



# Spatial Analysis of Temperate Forest Structure: A Geostatistical Approach to Natural Forest Potential

Jesús A. Prieto-Amparán<sup>1</sup>, Eduardo Santellano-Estrada<sup>1</sup>, Federico Villarreal-Guerrero<sup>1</sup>, Martin Martinez-Salvador<sup>1</sup>, Alfredo Pinedo-Alvarez<sup>1</sup>, Griselda Vázquez-Quintero<sup>2</sup>, María C. Valles-Aragón<sup>2</sup> and Carlos Manjarrez-Domínguez<sup>2,\*</sup>

- <sup>1</sup> Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua, Periférico Francisco R. Almada Km. 1, Chihuahua, Chihuahua 31453, Mexico; jesus\_prieto06@hotmail.com (J.A.P.-A.); esantellano@uach.mx (E.S.-E.); fvillarreal@uach.mx (F.V.-G.); msalvador@uach.mx (M.M.-S.); apinedo@uach.mx (A.P.-A.)
- <sup>2</sup> Facultad de Ciencias Agrotecnológicas, Universidad Autónoma de Chihuahua, Ave. Pascual Orozco s/n, Campus 1, Chihuahua, Chihuahua 31350, Mexico; gquintero@uach.mx (G.V.-Q.); valles.cecilia@gmail.com (M.C.V.-A.)
- \* Correspondence: manjarrez.carlos@gmail.com; Tel.: +52-(614)-434-0363; Fax: +52-(614)-434-0345

Received: 18 January 2019; Accepted: 14 February 2019; Published: 16 February 2019



Abstract: Forest ecosystems represent an important means of ecosystem services; they are key as carbon sinks, water collectors, soil stabilizers, suppliers of great biological diversity, among other benefits. In addition, regionalization based on forest conditions provides a valuable approach to understanding and analyzing spatial patterns, which is useful as a tool for the implementation of forest ecosystem protection and conservation programs. In this research, the structure of a temperate forest in the western Sierra Madre region of Mexico was analyzed and characterized. The study unit was the watershed and the analysis used a geospatial approach combined with multivariate techniques such as: Principal Component Analysis, Cluster Analysis (CA), Discriminant Analysis (DA) and Multivariate Analysis of Variance. We evaluated the relationships among spectral satellite data, thematic maps and structural forest variables. A total of 345 watersheds were grouped based on these variables. The grouping of watersheds under low, medium and high production conditions was carried out with CA, defining 3 groups. The validation of the grouping was performed through DA, estimating errors with the restitution method, as well as with the cross-validation method. Significant differences were found among the groups. The grouping of watersheds provides observable evidence of the variability of the forest condition throughout the area. This study allows identifying forest areas with different levels of productivity and can help to detect levels of vulnerability and ecological fragility in natural forests in temperate zones.

**Keywords:** multivariate analysis; watershed; landsat; classification; GIS; forest structure; regionalization; spatial clustering

# 1. Introduction

Forest ecosystems provide essential benefits for humanity, including the protection of biodiversity, climate regulation, carbon sinks, among others [1,2]. The assessment of changes in forest ecosystems and the understanding of their causes are of great concern. Factors that can alter the forest structure such as fire, pests, human activities related to the settlement or opening of agricultural land and the extraction of resources, generate loss of biomass that influences biogeochemical cycles [3]. Climatic characteristics determine the health of the world's forest ecosystems [4]. With a continental area of approximately two million square kilometers, Mexico is one of the richest countries in biodiversity [5].



The temperate forests of Mexico cover approximately 32 million hectares, equivalent to 18% of the territory. In these ecosystems we can find a great diversity of associations between pines and oaks that are not present in another part of the world [6]. These forests are mainly composed of conifers and broad-leaved trees, where pine trees (*Pinus sp.*), Oyameles (*Abies sp.*), Pinabetes (*Picea sp., Pseudotsuga sp.*) and Oak (*Quercus sp.*) are the dominant genres in their composition and structure, predominating

However, Mexico is also the country with the highest rates of deforestation worldwide [8]. Many of its native ecosystems are gradually being reduced to small remnants of the original vegetation [9]. To ensure the sustainability of forests, we need to characterize their structure [10]. This implies the understanding of the different variables associated with the structure of forests [11]. Forest structural variables, such as volume or biomass, average diameter and height, are important data needed to assess forest resources [12,13]. Due to the extensive surface of temperate forests that are distributed in Mexico, together with a very rugged orography, only a portion of the forest can be sampled [14]. Under these conditions, it is important to determine the spatial and temporal distribution of the forest structure. These actions allow classifying the forest coverage according to its natural potential [15].

in the temperate zones of the main mountainous regions of the country [7].

The forest condition is defined in this study as the related set of forest characteristics determined by the number of trees per hectare, their spectral response, their distance from anthropogenic activities, among others. A good condition would present a high number of trees per hectare, a low spectral response and a high distance from anthropogenic activities. At the regional level, a feasible procedure to monitor the forest condition on a regular and continuous basis is based on satellite data, as well as field information [16]. Satellite image data represent a source of key information in the monitoring of the forest condition [17]. The measurement of the forest structure through remote sensing reduces costs by covering large areas of land [18]. In addition, the analyzed data serve to determine the health status of the forests [19]. There is a great variety of studies where satellite images, as well as field data, have been used to characterize the structure of the forest [20–23]. Many of them explain the relationship of the forest structure with the spectral satellite data. Traditionally, the measurement of the forest structure considers the forest's dasometric variables. However, there are environmental variables that influence the productivity of the forest and represent an important element in forest management [24,25]. A large number of biophysical, biological, topographic and anthropogenic variables are associated with forest productivity [26].

