

Article

Estimation of Vegetation Cover Using Digital Photography in a Regional Survey of Central Mexico

Víctor Salas-Aguilar ^{1,*}, Cristóbal Sánchez-Sánchez ², Fabiola Rojas-García ²,
Fernando Paz-Pellat ³, J. René Valdez-Lazalde ² and Carmelo Pinedo-Alvarez ⁴

¹ Programa Mexicano del Carbono, Texcoco 56230, Mexico

² Postgrado en Ciencias Forestales, Colegio de Postgraduados, Texcoco 56230, Mexico;
crisdansanchez@gmail.com (C.S.-S.); fabiosxto1981@gmail.com (F.R.-G.); jrene.valdez@gmail.com (J.R.V.-L.)

³ Postgrado en Hidrociencias, Colegio de Postgraduados, Texcoco 56230, Mexico; ferpazpel@gmail.com

⁴ Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua, Chihuahua 33820, Mexico;
cpinedo@uach.mx

* Correspondence: vsalasaguilarl@gmail.com; Tel.: +52-595-120-2056

Received: 9 September 2017; Accepted: 9 October 2017; Published: 15 October 2017

Abstract: The methods for measuring vegetation cover in Mexican forest surveys are subjective and imprecise. The objectives of this research were to compare the sampling designs used to measure the vegetation cover and estimate the over and understory cover in different land uses, using digital photography. The study was carried out in 754 circular sampling sites in central Mexico. Four spatial sampling designs were evaluated in three spatial distribution patterns of the trees. The sampling designs with photographic captures in diagonal form had lower values of mean absolute error (MAE < 0.12) and less variation in random and grouped patterns. The Carbon and Biomass Sampling Plot (CBSP) design was chosen due to its smaller error in the different spatial tree patterns. The image processing was performed using threshold segmentation techniques and was automated through an application developed in the Python language. The two proposed methods to estimate vegetation cover through digital photographs were robust and replicable in all sampling plots with different land uses and different illumination conditions. The automation of the process avoided human estimation errors and ensured the reproducibility of the results. This method is working for regional surveys and could be used in national surveys due to its functionality.

Keywords: automated classification; sampling design; adaptive threshold; over and understory cover

1. Introduction

Vegetation cover is used in studies of the atmosphere, pedosphere, hydrosphere, and biosphere, as well as their interactions [1]. Remote sensing (RS) technology, particularly the development of vegetation indices, has allowed researchers to estimate vegetation cover at a regional and global scale [2].

Validation of high and medium resolution satellite products is a critical aspect of its usefulness in operational approaches [3]. The feasibility and precision of RS must be verified before data can be applied [4]. One way of validating and re-scaling RS products is the use of field measurements, especially the application of digital photography [5,6].

The use of digital photography to estimate the understory cover (shrub and herbaceous—nadir angle) and overstory cover (arboreal—zenith angle) has been advocated in recent years as one of the most accurate methods to estimate these variables [7,8]. According to Liang et al. [1] and White et al. [9], this technique is the most reliable and can be easily employed to extract vegetation cover information in different physiographic conditions.

In relation to shrub and herbaceous cover (understory), robust segmentation algorithms have been developed to discriminate bare soil and vegetation regardless of the type of vegetation present [10,11] and type of luminosity (presence of shadows) [12]. On the other hand, vegetation cover estimations have also progressed in terms of the classification methods [13,14], automation [15], and classification when there is a mixture of vegetation-sky pixels [16].

The advantage of using a non-destructive method such as digital photography to estimate the vegetation cover is that it can be related to the biophysical variables of an ecosystem at a lower cost and time [17,18]. However, few operational studies (local, regional, and national surveys) contemplate measuring this variable for lack of knowledge on how to pursue it efficiently [19] or for the inconsistency of associating forest attributes to a sampling site with an estimated vegetation cover that may exceed the site surface [20].

A disadvantage of the mentioned methods is that sampling sites in which the experiments are made are generally small and homogeneous (low slope and similar land use conditions and plant architecture) [21]. Therefore, the application of these techniques in operational inventories must be planned to efficiently provide the information for which they are developed, and this condition includes considerations of the accuracy of estimates, costs of data collection and processing, and the speed of the process from the planning stage to the presentation of results [22].

At present, there is no agreement among researchers on how to define the optimum sampling design to derive the leaf area index and vegetation cover in field measurements [23]. In a sampling plot, the vegetation cover and its spatial distribution may vary when considering the effects of management [24].

A rapid, reliable, and economical way to compare vegetation cover sampling designs is by predicting the crown diameter through allometric ratios [25] and by estimating the spatial patterns of trees in the sample site. The advantage is that the crown diameter and spatial clustering of trees can be projected into a geographic information system [26], avoiding the intensive work of conducting and comparing them directly in a survey campaign [27].

Due to the lack of an accurate vegetation cover estimation method for forest surveys in México, the objectives of the present study were: (1) to compare the sampling patterns used to measure the vegetation cover; and (2) to estimate the overstory (trees) and understory (shrub and herbaceous) cover in sampling sites of the State of Mexico, Mexico, using a practical procedure and easily reproducible method.

2. Materials and Methods

The State of Mexico is located in the southern part of the southern plateau of Mexico, between parallels 18°22' and 20°17' North and meridians 98°36' and 100°37' West, in an area of 22,333 km². In this region and particularly around the Valley of Mexico, there are specific environmental and historical conditions that have resulted in great biological and cultural diversification along mountain ranges, basins, rivers, and forests.

Ceballos et al. [28] consider that the vegetation of the State of Mexico is represented by three main ecosystems with variations: temperate-cold (temperate forests), semi-warm and sub-humid warm (low deciduous forests), and arid zones (arid and semi-arid vegetation).

The study was carried out from January to September 2015; 754 circular sampling sites of 1000 m² were established and distributed in eight forest regions of the State of Mexico (Figure 1) [29].

In each region, we collected information on the type of vegetation cover and land use. The classification of vegetation was established according to the land use and vegetation chart, Series IV, scale 1:250,000 [30], and was verified in the field.

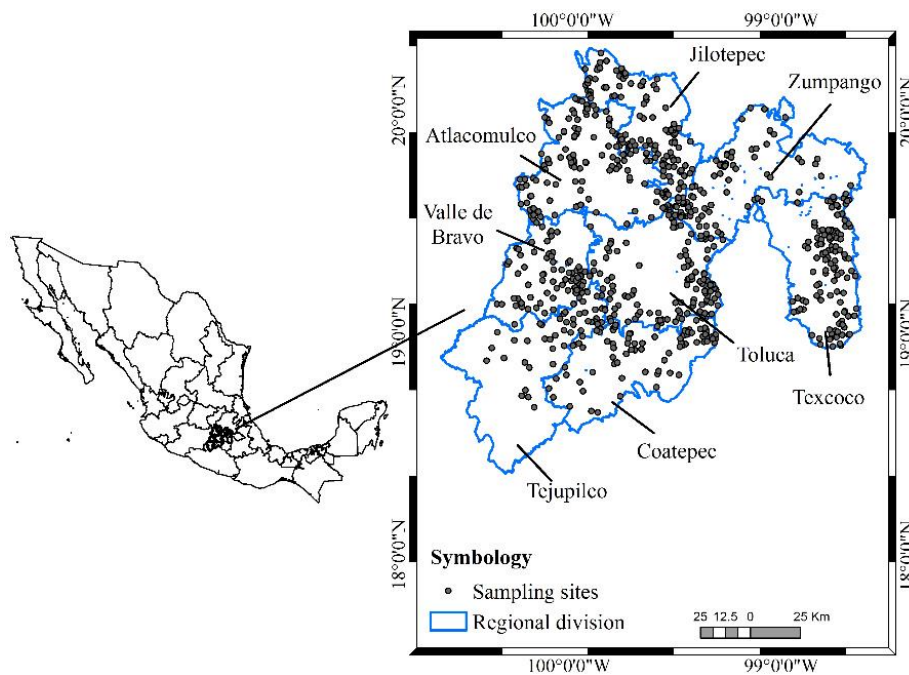


Figure 1. Spatial location of the sampling sites, within the forest regions of the State of Mexico.

2.1. Spatial Projection to Evaluate Sampling Designs

The regional survey was planned as a complement to the National Forest and Soil Inventory (NFSI) in a simplified way, where the height and crown diameters were not measured as done in the NFSI. These variables were planned to be estimated from the state and national surveys.

