *10*, 952. [[CrossRef](#)]
24. Tridawati, A.; Wikantika, K.; Susantoro, T.M.; Harto, A.B.; Darmawan, S.; Yayusman, L.F.; Ghazali, M.F. Mapping the Distribution of Coffee Plantations from Multi-Resolution, Multi-Temporal, and Multi-Sensor Data Using a Random Forest Algorithm. *Remote Sens.* **2020**, *12*, 3933. [[CrossRef](#)]
25. Ortega-Huerta, M.A.; Komar, O.; Price, K.P.; Ventura, H.J. Mapping coffee plantations with Landsat imagery: An example from El Salvador. *Int. J. Remote Sens.* **2012**, *33*, 220–242. [[CrossRef](#)]
26. Chemura, A.; Mutanga, O.; Odindi, J. Modelling Leaf Chlorophyll Content in Coffee (*Coffea Arabica*) Plantations Using Sentinel 2 Msi Data. In Proceedings of the IGARSS IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 8228–8231. [[CrossRef](#)]
27. Figueroa-Hernández, E.; Pérez-Soto, F.; Godínez-Montoya, L. La producción y el consumo del café. *ECORFAN* **2015**, *1*, 64–82.

28. INEGI. Continuo de Elevaciones de México (CEM 3.0). 2019. Available online: <https://www.inegi.org.mx/app/geo2/elevacionemex/> (accessed on 7 January 2019).
29. SMN. *Resúmenes Mensuales de Temperaturas y Lluvia*; SMN: Mexico City, Mexico, 2019.
30. González-Espinosa, M.; Ramírez-Marcial, N.; Gómez-Pineda, E.; Parra-Vázquez, M.R.; Díaz-Hernández, B.M.; Musálem-Castillejos, K. Vulnerabilidad Ambiental y Social. Perspectivas de Restauración de Bosques en las Partes Altas de la Sierra Madre de Chiapas. *Investig. Ambient. Cienc. Y Política Pública* **2015**, *6*, 90–91.
31. Moguel, P.; Toledo, V. Biodiversity Conservation in Traditional Coffee Systems of Mexico. *Conserv. Biol.* **1999**, *13*, 11–21. [[CrossRef](#)]
32. Hernández-Stefanoni, J.L.; Castillo-Santiago, M.A.; Mas, J.F.; Wheeler, C.E.; Andres-Mauricio, J.; Tun-Dzul, F.; George-Chacón, S.P.; Reyes-Palomeque, G.; Castellanos-Basto, B.; Vaca, R.; et al. Improving aboveground biomass maps of tropical dry forests by integrating LiDAR, ALOS PALSAR, climate and field data. *Carbon Balance Manag.* **2020**, *15*, 15. [[CrossRef](#)] [[PubMed](#)]
33. Brodu, N. Super-Resolving Multiresolution Images With Band-Independent Geometry of Multispectral Pixels. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4610–4617. [[CrossRef](#)]
34. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [[CrossRef](#)]
35. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated Narrow-band Vegetation Indices for Prediction of Crop Chlorophyll Content for Application to Precision Agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [[CrossRef](#)]
36. Filipponi, F. Sentinel-1 GRD Preprocessing Workflow. *Proceedings* **2019**, *18*, 11. [[CrossRef](#)]
37. Hunt, E.R.; Daughtry, C.; Eitel, J.U.H.; Long, D.S. Remote Sensing Leaf Chlorophyll Content Using a Visible Band Index. *Agron. J.* **2011**, *103*, 1090–1099. [[CrossRef](#)]
38. Main, R.; Cho, M.A.; Mathieu, R.; O’Kennedy, M.M.; Ramoelo, A.; Koch, S. An Investigation Into Robust Spectral Indices for Leaf Chlorophyll Estimation. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 751–761. [[CrossRef](#)]
39. Escadafal, R.; Belghith, A.; Ben-Moussa, H. Indices Spectraux Pour la Dégradation des Milieux Naturels en Tunisie Aride. In Proceedings of the 6eme Symposium International Sur les Mesures Physiques et Signatures en Télédétection, Val d’Isère, France, 17–24 January 1994; pp. 253–259.
40. Bannari, A.; Morin, D.; Bonn, F.; Huete, A.R. A Review of Vegetation Indices. *Remote Sens. Rev.* **1995**, *13*, 95–120. [[CrossRef](#)]
41. Darst, B.F.; Malecki, K.C.; Engelman, C.D. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. *BMC Genet.* **2018**, *19*, 65. [[CrossRef](#)]
42. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
43. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
44. Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131. [[CrossRef](#)]
45. Esri. World Imagery. Available online: <https://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08febac2a9> (accessed on 15 January 2019).
46. Cordero-Sancho, S.; Sader, S.A. Spectral analysis and classification accuracy of coffee crops using Landsat and a topographic-environmental model. *Int. J. Remote Sens.* **2007**, *28*, 1577–1593. [[CrossRef](#)]
47. Maskell, G.; Chemura, A.; Nguyen, H.; Gornott, C.; Mondal, P. Integration of Sentinel optical and radar data for mapping smallholder coffee production systems in Vietnam. *Remote Sens. Environ.* **2021**, *266*, 112709. [[CrossRef](#)]
48. Gomez, C.; Mangeas, M.; Petit, M.; Corbane, C.; Hamon, P.; Hamon, S.; De Kochko, A.; Le Pierres, D.; Poncet, V.; Despinoy, M. Use of high-resolution satellite imagery in an integrated model to predict the distribution of shade coffee tree hybrid zones. *Remote Sens. Environ.* **2010**, *114*, 2731–2744. [[CrossRef](#)]
49. Moreira, M.A.; Adami, M.; Rudorff, B.F.T. Spectral and temporal behavior analysis of coffee crop in Landsat images. *Pesqui. Agropecuária Bras.* **2004**, *39*, 223–231. [[CrossRef](#)]
50. Bernardes, T.; Moreira, M.A.; Adami, M.; Giarolla, A.; Rudorff, B.F.T. Monitoring Biennial Bearing Effect on Coffee Yield Using MODIS Remote Sensing Imagery. *Remote Sens.* **2012**, *4*, 2492–2509. [[CrossRef](#)]
51. Marín-Garza, T.; Gómez-Merino, F.C.; Aguilar-Rivera, N.; Murguía-González, J.; Trejo-Téllez, L.I.; Pastelín-Solano, M.C.; Castañeda-Castro, O. Composición Bioactiva De Hojas De Café Durante Un Ciclo Anual. *Rev. Fitotec. Mex.* **2018**, *41*, 365–372. [[CrossRef](#)]
52. Castañeda-Castro, O. Variaciones en Área Foliar y Concentraciones de Clorofilas y Nutrientos Esenciales en Hojas de Café Robusta (*Coffea Canephora* P.) Durante un Ciclo Anual. *Agro Product.* **2018**, *11*. [[CrossRef](#)]
53. Júnior, A.F.C.; de Carvalho Júnior, O.A.; de Souza Martins, E.; Guerra, A.F. Phenological characterization of coffee crop (*Coffea arabica* L.) from MODIS time series. *Braz. J. Geophys.* **2013**, *31*, 569–578. [[CrossRef](#)]
54. Ruiz-de Oña, C.; Merlín-Urbe, Y. New Varieties of Coffee: Compromising the Qualities of Adaptive Agroforestry? A Case Study From Southern Mexico. *Front. Sustain. Food Syst.* **2021**, *5*, 123. [[CrossRef](#)]