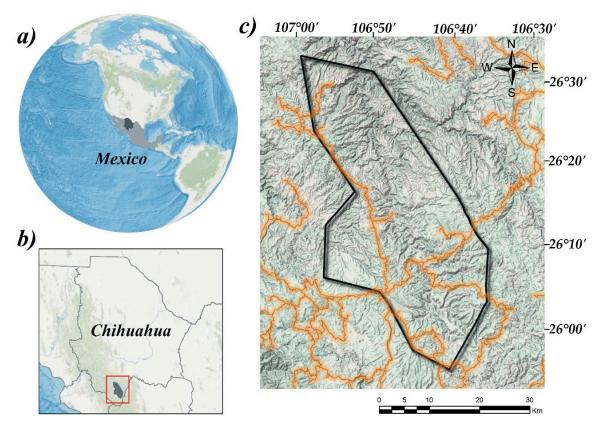
To identify these different forest conditions and variability at the regional level, regionalization or spatial grouping of data is applied. The process of delimitation and classification of areas with homogeneous characteristics according to their environmental properties distributed in the forest landscape allows the definition of environmental units on a spatial scale. In relation to the above, the classification, modeling and interpretation of the monitored data are the most important steps in the evaluation of the forest structure. The spatial identification of regions with homogeneous characteristics often lacks an accepted and clearly articulated theoretical basis. Regions are typically delineated by expert criteria, which is sometimes subjective. The spatial continuity of the resulting regions is rarely managed quantitatively. In this sense, geostatistical multivariate techniques allow us to study the spatial variability of forests, forest structure, topographic and biophysical factors, as well as the relationships that may exist among these variables. Principal Component Analysis (PCA), Cluster Analysis (CA) and Discriminant Analysis (DA) have been used in recent years to analyze environmental variables and extract significant information [27–30]. The use of these techniques allows defining new variables or groups that provide information about the spatial variability of forests [31]. There are few studies that integrate variables of the forest structure by using remote sensing, biophysical, topographic or proximity (roads, communities) data and their interaction [32–34]. Such studies have been carried out mainly in developing countries like Mexico, where there are large gaps of information that can be reinforced with the integration of different sources.

The objective of this study was to analyze the forest condition through environmental, topographic, proximity and spectral variables, employing multivariate techniques with a geospatial approach in a forest area located in the municipality of 'Guadalupe y Calvo', Chihuahua, Mexico.

#### 2. Materials and Methods

# 2.1. Study Area

The study area is located in the municipality of 'Guadalupe y Calvo', in the state of Chihuahua, Mexico, between parallels 107°00' W–26°30' N and 106°30' W–26°00' N (Figure 1). The area comprehends 113.73 ha, of which 95.13 ha are temperate forests (15% of the total forest area of the municipality); 42.60 ha are exploitable pine and oak woods [35]. The region is characterized by its high number of endemic species, estimated at around 4000 plant species [36]. The area is one of the forest regions with the highest biodiversity, which supports various environmental services for the region. It has a unique and wide system of deep canyons, resulting in a mixture of temperate and tropical ecosystems [37]. The main land uses of the study area are coniferous and hardwood forests. Forestry is the main productive activity, representing 42% of the income for the inhabitants. It includes the extraction of wood and the use of dead wood [38]. The second productive activity is livestock for self-consumption, which represents 24% of the economy [39].



**Figure 1.** Geographical location of the study area: (**a**) Mexico, (**b**) "Ejido Chinatu" location in Chihuahua, Mexico, (**c**) "Ejido Chinatu". Boundary of the studied area (—), principal roads (—).

#### 2.2. Forest Data

The data were collected between January and May of 2014 and comprised four thousand 0.1-ha circular ground survey plots (Figure 2). Sample measurements included tree height recorded using digital hypsometer and tree diameter at the breast height (dbh) using a tape. In all plots, the position was determined with a GPS. Canopy closure status was determined by recording the type and proportion of shrub stratum vegetation present in each sample plot using a 1-m square quadrat split

into 4 equal quadrants. Plots with more than 50% dead vegetation on the forest floor were classified as closed canopy. Conversely, those with less than 50% of under growth vegetation were classified as open canopy.

#### 2.3. Geospatial Data

We used data from the Landsat 8 OLI satellite. The original bands were transformed into physical reflectance tools. The spectral value was extracted using a  $3 \times 3$  window to minimize the error [40]. The Normalized Difference Vegetation Index (NDVI) and the Modified Soil Adjusted Vegetation Index 2 (MSAVI2) were generated (Equations (1) and (2)).

$$NDVI = \frac{p_{NIR} - p_{red}}{p_{NIR} + p_{red}} \tag{1}$$

where:  $p_{NIR}$  = near infrared,  $p_{red}$  = red band.

$$MSAVI2 = \frac{2p_{NIR} + 1 - \sqrt{(2p_{NIR})^2 - 8(p_{NIR} - p_{red})}}{2}$$
(2)

where:  $p_{NIR}$  = near infrared,  $p_{red}$  = red band.

The variables considered for the analysis were recorded in the same reference system (WGS84) and were converted to raster format with the same number of rows and columns by using ArcGis 10.3<sup>©</sup> software (Environmental Systems Research Institute, Redlands, CA, USA.; https://www.esri.com/en-us/home). We used 13 variables for the analysis (Table 1). They were selected based on the relationships found in other studies [41–43].

Table 1. Variables used in the data analysis.

Variable	Unit	Source	Acronym
Average total volume per tree	m <sup>3</sup>	Sampling	ATVT
Number of trees per hectare	No	Sampling	NTPH
Basal area per hectare	m <sup>2</sup>	Sampling	BAPH
Total volume of trees per hectare	m <sup>3</sup>	Sampling	TVTPH
Average quadratic diameter	Cm	Sampling	AQD
Spectral band 3	$W/(m^2 \operatorname{sr} \mu m)$	USGS	SB3
Spectral band 7	$W/(m^2 sr \mu m)$	USGS	SB7
Normalized difference vegetation index	Adimensional	Own source	NDVI
Modified soil-adjusted vegetation index 2	Adimensional	Own source	MSAVI2
Distance to roads	m	Own source	DR
Distance to water bodies	m	Own source	DWB
Slope	Degrees	INEGI	Slope
Mean annual temperature	°C	CONAGUA	MÂT

The variables were measured in 345 watershed, delimited within the study area (Figure 2). The size of the watershed varied from 100 to 2100 ha, with an average of 424.02 ha and a standard deviation of 274.07 ha. The geographic information system (GIS) was used to delineate watershed. A digital terrain model [44] was used for the generation of the watersheds [45]. The details of the procedure for the delimitation of watersheds in GIS can be found in the study by Hamdy et al. [46].

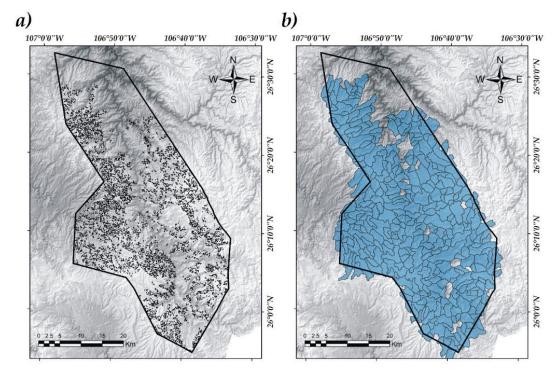


Figure 2. Sampling of structural variables of the forest (a), study unit: watershed (b).

#### 2.4. Statistical Analyses

To identify the relationships among the set of variables (forest and geospatial), a correlation analysis was used in the SAS<sup>©</sup> software, version 9.1.3 [47]. Subsequently, a PCA was performed through the routine PRINCOMP with the SAS software. This analysis helped to reduce the amount of data; it was carried out using Equation (3).