Before the survey phase, a pre-survey of 30 sites was carried out in the Texcoco forest region to evaluate the spatial pattern of trees in four sampling designs of vegetation cover. Comparisons were made between VALERI [31] and SLAT [32] designs, along with two alternative samples: CBSP (carbon and biomass sampling plots) and RM (regular mesh). The VALERI design is composed of 13 samples, SLAT of 15 samples, CBSP of 21 samples, and RM of 37 samples (Figure 2a).

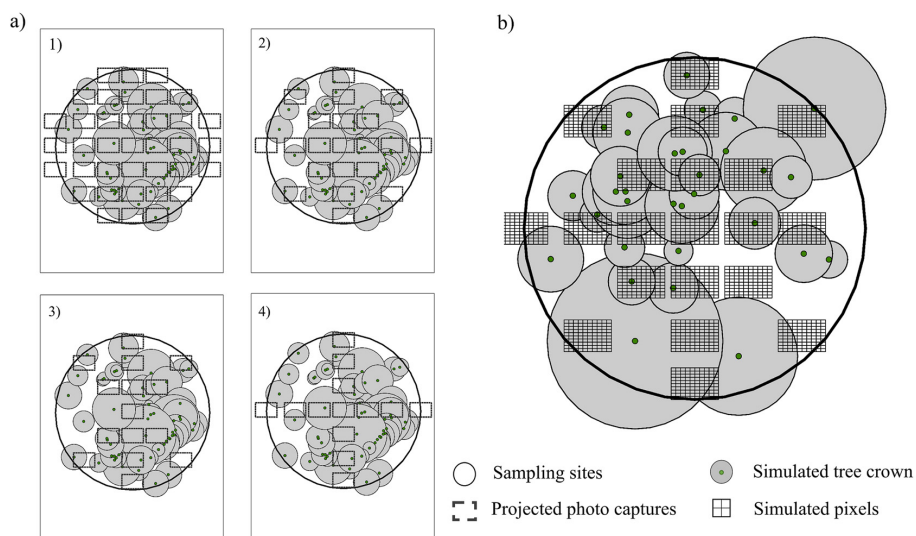


Figure 2. (a) Photographic captures by sampling design: (1) RM, (2) CBSP, (3) SLAT, (4) VALERI; (b) Projected photo capture areas (pixels) in a CBSP sampling design.

Due to the difficulties of knowing the real vegetation cover within a sampling site, the comparison of sampling designs was performed within a geographic information system. Initially, we recorded the central coordinates of each site (as planned in the regional survey), the distance of the trees to the central point, and their azimuth. Then we calculated the location of trees using the central location of the plot and the azimuth and distance to the central point of the respective tree.

Given that no sampling of the tree crowns diameter was recorded in the survey, an allometric relationship was established between the crown diameter and diameter at breast height [25]. The data of this function were obtained from the National Forest and Soil Inventory [33]. The function is the following: $DC = 0.1553 + 0.1859 (Dn)$ ($R^2 = 0.79$, $p < 0.001$), where DC is the tree crown diameter and Dn is the diameter at breast height. The linear model is generalized because it comprised all of the timber species found in the survey. The estimated DC allowed us to construct the crown influence area projection of the trees, assuming a circular shape.

The spatial patterns of the trees were evaluated using the Average Nearest Neighbor (ANN) equation [34]. If the pattern of the tree distribution is completely random $ANN = 1$, if $ANN < 1$ the trees are grouped, and if $ANN > 1$ the tree mass is regular (dispersed). The ANN analysis was performed within the ArcGIS (10.3, Esri, Redlands, CA, United States).

2.2. Projected Photographic Captures within the GIS

A single photographic capture area of 16.38 m^2 was established to determine the projected cover per site and type of sampling (the procedure for estimating the area is described below). In each area we built a grid (10×10) with the purpose of simulating the pixels of a photographic camera (Figure 2b).

The total observed cover was calculated by dividing the area of overlapping crowns between the sizes of the plot (1000 m^2). On the other hand, the estimated cover resulted from the following equation:

$$\sum_{i=0}^n \frac{NPS_i}{NTP} \quad (1)$$

where NPS_i is the number of projected pixels (grid) intersecting with the tree crown area and NTP is the total number of pixels per sampling design.

As a quantitative measure of the error, the mean absolute error (MAE) was estimated:

$$MAE = N^{-1} \sum_{I=1}^N |O_i - E_i| \quad (2)$$

where O is the observed value of the total projected cover, E is the estimated value (Equation (1)), and N is the number of captures per sampling design.

2.3. Field Sampling

Sampling sites were targeted to include vegetation succession and degradation among land uses in Central Mexico [29]. Information was collected on sites with and without anthropogenic intervention.

2.4. Photo Features

The photographic images were taken at a resolution of 5184×3456 pixels in JPG format. We used a Canon EOS Rebel T5^{RM} camera. The camera lens was adjusted to a range of 18 to 55 mm focal length and an ISO 200 with aperture and exposure in automatic mode.

2.5. Taking Photos at the Sampling Sites

We applied the CBSP sampling design with 21 captures to nadir and zenith, according to Figure 3a. The lines represent the transects within the sampling site (L1–L4) and each letter represents

a photographic capture. Figure 3b shows the photograph taken at zenith, where the distance between the camera and the ground is 1.5 m.

Figure 3c shows the process of shooting understory (shrub and herbaceous strata), where the interference of the personnel in the photograph was avoided using a stick of five meters long; in this case, the distance between the camera and the ground was 3 m. Two levels of bubble were used to control the angle in which the photographs were taken, one near the operator and the other one stuck to the side of the camera.

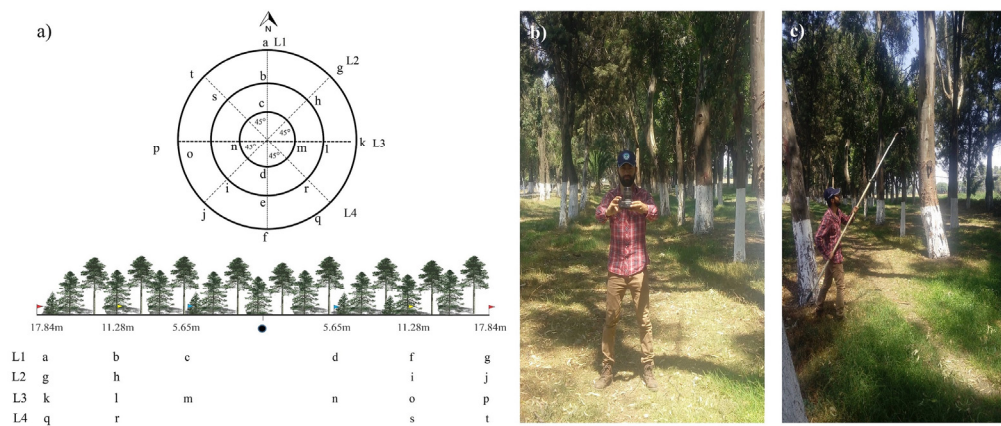


Figure 3. Photographic capture process at sampling sites. (a) CBSP design, the letters correspond to the distance of capture from the center of the sampling site; (b) photographic capture to zenith; (c) photographic capture to nadir.

The purpose of the CBSP sampling was to capture the largest possible physical area with the fewest samples. To do so, the visual field angle of the lens (θ) was adjusted depending on the size of the sensor (n) and its focal length (f):

$$\theta = 2 \times \tan^{-1} \left(\frac{n}{2 \times f} \right) \quad (3)$$

The real area covered by a photograph depends on three variables: sensor size (n_{ij}), focal length of the lens (f), and distance of the lens to the object (h). In the case of nadir, h is the distance between the camera lens and the ground. For zenith, h corresponds to the distance between the lens and the tree crowns. Equation (4) defines the calculation of the real area of the photograph:

$$G_{ij} = \frac{n_{ij} \times h}{f} \quad (4)$$

where G is the actual length of the object in the horizontal (i) and vertical (j), where the horizontal distance of the n_i sensor size of the camera used was 22.3 mm and the vertical distance n_j was 14.9 mm.

The value of f for the nadir photographs was set at 18 mm because h was established at 3 m. By solving Equation (4), we estimated a real area to nadir of 9.2 m².

In the case of zenith photographs, a larger real area was required to be representative at the sampling site. The minimum average height of the tree crowns in the forested areas of the region was 4 m. At this point, the value of θ must be adjusted to reach the largest surface, so the value of f was set at 18 mm. Then, by solving Equation (4), the real area at zenith is 16.38 m².

In heterogeneous forests, such as the study area, the height of the trees can vary in short distances, so in order to maintain the area captured independently of the height of the tree, the value of f can be adjusted by multiplying the average height of the trees at the point of capture by the constant 4.5 ($f = 4.5 \times h$). If the height value was four meters, the camera was placed as close as possible to the

ground; in the case of exceeding six meters in height, the camera was placed at a fixed distance of 1.5 m on the ground.

2.6. Estimation of over and Understory Cover

The processing of images to estimate the vegetation cover is different in nadir and zenith projections; in the first case a robust classifier is needed to distinguish the shade of the vegetation, whereas the second one requires a methodology that distinguishes the cover of the canopy in contrast to the sky (atmosphere). Due to the large number of photographs that needed to be processed (24,182 photographs), a code was written in the Python 2.7^{RM} language (Python Software Foundation (PSF): Wilmington, DE, USA) to optimize the process (the software can be requested from the authors). The following sections describe the methodology used.

2.6.1. Estimation of Overstory Cover

The photographs were taken in the morning and in the afternoon, before the sun surpassed 130° of azimuth or after 230° to avoid confusion due to the brightness of the leaves in association with the sky. We used the SunEarthTool tool (http://www.sunearthtools.com/dp/tools/pos_sun.php) to identify the appropriate times to take the photographs.

The methodology of Fuentes et al. [15] was adjusted within the Python language for image processing. The images were converted to vector format in order to separate the three color channels (R, G, B). The blue channel (B) was used to filter the clouds from the image because it gives the best contrast between the cover of the foliage and the sky with the presence of clouds. The adaptive threshold method was used to classify the image [35].

The method consists of dividing the image into sub-images. The threshold (M) of the sub-image is calculated using the mean or median Gaussian methods. In this case, the median was used as a threshold to perform the separation. The size of the blocks used to divide the image was 200×200 pixels. The sum of the proportions of the number of pixels with vegetation in each block to the total number of pixels of the photograph was the cover of the canopy per photograph. Figure 4 shows an outline of the threshold calculation using this method. The overstory cover includes branches and the upper stems of trees.

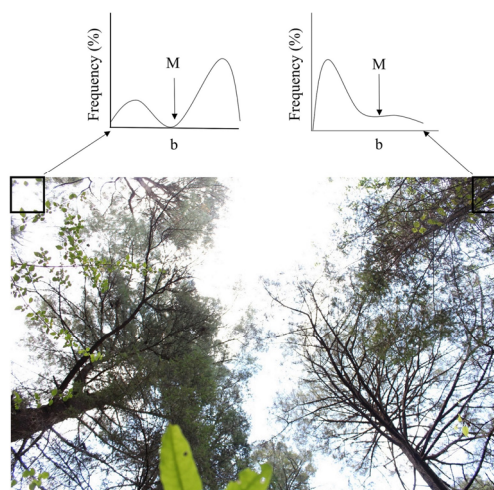


Figure 4. Outline of thresholds in blocks by the adaptive threshold method; the left part of the histogram corresponds to pixels with vegetation, the right side of the histogram corresponds to pixels without vegetation. The M value is the threshold for each block.

2.6.2. Estimation of Understory Cover

The classification of green vegetation and soil was achieved by calculating a threshold within a two-dimensional space. The images were transformed in the color space $L^*a^*b^*$ [10]. The green-red component a^* was used to distinguish the vegetation from the bare soil, where the values skewed to the left of the histogram indicated green pixels (vegetation) and those skewed to the right showed pixels in red (bare soil). The assumption of the methodology is that the distribution of this component tends to be a bimodal Gaussian distribution.

$$F(x) = \frac{W_1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + \frac{W_2}{\sqrt{2\pi\sigma_2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} \tag{5}$$

where μ_1 and μ_2 are the green vegetation and soil average, respectively; and σ_1 and σ_2 are the standard deviations of green vegetation and soil, respectively. The value W_1 is a weighting of the pixels in green and W_2 is the respective weighting for soil. The image is scaled to values of 0–255.

Threshold Adjustment

Regardless of the land use, the value of the pixels is between 75 and 150 in all photographs; to make an optimal adjustment, an initial threshold was set, which was obtained in the middle of the range 75–150 ($T_0 = 112$). The optimal value of the threshold (T_1) occurred when the functions of Equation (5) were equal (Figure 5). In this case, the error of omission of vegetation and soil classification, represented by areas S_1 and S_2 , is minimal. The following equation was used to solve T_1 :

$$AT_1^2 + BT_1 + C = 0 \tag{6}$$

where:

$$\begin{aligned} A &= \sigma_1^2 - \sigma_2^2 \\ B &= 2 \times (\mu_1\sigma_2^2 - \mu_2\sigma_1^2) \\ C &= \sigma_1^2\mu_2 - \mu_1\sigma_2^2 + 2\sigma_1^2\sigma_2^2 \ln(\sigma_2W_1/\sigma_1W_2) \end{aligned} \tag{7}$$

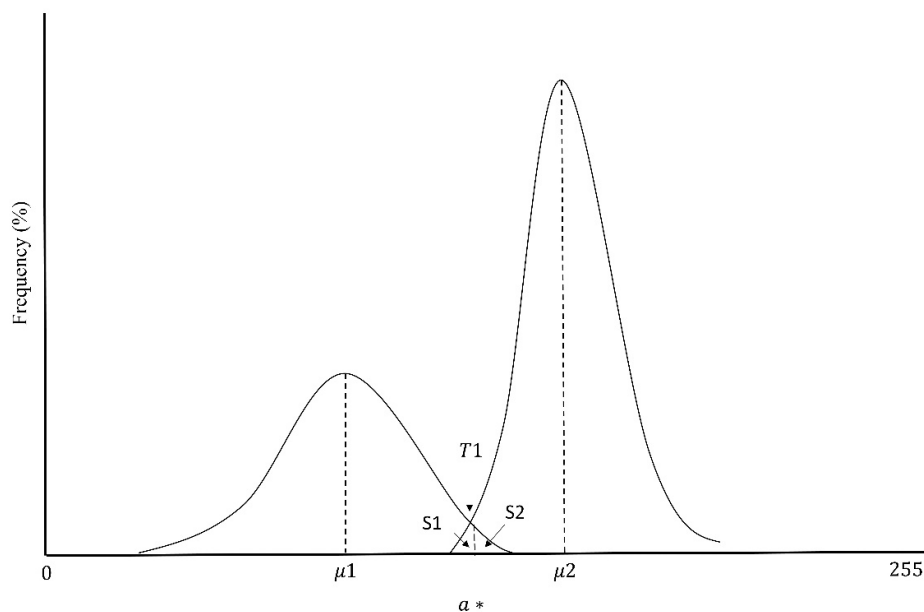


Figure 5. Identification of the optimal T_1 threshold in a bimodal distribution.

In extreme situations where bimodality is not evident, that is, photographs where there is only vegetation or bare soil, we applied the algorithm proposed by Liu et al. [10].