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + \dots + a_{im}x_{mj}$$
(3)

where: *z* represents the component coefficient, *a* represents the weight of the component, *x* represents the measured value of the variable, *i* corresponds to the number of the component, *j* represents the number of the sample and *m* represents the total number of variables.

The watersheds studied were classified by Ascending Group Analysis (AHC) by using the Ward criterion (SAS program, CLUSTER rutine). Ward's criterion was applied to the AHC by using the square Euclidean distance to explore the clustering patterns of the 345 watersheds. The group definition was determined based on the coefficient of determination ( $R^2$ ), the pseudo T<sup>2</sup>, the Cubic Clustering Criterion (CCC) and the pseudo F statistic [48].

Group formation was evaluated based on the original variables by using Discriminant Analysis (DA) (Equation (4)). Data was analyzed to determine whether the variables contribute to discrimination among the groups. Through this analysis, we were able to determine the association of each watershed with a group. The classification error of the generated DA was calculated with the restitution method as well as with the cross-validation method. Finally, the variance was analyzed to verify if there were significant differences between the classification and the independent variables as a whole (MANOVA, [49]).

$$f(Gi) = k_i + \sum_{j=1}^{n} w_{ij} P_{ij}$$
(4)

where: *Gi* represents the number of groups,  $k_i$  = represent an inherent constant for each group, *n* number of parameters, and  $w_{ij}$  weight of the coefficient assigned by the DA for a given parameter  $P_{ij}$ .

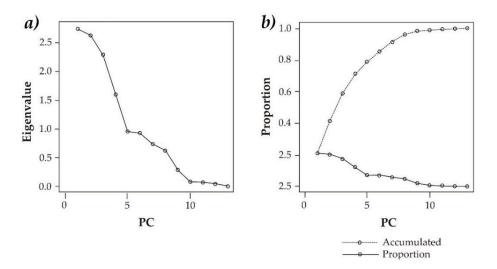
## 3. Results

In the correlation analysis, we found significant positive correlations among the variables (Table 2). The evaluation shows different relationship conditions for the forest in the study area; this condition allowed us to perform a differential characterization as a measure to explain the forest heterogeneity. The PCA showed that the first four components explain 71% of the total variability provided by the 13 original variables. The values of the eigenvalue and the accumulated variability explain the proportion of the original variables (Figure 3).

**Table 2.** Matrix of different variable correlations considered in the analysis.

	ATVT	NTPH	BAPH	TVTPH	DQ	SB3	SB7	NDVI	MSAVI2	DR	DWB	Slope	MAT
ATVT	1.00												
NTPH	$^{-0.41}_{**}$	1.00											
BAPH	0.10	0.72 **	1.00										
TVTPH	0.24 **	0.70 **	0.89 *	1.00									
DQ	0.92 **	-0.47 **	0.11 *	0.18 *	1.00								
SB3	0.05	-0.06	0.29 *	-0.05	0.03	1.00							
SB7	-0.17 *	-0.07	-0.17 *	-0.18 *	-0.15 *	0.35 **	1.00						
NDVI	0.18 *	-0.05	0.11 *	0.12 *	0.23 **	-0.15 *	-0.40 **	1.00					
MSAVI2	0.01	-0.10	-0.05	-0.06	0.07	0.21 *	0.58 **	0.45 **	1.00				
DR	0.03	-0.01	0.03	-0.01	0.08	0.11 *	0.21 **	0.02	0.22 **	1.00			
DWB	0.19 *	-0.16 *	-0.04 *	-0.04	0.24 *	-0.05	-0.09	0.00	-0.13 *	-0.02	1.00		
Slope	0.14 *	-0.07	-0.01	0.01	0.10	-0.06	-0.27 **	-0.13 *	-0.41 **	-0.24 **	0.22 **	1.00	
ANT	0.19 *	-0.11 *	-0.08	-0.04	0.09	0.03	-0.10	-0.31 **	-0.37 **	-0.30 **	0.47 **	0.61 **	1.00

ATVT = Average total volume per tree, NTPH = Number of trees per hectare, BAPH = Basal area per hectare, TVTPH = Total volume of trees per hectare, AQD = Average quadratic diameter, SB3 = Spectral band 3, SB7 = Spectral band 7, NDVI = Normalized difference vegetation index, MSAVI2 = Modified soil-adjusted vegetation index 2, DR = Distance to roads, DWB = Distance to water bodies, MAT = Mean annual temperature. \* = Significant p > 0.05, \*\* = Highly significant p > 0.001.



**Figure 3.** Plot indicating the eigenvalues and their contribution to the total variance (**a**) and explained variance (**b**).

Table 3 shows the characteristics of the first four components and their eigenvectors. The variables that contribute the most to the PC1 (weak coefficients) are Average total volume per tree, Average quadratic diameter, Distance to roads, Slope and Mean annual temperature that represent the forest structural and biophysical variables. The moderate coefficients of PC2 correspond to Number of trees per hectare, Basal area per hectare and Total volume of trees per hectare, which are structural variables of the forest. PC3 has moderate-high coefficients for Average total volume per tree, Average quadratic diameter, NDVI and MSAVI2, which are part of the structural and spectral forest variables. Finally, PC4 has the highest coefficients for the spectral variables Spectral band 3, Spectral band 7 and NDVI. The grouping of the sites based on the variables is presented in Figure 4. The variables can be grouped

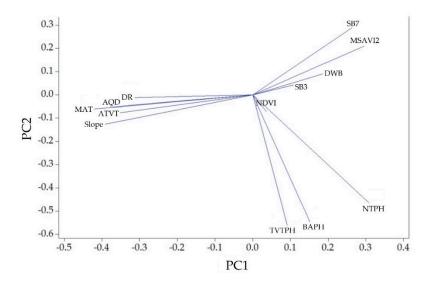
trees per hectare.

into 3: Group 1: Mean annual temperature, Distance to water bodies, Average quadratic diameter, Average total volume per tree, Slope; Group 2: Spectral band 7, MSAVI2, Distance to water bodies, Spectral band 3; and Group 3: Total volume of trees per hectare, Basal area per hectare and Number of

	PC1	PC2	PC3	PC4
ATVT	0.3520	0.0774	0.4682	0.1101
NTPH	-0.3087	0.4643	-0.2010	0.0330
BAPH	-0.1517	0.5470	0.1315	0.2141
TVTPH	-0.0915	0.5614	0.1527	0.0964
AQD	0.3326	0.0453	0.5076	0.0861
SB3	-0.1082	-0.0420	0.1070	0.5630
SB7	-0.2637	-0.2888	0.0276	0.5045
NDVI	-0.0359	0.0715	0.3801	-0.4627
MSAVI2	-0.2956	-0.2097	0.3703	0.0775
DR	-0.1854	-0.0899	0.2187	0.1282
DWB	0.3217	0.0119	0.0093	0.1521
Slope	0.3911	0.1258	-0.2137	0.0936
MÂT	0.4219	0.0606	-0.2338	0.2905