2.6.3. Calculation of Total Vegetation Cover

Total vegetation cover (TVC) was calculated as follows (Figure 6):

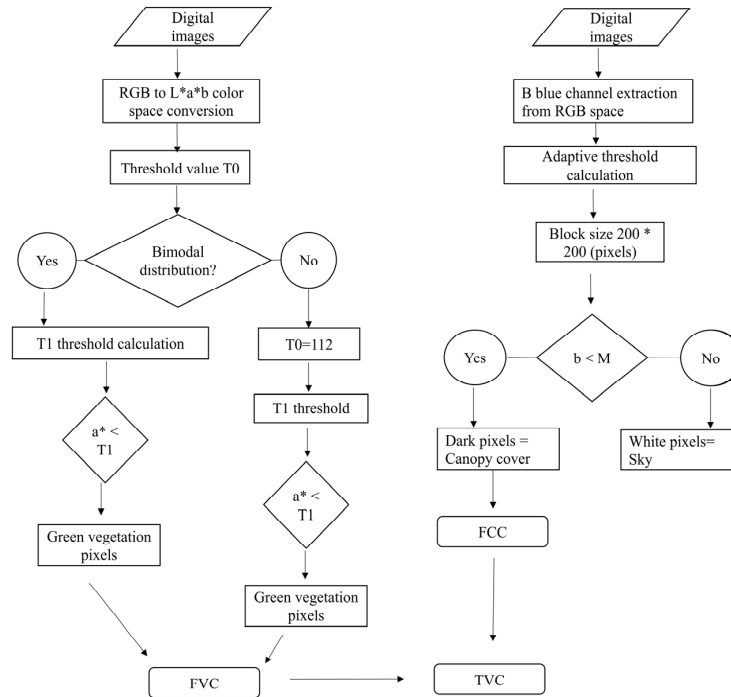


Figure 6. Flowchart for estimating the cover of vegetation vertical strata at sampling sites in the State of Mexico.

$$TVC = \frac{\sum_{i=0}^n FCC}{NPT} + \left(1 - \frac{\sum_{i=0}^n FCC}{NPT}\right) \times \left(\frac{\sum_{i=0}^n FCV}{NPT}\right) \quad (8)$$

where FCC is the proportion of the number of pixels classified as aerial (overstory) cover, FCV is the fraction of the vegetative cover of the understory (lower stratum), and NPT is the total number of pixels contained in the image. The sum of TVC in all images of the CBSP design was considered the total cover per plot. Figure 6 presents the flowchart of the process of classification.

2.6.4. Accuracy in Cover Classification

The accuracy of the cover estimates obtained using the proposed methodology was calculated through a comparison of these values using a visual classification of the images within the ENVI 5.0^{RM} program. We considered two classes to distinguish the colors in the photographs. In understory, all pixels in green were considered as leaves, and the rest were classified as bare soil. In overgrowth, all pixels corresponding to leaves, stems, and branches were classified as cover, and the rest of the pixels were classified as sky. As mentioned in [11], visual classification is considered as the real values of cover in the image and those are compared with the automated threshold proposed in this research.

Images of 12 zenith plots (252 images) and 11 plots to nadir (231 images) were used. The plots represented different land uses. The accuracy of the implemented classifier (AC) was evaluated using the following formula [11].

$$AC = 100 \times \left(1 - \frac{\sum_{i=1}^n \left| \frac{A-B}{A} \right|}{N}\right) \quad (9)$$

where A represents the number of pixels with a real presence of vegetation in the reference image (visual classification) and B represents the number of pixels classified as having vegetation in the applied methods. An average accuracy was obtained in each plot evaluated, where 100% corresponds to a classification without errors.

3. Results

3.1. Sampling Design

The observed (total area of projected crowns) and estimated (projected photographic captured area) values of the projected cover (%) per type of sampling design and spatial pattern of the trees are shown in Table 1. In the pattern analysis of the trees, 10 plots corresponded to a grouped (clustered) pattern, 15 to a random pattern, and five plots belonged to a dispersed cluster. The calculation of the estimated area is explained in Equation (1).

Table 1. Estimated and observed cover values (%) per type of sampling and spatial pattern of trees.

Design	Spatial Pattern	Observed (%)	Estimated (%)	MAE	Coefficient of Variation
RM	Grouped	87.1	80.9	0.175	0.24
	Random	64.0	58.1	0.173	0.34
	Dispersed	34.9	29.0	0.171	0.06
CBSP	Grouped	87.1	85.3	0.091	0.09
	Random	64.0	62.5	0.097	0.33
	Dispersed	34.9	32.9	0.089	0.16
SLAT	Grouped	87.1	86.9	0.079	0.22
	Random	64.0	62.2	0.119	0.38
	Dispersed	34.9	33.7	0.117	0.20
VALERI	Grouped	87.1	34.2	0.153	0.23
	Random	64.0	31.7	0.179	0.41
	Dispersed	34.9	29.2	0.155	0.22

RM: regular mesh; CBSP: Biomass and carbon sampling plots; SLAT: tree and land use sampling; VALERI: Remote sensing ground validation instrument.

The CBSP design showed the least error in two of the three types of spatial patterns. The second design that showed minor error was SLAT, which indicates that the sampling design with photographic captures in diagonal form exhibits better results. The RM and VALERI designs had the highest and lowest number of samples, respectively. However, their errors were the highest (MAE > 0.15). These results practically discard them from being considered in operational sampling.

The random spatial pattern showed a higher coefficient of variation (CV) in the four sampling designs due to the design geometry. The dispersed pattern was the second one with the highest CV in the CBSP, SLAT, and VALERI designs. Within the grouped pattern, the CBSP sample recorded the lowest variation.

3.2. Segmentation of Images

The use of the Python program allowed us to make the segmentation threshold selection consistent. The number of captured photos in the sampling makes it impractical to analyze the photographs in a supervised way or in semi-automated processes (photo by photo). The developed program classifies a sampling plot of 42 photographs at zenith and nadir in 30 s. The processor used has 2.6 GHz and 8 GB of memory.

3.3. Classification of Overstory Images

Figure 7 shows the classification of zenith images in four cover conditions. Figure 7a and its classification 7b represent the photograph with the highest cover in the whole 97% sampling (Oyamel fir forest). Figure 7c, d represent 50% cover (Oyamel fir forest secondary vegetation). Figure 7g

shows how the classifier correctly discriminates foliage from clouds (secondary vegetation of Pine Forest). Finally, Figure 7i presents a correct classification with minimum cover (secondary vegetation of Pine Forest).

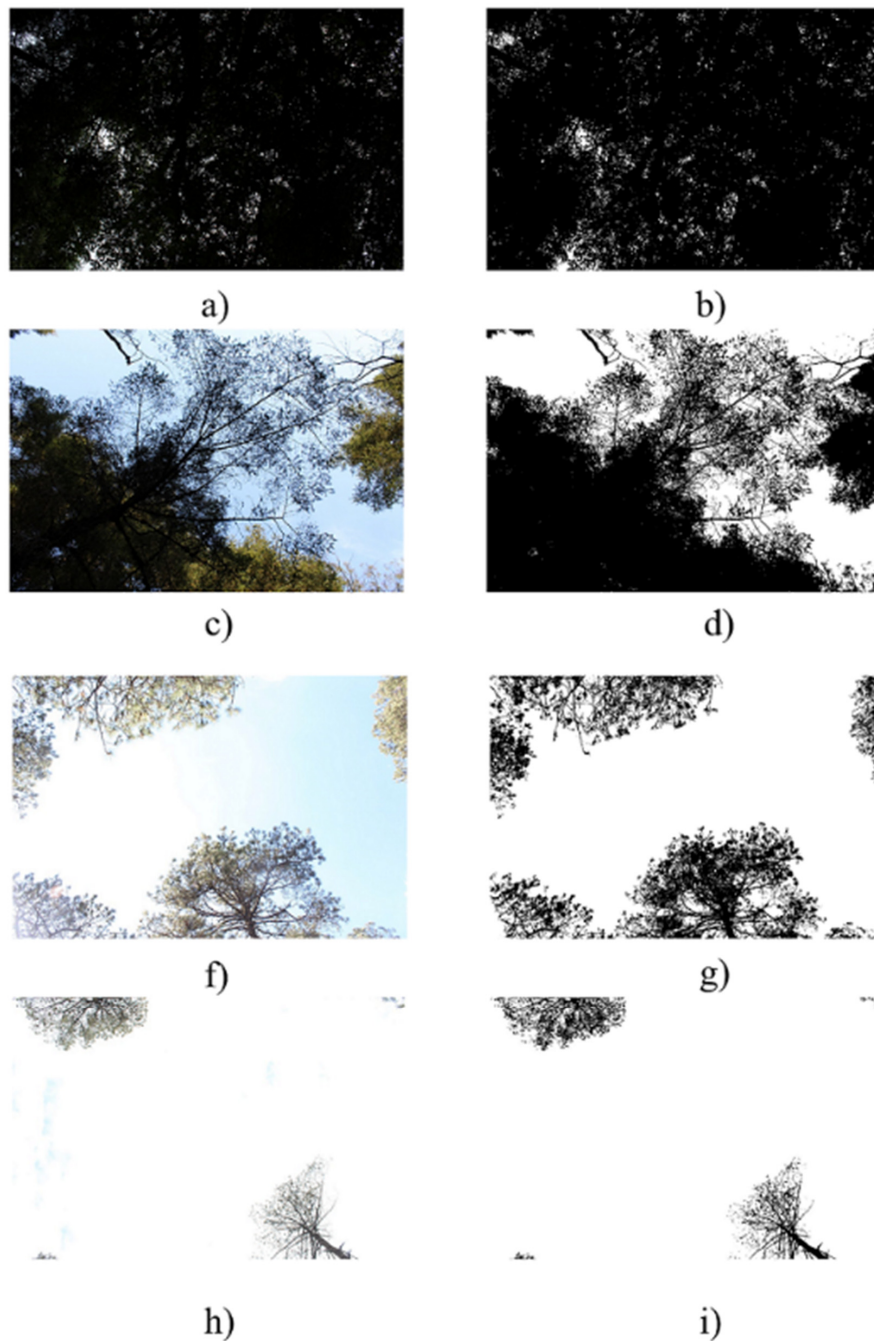


Figure 7. Classification of zenith images, with different cover conditions. Images (a,c,f,h) correspond to the original photograph. Images (b,d,g,i) correspond to the adaptive threshold classifier.

3.4. Classification of Undestory Images

Figure 8 presents a sample of four photographs in different land use and different illumination and vegetation cover conditions, as well as their respective classification and histogram. Figure 8a,b are presented within a plot of land with secondary vegetation of Oyamel fir forest; the threshold in this capture was $a^* = 119$ and the area under the curve with vegetation (V) was 43%. Figure 8c,d

correspond to herbaceous vegetation within an Oyamel fir forest; in this particular vegetation the illumination was reduced because of the high overstory cover, with an understory cover of 62%.

Figure 8f,g show the image within a Rainfed Agriculture area and we observed that the classifier adequately discriminated between bare soil vegetation and shadows; the threshold in this image was $a^* = 122$ and the vegetation cover was 35%. Figure 8i,h are presented within a plot without vegetation; the classifier was able to detect the minimal green cover found in the photograph. In this case, as the distribution of the histogram was unimodal, the threshold was set at $a^* < 105$ and the vegetation cover was 0.8%.

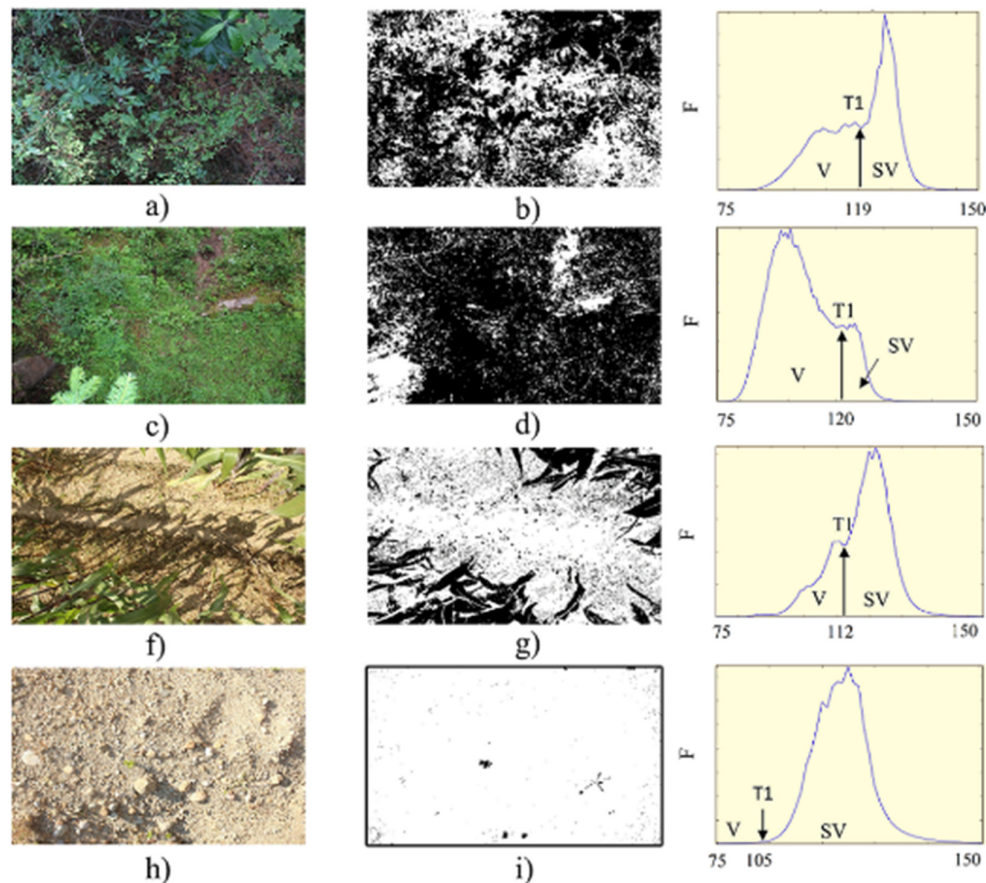


Figure 8. Classification of images in different land uses to nadir. Images (a,c,f,h) correspond to the original photograph. Images (b,d,g,i) correspond to the classified images. The third column shows the distribution of the histogram according to the cover.

3.5. Accuracy of the Classification

Table 2 presents the comparison of the classified cover (%) by the supervised classification method (visual classification) and the zenith method (Estimated). The accuracy was high in all land uses with an average of 93%. In relation to the coefficient of variation CV (representing the variability of the sampling design), we observed that this variability increased as the estimated cover of the different land uses declined. In primary forests (BQP, BQ, BA, BPQ), the variation of the sampling design was low; insofar as there was disturbance in the land use (VSA, VSA, VSh) the variation increased because the static designs were sensitive to the opening of the canopy.

Table 2. Accuracy evaluation of the visual interpretation of the images in 12 zenith sampling plots (overstorey cover).

Land Use	Description	Classified (%)	Estimated (%)	AC	CV
BQP	Oak-pine forest	84.0	83.0	99.0	0.06
BQ	Oak forest	86.0	84.0	96.0	0.09
BA	Oyamel fir forest	82.0	78.0	94.0	0.10
BPQ	Pine-oak forest	74.0	75.0	96.0	0.12
VSa/BQ	Secondary shrub vegetation of oak forest	52.0	48.0	90.0	0.13
VSA/BPQ	Secondary arboreal vegetation of pine-oak forest	69.0	67.0	94.0	0.17
VSa/BQP	Secondary shrub vegetation of oak-pine forest	71.0	69.0	96.0	0.21
VSa/BQP	Secondary shrub vegetation of oak-pine forest	50.0	49.0	96.0	0.22
VSA/BA	Secondary arboreal vegetation of oyamel fir forest	68.0	62.0	91.0	0.24
VSA/BP	Secondary arboreal vegetation of pine forest	22.0	22.0	91.0	0.28
VSh/BP	Secondary herbaceous vegetation of pine forest	16.0	15.0	92.0	0.66
VSh/BPQ	Secondary herbaceous vegetation of pine-oak forest	31.0	35.0	87.0	0.68

Table 3 presents the evaluation of the vegetative cover to nadir. The average accuracy was 94%. As in zenith, the estimated cover maintains a negative correlation with CV. For this classification, the greatest cover is found in secondary herbaceous vegetation (VSh) where the variability of the sampling is lower; as the vegetation transits to mature forest, the CV is high due to the fact that the understory cover is random. In sites with non-vascular vegetation (bryophytes), the classifier overestimates the percentage of vascular vegetation because it associates the green color with this type of cover.

Table 3. Accuracy evaluation of the visual interpretation of the images in 12 sampling plots at nadir (understorey cover).

Land Use	Description	Classified (%)	Estimated (%)	AC	CV
VSh/BPQ	Secondary herbaceous vegetation of pine-oak forest	82.10	85.00	89.0	0.06
VSh/BP	Secondary herbaceous vegetation of pine forest	78.43	79.00	96.0	0.10
TA	Rain-fed agriculture	27.37	28.00	90.0	0.16
VSh/BA	Secondary herbaceous vegetation of oyamel fir forest	84.00	83.50	98.0	0.19
VSa/BQ	Secondary shrub vegetation of oak forest	32.70	32.00	94.0	0.51
PC	Cultivated grassland	47.23	49.33	97.0	0.84
BP	Pine forest	29.30	27.67	95.0	0.86
VSA/BPQ	Secondary arboreal vegetation of pine-oak forest	37.65	34.42	92.0	0.87
BQP	Pine-oak forest	30.69	29.87	96.0	0.88
BQ	Oak forest	6.58	6.00	94.0	0.94
VSa/BQP	Secondary shrub vegetation of oak-pine forest	5.63	6.95	87.0	0.98

4. Application of CBSP Design to the Regional Survey

After validating that the CBSP design was robust and precise enough to be used in a regional survey, it was implemented in the campaign through a replication of the procedure used in the pre-survey evaluation in the 754 sampling plots.