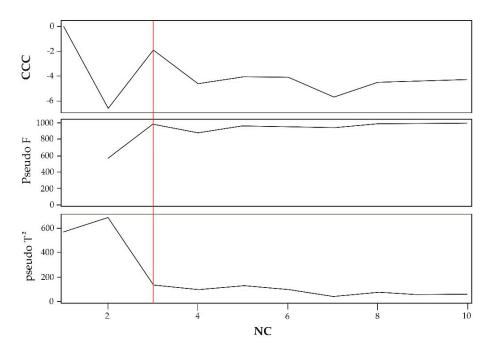
Table 3. Eigenvectors of the correlation matrix.

ATVT = Average total volume per tree, NTPH = Number of trees per hectare, BAPH = Basal area per hectare, TVTPH = Total volume of trees per hectare, AQD = Average quadratic diameter, SB3 = Spectral band 3, SB7 = Spectral band 7, NDVI = Normalized difference vegetation index, MSAVI2 = Modified soil-adjusted vegetation index 2, DR = Distance to roads, DWB = Distance to water bodies, MAT = Mean annual temperature. Bold letters indicate the dominant coefficients.



**Figure 4.** Loading plot of the first two principal components (PC) showing the contribution of each variable.

The PCA results indicate that most of the variations in the forest structure can be explained by Average total volume per tree, Average quadratic diameter, Distance to roads, Slope and Mean annual temperature. The CA determined three groups; 85.2% of the total variability is explained through the three determined groups. Other statistical criteria that reinforced this decision were the CCC, the pseudo  $T^2$  and the pseudo F (Figure 5).



**Figure 5.** Statistical criteria for determining the number of groups to be established. CCC = Cubic Clustering Criterion. NC = Number of clusters.

Figure 6 presents the grouping of the watersheds. A line was drawn around the value of 0.83 to illustrate the cut-off point that delimits the number of groups established according to the Euclidean distance. The line layout can vary according to the statistical and practical criteria defined, increasing or decreasing the number of groups.

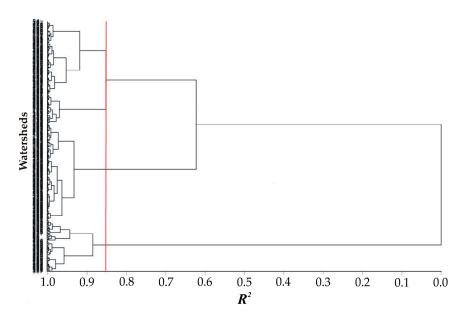
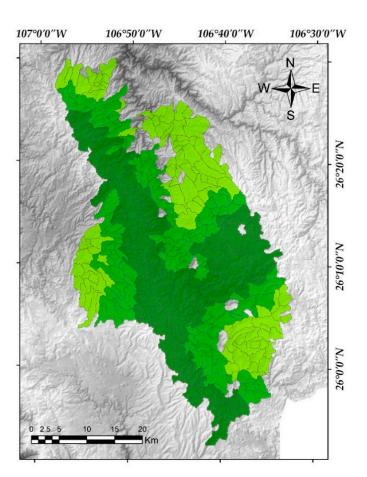


Figure 6. Group tree of watersheds using cluster analysis based on the coefficient of determination  $(R^2)$ .

The groups were evaluated and were statistically different ( $p \le 0.0001$ ). Group 1 comprises 133 watersheds, Group 2 comprises 140 watersheds and Group 3 comprises a total of 72 watersheds (Figure 7).



**Figure 7.** Spatial distribution of the tree clusters defined by group analysis. Group 1(•), Group 2(•), Group 3(•). Name of colors: Bilbao (•), Limeade (•), Inch Worm (•).

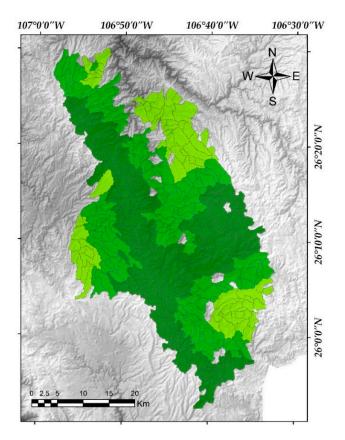
**Group 1** (*Bilbao group*, *n* = 133). Group 1 had the lowest average of the Average total volume per tree, a moderate value for the number of trees per hectare, the smallest value for the basal area per hectare and for the total volume of trees per hectare. Regarding the Average quadratic diameter, this group has an intermediate average value of 21.85 cm. Bands 3 and 7 in this group showed intermediate values between the other groups (0.036 and 0.08). The vegetation indexes showed higher values than Group 1 and 2 for NDVI with 0.55 and 0.18 for MSAVI2, indicating good coverage with respect to the photosynthetic activity of the vegetation. There are great distances to water bodies. The distance to the roads is shorter than distances of Group 2 and 3. Watersheds belonging to group 1 present the lowest slope averages of the three groups. The mean annual temperature is also the lowest of the groups.

**Group 2** (*Limeade group*, *n* = 140). Watersheds of this group are characterized by having the lowest average value of the Average total volume per tree (0.28) and Average quadratic diameter (0.2157). The highest values are for the Number of trees per hectare, the Basal area per hectare and the Total volume of trees per hectare. Bands 3 and 7 had the highest values (0.039 and 0.081). The NDVI and the MSAVI2 had the smallest values. These watersheds have intermediate values towards the Distance to water bodies and the Distance to roads, as well as with the Slope and the Mean annual temperature. Due to the characteristic values of the Average total volume per tree and the Average quadratic diameter, a large number of trees with small trunks are distributed in these watersheds. This group can be characterized as a group with medium forest productive potential, possibly a young forest.

**Group 3** (*Inch Worm group, n* = 72). This group is characterized by having the highest average value of the Average total volume per tree (0.33) and Average quadratic diameter (0.23). Low values were observed for the Number of trees per hectare, the Basal area per hectare and the Total volume of trees per hectare. Band 3 and band 7 showed the lowest reflectance values, indicating that they have

an abundant vegetation cover. In these watersheds, there are trees with large diameters. The NDVI showed the highest value for this group, which can be attributed to the high photosynthetic activity. This group of watersheds has a considerable Distance to water bodies, indicating that they are remote areas without alteration, supported by the Distance to the roads. The slope shows the highest values, which may be an indicator of remote areas with rugged topography. The Mean annual temperature was the highest of the groups.

The classification of the watersheds in the groups was evaluated through the methods of restitution and cross-validation of the DA. A total of 13 watersheds were poorly classified, which resulted in a total error of 0.03. With the method of cross-validation, 18 watersheds were reported as poorly classified with a total error of 0.04. The group watershed map generated in the CA was modified based on the DA (Figure 8). Using the results of the cross-validation, the misclassified observations were changed to those suggested.



**Figure 8.** Spatial distribution of the 3 groups modified by the method of cross-validation of the Discriminant Analysis. Group 1(•), Group 2(•), Group 3(•). Name of colors: Bilbao (•), Limeade (•), Inch Worm (•).

Finally, the multivariate analysis of variance (MANOVA) showed that there are significant differences among watershed groups with respect to the original variables (Wilk's Lambda = 0.1, Pr > F < 0.001), (Table 4).

Contrasts	Value	F-Value	DF	<b>Pr</b> > <i>F</i>
All	0.1026	53.84	26	<0001
1 vs 2 y 3	0.1389	157.31	13	<0001
2 vs 1 y 3	0.8128	5.85	13	<0001
3 vs 1 y 2	0.1579	135.37	13	<0001

Table 4. MANOVA test criteria and approximations of *F*.

DF = degree of freedom, Pr > F = p-value of F statics.

## 4. Discussion

The multivariate techniques and spatial information maps were useful to interpret the relationships among the multiple variables. The PCA has been previously used to examine and interpret the spatial behavior of the forest [50,51]. The relationship between variables from field sampling, thematic maps, and spectral maps provided significant information on the condition of the forest. In this study, four components were needed to explain the original data set. The four components showed the behavior of the variables in the watersheds.

According to Yidina et al. [52], in the analysis of dimensionality reduction of variables, PC1 usually represents the most important mixture of processes. The PCA was useful in the interpretation of qualitative variables. The dominant variables in the first main component were Average total volume per tree, Average quadratic diameter, corresponding to forest variables. Meanwhile, Distance to roads, Slope and Mean annual temperature, presented a strong relationship in the variation of the forest structure in the 'Ejido Chinatu' community. The above agrees with the study by Castillo-Rodríguez et al. [53]. The results of PC1 are consistent with the distribution of vegetation through the altitudinal gradient, where temperature and topography play an important role in the presence of certain species of panaceas [54,55]. This is also consistent with the mentioned by Helman et al. [56], where they found that mean annual temperatures influenced forest productivity in the Mediterranean Mount Carmel forests in Israel. The values of the coefficients within the PC1 present similarities that imply the variables have a similar influence on the variation of the forest structure. PC2 consisted of the variables of Number of trees per hectare, Basal area per hectare and Total volume of trees per hectare, which had weak and moderate coefficients. PC2 can be interpreted as a component that describes the structure of the forest in relation to the Number of trees, the Basal area and the volume measured per hectare. PC3 was represented by the Average total volume per tree, Average quadratic diameter and the spectral indices NDVI and MSAVI2. This is consistent with the mentioned by Kumar et al. [57], who related structural variables of the forest such as diameter and height with the NDVI, obtaining strong relationships of  $R^2 = 0.90$ . This component, similar to PC1, presents information on the forest structure based on the MSAVI2 spectral indices. As shown in the variable displacement plot (Figure 4), as the Average quadratic diameter and the Average total volume per tree decrease, the MSAVI2 increases. The MSAVI2 spectral index helps us to understand the reflectance emitted by bare soil. In this case, lower Average quadratic diameter and higher MSAVI2 indicate that a large number of trees are small.

In agreement with the spatial distribution of the groups shown in Figures 7 and 8, the grouping by CA presented a homogeneous geographical behavior. Group 1 represents the watersheds that correspond to the central corridor, where forest conditions for exploitation could be more accessible. The proximity to the roads is an indicator of anthropic presence [58], and the moderate slopes for this group of watersheds make wood extraction more feasible. Espinosa et al. [59] found that topography is related to species richness in dry forests in southern Ecuador. This is consistent with the distribution of forest variables across the 'Ejido Chinatu', where the increase in slopes ranges from flat to very steep slopes. The moderate spectral values of bands 3 and 7 indicate there is moderate absorption by vegetation and areas that are in a state of recovery. This group shows geographical continuity without being very isolated. Group 2 corresponds to watersheds that have average values with respect to almost all the variables, considering it as a group of watersheds that can be under a moderate

anthropic intervention (transition zone) [60]. It is divided into several homogeneous groups and does not present cases of isolated watersheds. Finally, Group 3 contains watersheds that, due to their spatial distribution, are further away from anthropogenic activities and may not have forest management. The low reflectivity in bands 3 and 7, and the high photosynthetic activity represented by the NDVI values indicate good vegetation conditions [61]. These watersheds are also divided into groups and present an isolated watershed.

The PCA detects the factors that control the spatial variation of the forest structure. Although the PCA is generally considered as a statistical exploratory technique [62], it is capable of being incorporated into the results that explain the distribution of a particular landscape [63]. GIS-based maps have the ability to visualize the spatial relationships between environmental data and other attributes, as reported by Facchinelli et al. [64]. The CA was useful to identify those watersheds that are similar in terms of their multiple environmental characteristics. Riitters and Coulston [65] used spatial statistics to identify the geographical concentrations of forests located near deforested or clear areas. Wang et al. [66] used the multivariate techniques PCA and CA for the delimitation of environmental units. Kupfer et al. [67] regionalized 2100 watersheds based on landscape metrics. Trakhtenbrot and Kadmon [68] used CA for the identification of sites that represent the diversity of species. Ramachandra et al. [69] analyzed information on landscape metrics and social variables using multivariate techniques in forests in Uttara Kannada District, India.

Based on the results of PCA, CA and DA it was possible to understand the multivariate relationship of the set of variables. At the same time, with the combination of geospatial data and multivariate techniques, it was possible to analyze the spatial variability from a point of view of reducing the dimensionality of the information. The combination of both was useful to examine patterns in common group of variables, allowing us to summarize the multiple relationships of variables in geographic regions to use in forest management analysis.

The identification and delimitation of geographic regions in forests is an active area of ongoing research. Technological and methodological advances allow analyzing data from free sources and field information. This study provides a perspective for the analysis of this information, helping us to enhance its interpretation with a more quantitative approach. Traditional grouping and classification methods are widely used, leading to the delimitation of coherent forest region classifications from a geospatial point of view.