It is identified that the application of the sampling design allowed us to capture photographs in an easy and fast way in the majority of land uses.

Figure 9 shows the average total vegetation cover of the main land uses in the regions of the State of Mexico. The cover data are presented from highest to lowest and from mature forest to agriculture. Mature forests have a tendency of 50–90% cover in the eight regions, and there was higher average coverage in the region of Toluca (70–90%) (Figure 9a); in wooded areas with secondary tree vegetation (VSA) and secondary shrub and herbaceous vegetation (VSa and Vsh), the cover ranges from 20 to 90%. This is because the limit of tree vegetation in these plots is a mixture of perennial and deciduous cover, therefore presenting a seasonal change of vegetation cover because of the weather pattern. In this case, the region with the highest coverage was Zumpango (Figure 9b) and that with the lowest coverage was Coatepec (Figure 9f).

With regard to cover in agricultural areas (RA, TA), the development of cover over time follows a spatial pattern associated with the time of planting and growth. Cover starts at <10% in all regions and the percentage increases up to 100%, as in the Toluca and Texcoco regions (Figure 9a,c).

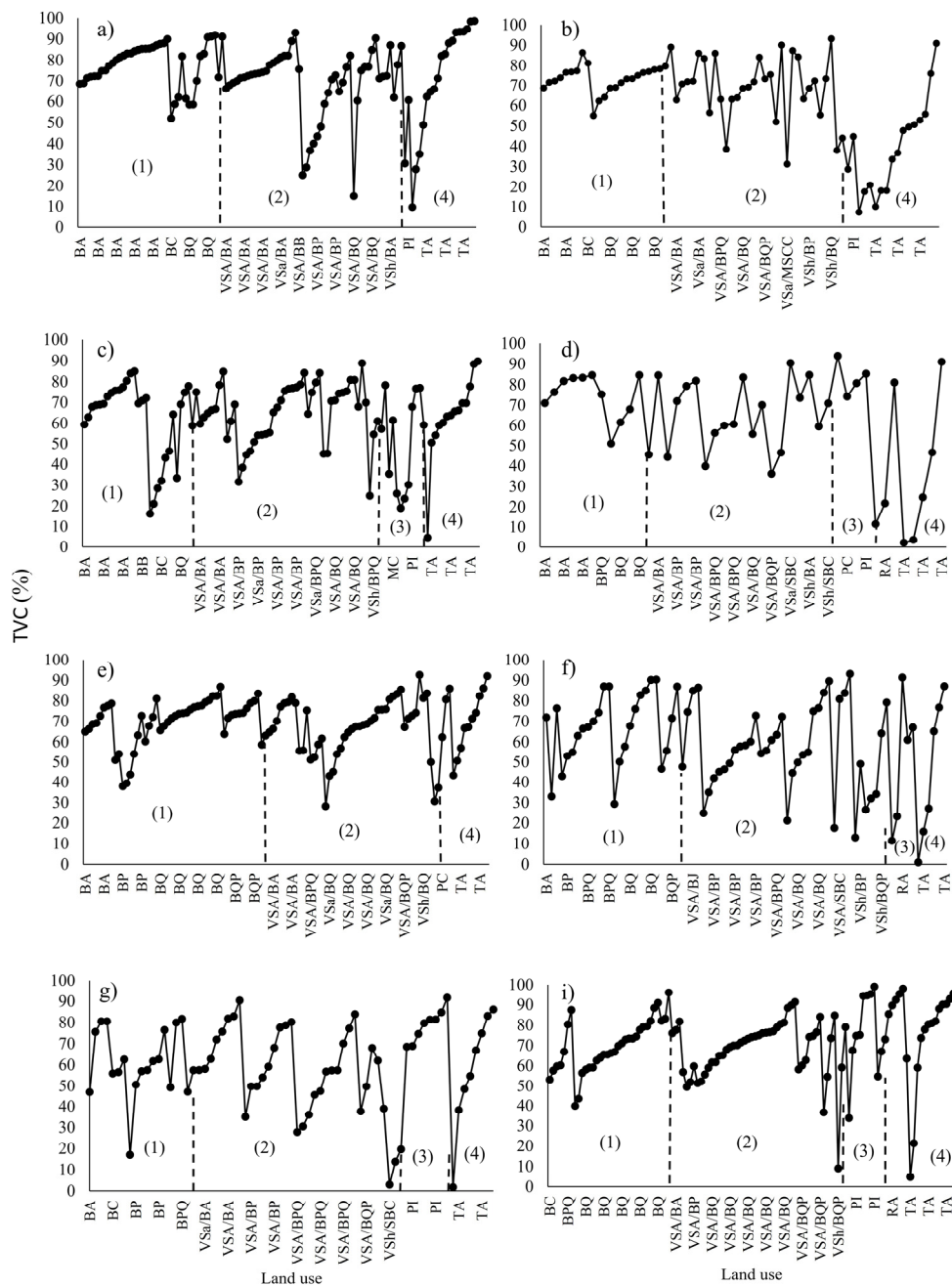


Figure 9. Average total vegetation cover by land use in the eight regions of the State of Mexico: (a) Toluca Region; (b) Zumpango Region; (c) Texcoco Region; (d) Tejupilco Region; (e) Atlacomulco Region; (f) Coatepec Region; (g) Valle de Bravo Region; (i) Jilotepec Region. (1) Mature forest; (2) disturbed forest; (3) grassland; (4) agricultural area.

5. Discussion

The use of digital photography for vegetation cover estimations is an easy, low-cost, and potentially suitable approach for hard-to-reach places. These properties give this method an advantage over direct (destructive) and indirect (fisheye cameras) methods [13]. In Mexico, operational surveys such as the National Forest and Soils Survey [33] generate vegetation cover estimations that rely on the technician criteria in the field. This research proposes an accurate survey design method which is potentially suitable for the forest sector in the country.

The advantage of using field data is its strength for the validation of satellite information, as a way to use the cover values at greater scales [6]. On the other hand, a disadvantage of field photography data is the requirement of several photographs in order to produce a reliable estimate. However, there are software methods for the automation of these processes, such as the one proposed in this research.

5.1. Sampling Designs Comparison

One important consideration in field vegetation cover measurements is the need to determine an appropriate sampling design [31]. The methodologies for measuring vegetation cover are ambiguous in reference to which method should be used to ensure a correct estimation within a sampling plot. The projection made in the GIS allowed us to observe differences in vegetation cover estimations with a simple scheme. The results provide important information for choosing a sampling design and reducing the costs of the collection and processing of field data [22], considering the large amount of samples of this survey.

Martens et al. [36] show that the spatial patterns influence the height, cover, and distribution of vegetation in their different strata. Our results showed that, independently of the spatial pattern of the survey sites, sampling designs that captured diagonal photographs (CBSP and SLAT) exhibited the least estimation error.

The advantages of these designs are the low number of photographs needed (42 photographs) and their easy field implementation. This ratifies the vegetation cover estimation operability using this method and how it can be related to grouping indices to evaluate the effect of the forest management of zones without disturbance on disturbed zones [24].

5.2. Sampling Survey

Forest surveys in several parts of the world, including Mexico, are carried out in circular plots of 1000 m² (17.85 m radius) [22]. In this research, we adopted this design to evaluate the biomass and carbon storage in different land uses within the State of Mexico. The CBSP design proved to be feasible for its implementation in different land uses and spatial patterns of the trees; this simple design allowed the application in distant sampling points and rugged terrain, which makes it an operative method to capture vegetation cover with digital photographs.

An example of the sampling operability is that bubble levels were used to stabilize the camera at each sampling point, instead of using tripods to fix it. This technique helped to reduce the time to take the photographs and presented no considerable error when compared to tripod shots [37].

The number of samples and their arrangement was another variable to be considered for the sampling to be operative and related to other variables measured in the plot (i.e., biomass, carbon), so captures were fixed at the ends of the sites [20]. The CBSP design obtained the best results when estimating the vegetation cover. The efficiency of its application and smaller errors in the cover estimation turns it into a practical design for this type of application, besides saving storage and time when processing the images.