From a methodological point of view, the study provides the identification of forest regions that are possibly more vulnerable to the effects of anthropogenic activities, such as changes in land use/ land cover, deforestation and degradation. In addition, the methodology serves to identify the forest geographical areas with potential for conservation. Such areas are of great importance for the development of conservation plans and protection of forest areas, promoted mainly by government agencies.

In particular, the increasingly availability of free and georeferenced data sources, such as those from remote sensing, spectral indicators, digital elevation models and information derived from them, represents a valuable resource for quantitative approaches to zoning. The presentation of geospatial information from forest areas allows users to examine the characteristics of each area, their variability and their level of productivity. However, the detailed presentation of these variations in a broader regional context, where transition zones could be detected, is complex but promising for future studies.

#### 5. Conclusions

The use of GIS associated with geostatistical techniques represents a solid scientific tool for regionalization and grouping of landscape features. The features and regionalization of hydrological watersheds based on environmental attributes and vegetation structures are key to the planning and environmental management of the territories. The application of this methodology allows the rapid integration of several environmental and biological attributes, which can be grouped according to their characteristics. In this way, we can define the productive potentials through the regionalization

analysis. Forest regionalization mapping can benefit a wide variety of management, conservation and protection activities. Thus, if a forest pattern has been identified in one region as favorable or problematic, it can be replicated in another region with similar conditions.

Multivariate statistical methods, forest structure, topographies and satellites data can be useful to group and regionalize forest areas. The multivariate geographic information system approach showed the spatial relationships between the different variables. Although, the results of the present study provided preliminary conclusions about the characteristics of the watershed groups, more studies such as multicriteria evaluation techniques, interpolation methods, among others, are needed to obtain a better understanding of the source of variation in the forest structure.

Author Contributions: Conceptualization, J.A.P.-A.; Formal analysis, E.S.-E. and M.C.V.-A.; Methodology, A.P.-A.; Resources, M.M.-S.; Supervision, F.V.-G. and G.V.-Q.; Writing—original draft, J.A.P.-A., A.P.-A. and C.M.-D.; Writing—review & editing, F.V.-G., M.C.V.-A. and C.M.-D.

**Acknowledgments:** We deeply thank Consejo Nacional de Ciencia y Tecnología (CONACYT-Mexico) for supporting the graduate studies of the first author of this manuscript. We deeply thank Manuel Galvan-Moreno, responsible for the 'Unidad de Manejo Forestal' of 'Guadalupe y Calvo', Chihuahua, for the data provided.

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

# References

- Randolph, J.C.; Green, G.M.; Belmont, J.; Burcsu, T.; Welch, D. Forest Ecosystems and the Human Dimension. In Seeing the Forest and the Trees: Human–Environment Interactions in Forest Ecosystems; Moran, E.F., Ostrom, E., Eds.; MIT Press: Cambridge, MA, USA, 2005; pp. 81–103.
- Bolton, D.K.; Coops, N.C.; Hermosillo, T.; Wulder, M.A.; White, J.C. Assessing variability in postfire forest structure along gradients of productivity in the Canadian boreal using multisource remote sensing. *J. Biogeogr.* 2017, 44, 1294–1305. [CrossRef]
- 3. Hilker, T.; Wulder, M.A.; Coops, N.C.; Linke, J.; McDermid, G.; Masek, J.G.; Gao, F.; White, J.C. A new data fusion model for high spatial-and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sens. Environ.* **2009**, *113*, 1613–1627. [CrossRef]
- Sáenz, R.C.; Rehfeldt, G.E.; Crookston, N.L.; Duval, P.; Amant, R.S.T.; Beaulieu, J.B.; Richardson, A. Spline models of contemporary, 2030, 2060 and 2090 climates for Mexico and their use in understanding climate-change impacts on the vegetation. *Clim. Chang.* 2010, 102, 595–623. [CrossRef]
- 5. Groombridge, B.; Jenkins, M.D. *Global Biodiversity: Earth's Living Resources in the 21st Century;* World Conservation Press: Cambridge, MA, USA, 2000.
- 6. González, E.M.; Meave, J.A.; Marcial, N.R.; Aceves, T.T.; Hernández, F.G.; Manríquez, G.I. Los bosques de niebla de México: conservación y restauración de su componente arbóreo. *Rev. Ecosistemas* **2012**, *21*, 1–2.
- Challenger, A. La zona ecológica templada húmeda. In *Utilización y Conservación de Los Ecosistemas Terrestres de México: Pasado, Presente y Futuro;* CONABIO-Instituto de Biología, UNAM Sierra Madre: Mexico City, México, 1998; pp. 433–518.
- Velázquez, A.; Mas, J.F.; Gallegos, J.R.D.; Saucedo, R.M.; Alcántara, P.C.; Castro, R.; Palacio, J.L. Patrones y tasas de cambio de uso del suelo en México. *Gaceta Ecológica* 2002, 62. Available online: https://www. redalyc.org/articulo.oa?id=53906202 (accessed on 18 January 2019).
- Aguilar, S.V.; Castillo, M.A.; Guerrero, V.F.; Wehenkel, C.; Pinedo, A.C. Modeling the potential distribution of *Picea chihuahuana* Martínez, an endangered species at the Sierra Madre Occidental, Mexico. *Forest* 2015, 6, 692–707. [CrossRef]
- Jupp, D.L.; Walker, J. Detecting structural and growth changes in woodlands and forests: The challenge for remote sensing and the role of geometric-optical modelling. In *The Use of Remote Sensing in the Modeling of Forest Productivity*; Springer: Dordrecht, Netherlands, 1997; pp. 75–108.
- 11. Ludwig, B.; Khanna, P.K.; Raison, R.J.; Jacobsen, K. Modelling changes in cation composition of a soil after clearfelling a eucalypt forest in East Gippsland, Australia. *Geoderma* **1997**, *80*, 95–116. [CrossRef]
- 12. Achard, F.; Mollicone, D.; Stibig, H.J.; Aksenov, D.; Laestadius, L.; Li, Z.; Potapov, P.; Yaroshenko, A. Areas of rapid forest-cover change in boreal Eurasia. *For. Ecol. Manag.* **2006**, *237*, 322–334. [CrossRef]

- 13. Bonan, G.B. Forests and climate change: forcing, feedbacks, and the climate benefits of forests. *Science*. 2008, 320, 1444–1449. [CrossRef]
- Chambers, J.Q.; Asner, G.P.; Morton, D.C.; Anderson, L.O.; Saatchi, S.S.; Espírito-Santo, F.D.; Souza, C. Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends Ecol. Evol.* 2007, 22, 414–423. [CrossRef]
- 15. Zhu, Z.; Woodcock, C.E.; Olofsson, P. Continuous monitoring of forest disturbance using all available Landsat imagery. *Remote Sens. Environ.* **2012**, *122*, 75–91. [CrossRef]
- 16. Franklin, S.E. Remote sensing for sustainable forest management; CRC Press: Boca Raton, FL, USA, 2001.
- 17. Wang, X.; Shao, G.; Chen, H.; Lewis, B.J.; Qi, G.; Yu, D.; Dai, L. An Application of remote sensing data in mapping landscape-level forest biomass for monitoring the effectiveness of forest policies in Northeastern China. *Environ. Manag.* **2013**, *52*, 612–620. [CrossRef]
- 18. Pflugmacher, D.; Cohen, W.B.; Kennedy, R.E. Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. *Remote Sens. Environ.* **2012**, 122, 146–165. [CrossRef]
- 19. Ingram, J.C.; Dawson, T.P.; Whittaker, R.J. Mapping tropical forest structure in southeastern Madagascar using remote sensing and artificial neural networks. *Remote Sens. Environ.* 2005, *94*, 491–507. [CrossRef]
- Powell, S.L.; Cohen, W.B.; Healey, S.P.; Kennedy, R.E.; Moisen, G.G.; Pierce, K.B.; Ohmann, J.L. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sens. Environ.* 2010, 114, 1053–1068. [CrossRef]
- 21. Wolter, P.T.; Townsend, P.A.; Sturtevant, B.R. Estimation of forest structural parameters using 5 and 10 meter SPOT-5 satellite data. *Remote Sens. Environ.* **2009**, *113*, 2019–2036. [CrossRef]
- 22. Dube, T.; Mutanga, O. Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in uMgeni catchment, South Africa. *ISPRS J. Photogramm.* **2015**, *101*, 36–46. [CrossRef]
- Soenen, S.A.; Peddle, D.R.; Hall, R.J.; Coburn, C.A.; Hall, F.G. Estimating aboveground forest biomass from canopy reflectance model inversion in mountainous terrain. *Remote Sens. Environ.* 2010, 114, 1325–1337. [CrossRef]
- 24. Bach, K.; Schawe, M.; Beck, S.; Gerold, G.; Gradstein, S.R.; Moraes, R. Vegetación, suelos y clima en los diferentes pisos altitudinales de un bosque montano de Yungas, Bolivia: Primeros resultados. *Ecología en Bolivia* **2003**, *38*, 3–14.
- Sanderson, M.; Santini, M.; Valentini, R.; Pope, E. *Relationships between forests and weather. EC Directorate General of the Environment*; Met Office: Devon, UK, 2012. Available online: http://ec.europa.eu/environment/forests/pdf/EU\_Forests\_annex1.pdf (accessed on 19 October 2018).
- 26. Vázquez-Quintero, G.; Solís-Moreno, R.; Pompa-García, M.; Villarreal-Guerrero, F.; Pinedo-Alvarez, C.; Pinedo-Alvarez, A. Detection and projection of forest changes by using the Markov Chain Model and Cellular Automata. *Sustainability* **2016**, *8*, 236. [CrossRef]
- 27. Pasher, J.; King, D.J. Multivariate forest structure modelling and mapping using high resolution airborne imagery and topographic information. *Remote Sens. Environ.* **2010**, *114*, 1718–1732. [CrossRef]
- 28. Bhuiyan, M.A.; Rakib, M.A.; Dampare, S.B.; Ganyaglo, S.; Suzuki, S. Surface water quality assessment in the central part of Bangladesh using multivariate analysis. *Ksce J. Civ. Eng.* **2011**, *15*, 995–1003. [CrossRef]
- 29. Batayneh, A.; Zumlot, T. Multivariate statistical approach to geochemical methods in water quality factor identification; application to the shallow aquifer system of the Yarmouk basin of north Jordan. *Res. J. Environ. Earth Sci.* **2012**, *4*, 756–768.
- 30. Oketola, A.A.; Adekolurejo, S.M.; Osibanjo, O. Water quality assessment of River Ogun using multivariate statistical techniques. *J. Environ. Prot. Ecol.* **2013**, *4*, 466. [CrossRef]
- Singh, K.P.; Malik, A.; Mohan, D.; Sinha, S. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—A case study. *Water Res.* 2004, *38*, 3980–3992. [CrossRef]
- 32. Kalabokidis, K.D.; Koutsias, N.; Konstantinidis, P.; Vasilakos, C. Multivariate analysis of landscape wildfire dynamics in a Mediterranean ecosystem of Greece. *Area* 2007, *39*, 392–402. [CrossRef]
- 33. Verdú, F.; Salas, J.; García, C.V. A multivariate analysis of biophysical factors and forest fires in Spain, 1991–2005. *Int J. Wildland Fire* **2012**, *21*, 498–509. [CrossRef]
- 34. Tritsch, I.; Sist, P.; Narvaes, I.D.S.; Mazzei, L.; Blanc, L.; Bourgoin, C.; Gond, V. Multiple patterns of forest disturbance and logging shape forest landscapes in Paragominas, Brazil. *Forests* **2016**, *7*, 315. [CrossRef]