The spatial distribution of trees within the site is an important element for the dynamics of the forest ecosystem [26]. In the case of overstory cover, we observed that the applied sampling design showed a negative trend between the estimated cover and its coefficient of variation (CV). The primary forests presented high and compact canopy covers (low CV); when reducing it, the canopy cover tends to be dispersed (high CV). In the case of understory cover, the opposite occurs. In disturbed areas (secondary vegetation), the cover is larger and compact; when this cover is reduced, the pattern is dispersed.

The real area of a photograph is another important aspect that has been little explored in vegetation cover sampling design. Researchers generally use a fixed lens viewing angle (35–40°) to estimate the canopy cover, so a greater opening angle would be measuring the closure of the canopy. Jennings et al. [38] describe in detail the difference between these two concepts. In this study, we observed that in real situations the height of the trees in a sampling site is heterogeneous. For this

reason, we proposed to adjust the focal length of the camera at the point of capture by a constant as this ensures that one is able to approximate and make repeatable the capture area independently of the architecture of the trees.

5.3. Automated Classification of Images

The automation of vegetative cover classification is a process that avoids the error of the human component and ensures the possibility of reproducing the results [11]. The efforts to automate image classification have focused on programs such as MATLAB [12,15,39], WinScanopy [40], or Photoshop [41], among others. But although the former meets the requirement of making the batch processing of the images operative, it has the disadvantage of having an additional cost. The other programs have the disadvantage of processing the images photo by photo, which discarded them for the analysis of cover in this study. The Python^{RM} program was chosen for the versatility of the specialized libraries available, and for being a free access program. The written code enabled us to process in batch the images of the sampling in a time similar to that described in programs like MATLAB.

5.3.1. Overstory Cover

In the classification of digital photographs, binary methods (global thresholds) are generally used to estimate canopy cover [42]. The colors of the classified images have gray tonalities for vegetation and white for the sky. According to Chityala and Pudipeddi [35], the accuracy in the classification for the global threshold methods is low. The problem is trying to find the maximum variance between two logical groups of segments within the whole image, which can cause confusion in the classification by overexposure in the camera or image capture at inappropriate times. In this research, we propose the use of the adaptive threshold in the blue space of the image [15]. The method is based on the same principles of the global threshold, but the segmentation statistics are done at the block level within the image, allowing greater accuracy in the overall classification.

The accuracy of the classification is high and related to the comparison of binary algorithms performed by Ghatthorn and Beckschäfer [42]. The sampling was directed to avoid the effect of the sun (the captures were made before noon and at near sunset); however, in the sampling, there were circumstances that caused the photographs to exceed the proposed range, such as the VSh/BPQ land use plot (Table 2), which showed the lowest precision (87%) [16]. Nonetheless, the development of methodologies applied to this problem are beyond the scope of this research.

5.3.2. Understory Cover

There are many methods to extract the vegetation cover fraction [11,12,43] so the degree of accuracy is an important factor in the efficiency of field measurements. For example, supervised classification has a high accuracy but low efficiency, while unsupervised classification has a high efficiency but low precision due to commission and omission errors [1].

The algorithm proposed by Liu et al. [10] has the property of being simple, easy to automate, and has a high degree of precision. When comparing it with the supervised classification in different sampling plots, the results in terms of precision were high (>87%); the main problem of the misclassification was the confusion of the vascular vegetation with bryophytes, which in the color space *a detects a green color that is difficult to discern. In conditions of low illumination due to the effect of the canopy, the algorithm had no problems in correctly classifying vegetation and shade [12], and with a single component predomination (bare soil) in the photograph, the classifier generated good results (Figure 8i).

The two methods proposed to evaluate the over and understory vegetation cover were robust and replicable in all sampling plots. The reason for estimating the foliar projective cover rather than the leaf area index is that the former is a more adequate variable to characterize vegetation [44] since the projective foliar cover captured with digital photographs contains information on individual plants

and their spatial distribution. With this perspective and considering the validation of the proposed sampling and plot area, we will contemplate the validation of biophysical variables calculated with remote sensors in future work of the research group.

6. Conclusions

The over and understory vegetation cover was estimated with digital photographs in sampling plots of the State of Mexico. The high efficiency and precision of the classification methods indicate that they are robust for discerning vegetation in different land uses and illumination conditions.

The use of digital photography reduces the ambiguity of vegetation cover estimations in regional and national surveys. The proposed method is easily reproducible in heterogeneous land and vegetation conditions.

The automation of the process using a free programming language avoided human errors and ensured the reproducibility of the results at a low cost.

The sampling methods using the capture of diagonal-angled photographs were the best way to obtain less biased information when taking digital photographs at a circular sampling site. The simulation showed that the CBSP design has a smaller error when considering the spatial distribution of trees within the sampling site. Its easy field management, the number of photographs per site, and its precision make it an operative design. One additional advantage of the proposed field survey is that the real area covered by the photograph is independent of the height of trees. This guarantees representability and avoids image superposition in the sampling site.

Mature forests have a high and compact overstory vegetation cover, which tends to be reduced in secondary forests. The greater cover of the understory is found in secondary forests, where it is denser. The cover of the undergrowth declines in mature forests.

The application of this method for regional and national surveys is recommended.

Acknowledgments: We appreciate the financial support given by the Programa Mexicano del Carbono and Protectora de Bosques del Estado de México (PROBOSQUE) which allowed us to conduct this study.

Author Contributions: Víctor Salas-Aguilar and Fernando Paz-Pellat contributed to the initial proposal of the methodology. Víctor Salas-Aguilar designed and developed the software to process the data in python for implementation on a larger scale. Fabiola Rojas-García, Cristóbal Sánchez-Sánchez, J. René Valdez-Lazalde, and Carmelo Pinedo-Alvarez conducted the research methods. All authors discussed the structure and commented on the manuscript at all stages.

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

References

1. Liang, S.; Li, X.; Wang, J. *Advanced Remote Sensing: Terrestrial Information Extraction and Applications*; Academic Press: Cambridge, MA, USA, 2012.
2. Gutman, G.; Ignatov, A. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *Int. J. Remote Sens.* **1998**, *19*, 1533–1543. [[CrossRef](#)]
3. Li, Y.; Wang, H.; Li, X.B. Fractional vegetation cover estimation based on an improved selective endmember spectral mixture model. *PLoS ONE* **2015**, *10*, e0124608. [[CrossRef](#)] [[PubMed](#)]
4. Mu, X.; Hu, M.; Song, W.; Ruan, G.; Ge, Y.; Wang, J.; Huang, S.; Yan, G. Evaluation of sampling methods for validation of remotely sensed fractional vegetation cover. *Remote Sens.* **2015**, *7*, 16164–16182. [[CrossRef](#)]
5. Li, F.; Chen, W.; Zeng, Y.; Zhao, Q.; Wu, B. Improving estimates of grassland fractional vegetation cover based on a pixel dichotomy model: A case study in Inner Mongolia, China. *Remote Sens.* **2014**, *6*, 4705–4722. [[CrossRef](#)]
6. Mu, X.; Huang, S.; Ren, H.; Yan, G.; Song, W.; Ruan, G. Validating GEOV1 fractional vegetation cover derived from coarse-resolution remote sensing images over croplands. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2015**, *8*, 439–446. [[CrossRef](#)]
7. Zhou, Q.; Robson, M.; Pilesjo, P. On the ground estimation of vegetation cover in Australian rangelands. *Int. J. Remote Sens.* **1998**, *19*, 1815–1820. [[CrossRef](#)]

8. 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](#)]
9. White, M.A.; Asner, G.P.; Nemani, R.R.; Privette, J.L.; Running, S.W. Measuring fractional cover and leaf area index in arid ecosystems: Digital camera, radiation transmittance, and laser altimetry methods. *Remote Sens. Environ.* **2000**, *74*, 45–57. [[CrossRef](#)]
10. Liu, Y.; Mu, X.; Wang, H.; Yan, G. A novel method for extracting green fractional vegetation cover from digital images. *J. Veg. Sci.* **2012**, *23*, 406–418. [[CrossRef](#)]
11. Coy, A.; Rankine, D.; Taylor, M.; Nielsen, D.C.; Cohen, J. Increasing the accuracy and automation of fractional vegetation cover estimation from digital photographs. *Remote Sens.* **2016**, *8*, 474. [[CrossRef](#)]
12. Song, W.; Mu, X.; Yan, G.; Huang, S. Extracting the green fractional vegetation cover from digital images using a shadow-resistant algorithm (SHAR-LABFVC). *Remote Sens.* **2015**, *7*, 10425–10443. [[CrossRef](#)]
13. Macfarlane, C.; Hoffman, M.; Eamus, D.; Kerp, N.; Higginson, S.; McMurtrie, R.; Adams, M. Estimation of leaf area index in eucalypt forest using digital photography. *Agric. For. Meteorol.* **2007**, *143*, 176–188. [[CrossRef](#)]
14. Chianucci, F.; Chiavetta, U.; Cutini, A. The estimation of canopy attributes from digital cover photography by two different image analysis methods. *iFor. Biogeosci. For.* **2014**, *7*, 255–259. [[CrossRef](#)]
15. Fuentes, S.; Palmer, A.R.; Taylor, D.; Zeppel, M.; Whitley, R.; Eamus, D. An automated procedure for estimating the leaf area index (LAI) of woodland ecosystems using digital imagery, MATLAB programming and its application to an examination of the relationship between remotely sensed and field measurements of LAI. *Funct. Plant Biol.* **2008**, *35*, 1070–1079. [[CrossRef](#)]
16. Macfarlane, C. Classification method of mixed pixels does not affect canopy metrics from digital images of forest overstorey. *Agric. For. Meteorol.* **2011**, *151*, 833–840. [[CrossRef](#)]
17. Tausch, R.; Tueller, P. Foliage biomass and cover relationships between tree-and shrub-dominated communities in pinyon-juniper woodlands. *Great Basin Nat.* **1990**, *50*, 121–134.
18. Muukkonen, P.; Mäkipää, R. Empirical biomass models of understorey vegetation in boreal forests according to stand and site attributes. *Boreal Environ. Res.* **2006**, *11*, 355–369.
19. Luna, J.A.N.; Hernández, E.H. Relaciones morfométricas de un bosque coetáneo de la región de El Salto, Durango. *Ra Ximhai* **2008**, *4*, 69–82.
20. Williams, M.S.; Patterson, P.L.; Mowrer, H.T. Comparison of ground sampling methods for estimating canopy cover. *For. Sci.* **2003**, *49*, 235–246.
21. Muir, J.; Schmidt, M.; Tindall, D.; Trevithick, R.; Scarth, P.; Stewart, J. *Field Measurement of Fractional Ground Cover: A Technical Handbook Supporting Ground Cover Monitoring for Australia*; Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES): Canberra, Australia, 2011.
22. Matern, B. *Recopilación de Notas Sobre Técnicas de Muestreo Usadas en Inventarios Forestales*; SARH-INIFAP Pub. Especial: Distrito Federal, Mexico, 1993.
23. Gobron, N.; Verstraete, M. Remote sensing and geoinformation processing in the assessment and monitoring land degradation and desertification state of art and operational perspectives. In *Assessment of the Status of the Development of the Standards for the Terrestrial Essential Climate Variables: Fraction of Absorbed Photosynthetically Active Radiation (FAPAR)*; GTOS Secretariat Food and Agriculture Organization of the United Nation (FAO): Rome, Italy, 23 April 2009.
24. Corral-Rivas, J.J.; Wehenkel, C.; Castellanos-Bocaz, H.A.; Vargas-Larreta, B.; Diéguez-Aranda, U. A permutation test of spatial randomness: Application to nearest neighbour indices in forest stands. *J. For. Res.* **2010**, *15*, 218–225. [[CrossRef](#)]
25. Hemery, G.; Savill, P.; Pryor, S. Applications of the crown diameter–stem diameter relationship for different species of broadleaved trees. *For. Ecol. Manag.* **2005**, *215*, 285–294. [[CrossRef](#)]
26. LeMay, V.; Maedel, J.; Coops, N.C. Estimating stand structural details using nearest neighbor analyses to link ground data, forest cover maps, and Landsat imagery. *Remote Sens. Environ.* **2008**, *112*, 2578–2591. [[CrossRef](#)]
27. Shaw, J.D. *Models for Estimation and Simulation of Crown and Canopy Cover*; General Technical Report (GTR); US Forest Service: Washington, DC, USA, 2005.
28. Ceballos, G.; List, R.; Garduño, G.; López, R.; Muñozcano, M.; Collado, E.; San Román, J. *La Diversidad Biológica del Estado de México*; Estudio de Estado; Biblioteca Mexiquense del Bicentenario: Ventura, CA, USA, 2009.

29. Programa Mexicano del Carbono (PMC). Manual de Procedimientos Inventario de Carbono+. In *Estudio de Factibilidad Técnica Para el Pago de Bonos de Carbono en el Estado de México*; Programa Mexicano del Carbono: Texcoco, Mexico, 2015; p. 69.
30. INEGI Datos Vectoriales Escala 1:250,000 de Uso de Suelo y Vegetación. Available online: <http://www.inegi.org.mx/go/contends/recant/mussel/> (accessed on 24 May 2017).
31. Baret, F.; Weiss, M.; Allard, D.; Garrigues, S.; Leroy, M.; Jeanjean, H.; Fernandes, R.; Myneni, R.; Privette, J.; Morisette, J. VALERI: A network of sites and a methodology for the validation of medium spatial resolution land satellite products. *Remote Sens. Environ.* **2005**, *76*, 36–39.
32. Kuhnell, C.A.; Goulevitch, B.M.; Danaher, T.J.; Harris, D.P. Mapping woody vegetation cover over the state of Queensland using Landsat TM imagery. In Proceedings of the 9th Australasian Remote Sensing and Photogrammetry Conference, Sydney, Australia, 24 July 1998; pp. 20–24.
33. CONAFOR (Comisión Nacional Forestal). *Inventario Nacional Forestal y de Suelos Informe de Resultados 2004–2009 National Forest and Soils Survey, Results Report 2004–2009*; CONAFOR: Zapopan, Mexico, 2012; p. 212.
34. Clark, P.J.; Evans, F.C. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology* **1954**, *35*, 445–453. [[CrossRef](#)]
35. Chityala, R.; Pudipeddi, S. *Image Processing and Acquisition Using Python*; CRC Press: Boca Raton, FL, USA, 2014.
36. Martens, S.N.; Breshears, D.D.; Meyer, C.W. Spatial distributions of understory light along the grassland/forest continuum: Effects of cover, height, and spatial pattern of tree canopies. *Ecol. Model.* **2000**, *126*, 79–93. [[CrossRef](#)]
37. Origo, N.; Calders, K.; Nightingale, J.; Disney, M. Influence of levelling technique on the retrieval of canopy structural parameters from digital hemispherical photography. *Agric. For. Meteorol.* **2017**, *237*, 143–149. [[CrossRef](#)]
38. Jennings, S.; Brown, N.; Sheil, D. Assessing forest canopies and understorey illumination: Canopy closure, canopy cover and other measures. *Forestry* **1999**, *72*, 59–74. [[CrossRef](#)]
39. Korhonen, L.; Heikkinen, J. Automated analysis of in situ canopy images for the estimation of forest canopy cover. *For. Sci.* **2009**, *55*, 323–334.
40. Pekin, B.; Macfarlane, C. Measurement of crown cover and leaf area index using digital cover photography and its application to remote sensing. *Remote Sens.* **2009**, *1*, 1298–1320. [[CrossRef](#)]
41. Lee, K.-J.; Lee, B.-W. Estimating canopy cover from color digital camera image of rice field. *J. Crop Sci. Biotechnol.* **2011**, *14*, 151–155. [[CrossRef](#)]
42. Glatthorn, J.; Beckschäfer, P. Standardizing the protocol for hemispherical photographs: Accuracy assessment of binarization algorithms. *PLoS ONE* **2014**, *9*, e111924. [[CrossRef](#)] [[PubMed](#)]
43. Liu, J.; Pattey, E. Retrieval of leaf area index from top-of-canopy digital photography over agricultural crops. *Agric. For. Meteorol.* **2010**, *150*, 1485–1490. [[CrossRef](#)]
44. Poblete-Echeverría, C.; Fuentes, S.; Ortega-Farías, S.; Gonzalez-Tallice, J.; Yuri, J.A. Digital cover photography for estimating leaf area index (LAI) in apple trees using a variable light extinction coefficient. *Sensors* **2015**, *15*, 2860–2872. [[CrossRef](#)] [[PubMed](#)]



© 2017 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/>).