- 35. Torres, A.B.; Acuña, L.; Vergara, J.M. Integrating CBM into land-use based mitigation actions implemented by local communities. *Forests* **2014**, *5*, 3295–3326. [CrossRef]
- 36. Felger, R.S.; Johnson, M.B. Trees of the northern Sierra Madre Occidental and sky islands of southwestern North America. In *Biodiversity and Management of the Madrean archipelago: The Sky Islands of Southwestern* United States and Northwestern Mexico; US Forest Service General Technical Report: RM-GTR-264: Tucson, AZ, USA, 1995; pp. 71–83.
- 37. CONABIO. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. In *La biodiversidad en Chihuahua: Estudio de Estado*, 1st ed.; Comisión Nacional para el Conocimiento y Uso de la Biodiversidad: México D.F., Méx, 2014; p. 22.
- 38. Torres, R. 2013. Cuenca río Turuachi con atención a áreas críticas de deforestación y degradación forestal. Propuesta de Proyecto enviada por Ejido Chinatú como parte de la Convocatoria a proyectos de campo a través de Alianzas para la preparación a REDD+ del proyecto México de Reducción de Emisiones por Deforestación y Degradación Forestal (MREDD+); Alianza México REDD+: Mexico City, Mexico, 2012.
- 39. INEGI. *Síntesis de información geográfica del Estado de Chihuahua*, 1st ed.; Instituto Nacional de Estadística, Geografía e Informática: Aguascalientes, Méx, 2003; p. 113.
- Hall, R.J.; Skakun, R.S.; Arsenault, E.J.; Case, B.S. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. *Forest. Ecol. Manag.* 2006, 225, 378–390. [CrossRef]
- 41. Miranda, A.L.; Garza, E.J.; Pérez, J.J.; Calderón, O.A.; Tagle, M.A.; García, P.M.; Salado, A.C. Modeling susceptibility to deforestation of remaining ecosystems in North Central Mexico with logistic regression. *J. For. Res.* **2012**, *23*, 345–354. [CrossRef]
- Pérez, V.A.; Mas, J.F.; Zielinska, A.L. Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. *Environ. Modell. Softw.* 2012, 29, 11–23. [CrossRef]
- 43. Bax, V.; Francesconi, W.; Quintero, M. Spatial modeling of deforestation processes in the Central Peruvian Amazon. *J. Nat. Conserv.* **2016**, *29*, 79–88. [CrossRef]
- 44. INEGI, Instituto Nacional de Estadística, Geografía e Informática. Sistema de Descarga del Continuo de Elevaciones Mexicano. 2005. Available online: http://mapserver.inegi.gob.mx (accessed on 19 July 2018).
- 45. Maidment, D.R.; Djokic, D. *Hydrologic and Hydraulic Modelling Support: with Geographic Information Systems*; ESRI Inc.: Redlands, CA, USA, 2000; pp. 65–84.
- Hamdy, O.; Zhao, S.; Salheen, M.A.; Eid, Y.Y. Identifying the risk areas and urban growth by ArcGIS tools. *Geosciences* 2016, 6, 47. [CrossRef]
- 47. SAS software, Version 9.1.3; SAS Institute Inc.: Cary, NC, USA, 2006.
- 48. Eder, B.K.; Davis, J.M.; Bloomfield, P. An automated classification scheme designed to better elucidate the dependence of ozone on meteorology. *J. Appl. Meteorol.* **1994**, *33*, 1182–1199. [CrossRef]
- Johnson, R.A.; Wichern, D.W. Applied Multivariate Statistical Analysis; Prentice-Hall Inc.: Upper Saddle River, NJ, USA, 1988.
- Li, A.; Wang, A.; Liang, S.; Zhou, W. Eco-environmental vulnerability evaluation in mountainous region using remote sensing and GIS—A case study in the upper reaches of Minjiang River, China. *Ecol. Model.* 2006, 192, 175–187. [CrossRef]
- 51. Zhang, Q.F.; Wu, F.Q.; Wang, L.; Yuan, L.; Zhao, L.S. Application of PCA integrated with CA and GIS in eco-economic regionalization of Chinese Loess Plateau. *Ecol. Econ.* **2011**, *70*, 1051–1056. [CrossRef]
- 52. Yidana, S.M.; Banoeng-Yakubo, B.; Akabzaa, T.M. Analysis of groundwater quality using multivariate an spatial analyses in the Keta basin, Ghana. *J. Afr. Earth Sci.* **2010**, *58*, 220–234. [CrossRef]
- Castillo, R.; Blanco, J.L.; Salinas, E.M. A geomorphologic GIS-multivariate analysis approach to delineate environmental units, a case study of La Malinche volcano (central México). *Appl. Geogr.* 2010, *30*, 629–638. [CrossRef]
- Rzedowski, J. Diversidad y orígenes de la flora fanerogámica de México. In *Diversidad biológica de México:* orígenes y distribución; Ramamoorthy, T.P., Bye, R., Lot, A., Fa, J., Eds.; UNAM: Instituto de Biología, México, 1998; pp. 129–145.
- 55. Hlásny, T.; Trombik, J.; Bošeľa, M.; Merganič, J.; Marušák, R.; Šebeň, V.; Štěpánek, P.; Kubišta, J.; Trnka, M. Climatic drivers of forest productivity in Central Europe. *Agr. For. Meteorol.* **2017**, *234*, 258–273. [CrossRef]

- Helman, D.; Osem, Y.; Yakir, D.; Lensky, I.M. Relationships between climate, topography, water use and productivity in two key Mediterranean forest types with different water-use strategies. *Agr. Forest Meteorol.* 2017, 232, 319–330. [CrossRef]
- 57. Kumar, P.; Sajjad, H.; Mahanta, K.K.; Ahmed, R.; Mandal, V.P. Assessing suitability of allometric models for predicting stem volume of Anogeissus pendula Edgew in sariska Tiger Reserve, India. *Remote Sens. Appl. Soc. Environ.* **2018**, *10*, 47–55. [CrossRef]
- 58. Kirby, K.R.; Laurance, W.F.; Albernaz, A.K.; Schroth, G.; Fearnside, P.M.; Bergen, S.; Venticinque, E.M.; Costa, C.D. The future of deforestation in the Brazilian Amazon. *Futures*. **2006**, *38*, 432–453. [CrossRef]
- Espinosa, C.I.; De la Cruz, M.; Jara-Guerrero, A.; Gusmán, E.; Escudero, A. The effects of individual tree species on species diversity in a tropical dry forest change throughout ontogeny. *Ecography* 2016, *39*, 329–337. [CrossRef]
- 60. Barber, C.P.; Cochrane, M.A.; Souza, C.M., Jr.; Laurance, W.F. Roads, deforestation, and the mitigating effect of protected areas in the Amazon. *Biol. Conserv.* **2014**, 177, 203–209. [CrossRef]
- 61. Mancino, G.; Nolè, A.; Ripullone, F.; Ferrara, A. Landsat TM imagery and NDVI differencing to detect vegetation change: assessing natural forest expansion in Basilicata, southern Italy. *IFOREST* 2014, 7, 75. [CrossRef]
- 62. Johnson, D. *Métodos multivariados aplicados al análisis de datos*, 1st ed.; International Thompson Editors: México D.F., México, 1998; p. 566.
- 63. De La Cueva, A.V. Structural attributes of three forest types in central Spain and Landsat ETM+ information evaluated with redundancy analysis. *Int. J. Remote Sens.* **2008**, *29*, 5657–5676. [CrossRef]
- 64. Facchinelli, A.; Sacchi, E.; Mallen, L. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. *Environ. Pollut.* **2001**, *114*, 313–324. [CrossRef]
- 65. Riitters, K.H.; Coulston, J.W. Hotspots of perforated forest in the Eastern United States. *Enviro. Manag.* 2005, 35, 483–492. [CrossRef]
- 66. Wang, H.F.; Qureshi, S.; Qureshi, B.A.; Qiu, J.X.; Friedman, C.R.; Breuste, J.; Wang, X.K. A multivariate analysis integrating ecological, socioeconomic and physical characteristics to investigate urban forest cover and plant diversity in Beijing, China. *Ecol. Indic.* **2016**, *60*, 921–929. [CrossRef]
- 67. Kupfer, J.A.; Gao, P.; Guo, D. Regionalization of forest pattern metrics for the continental United States using contiguity constrained clustering and partitioning. *Ecol. Inform.* **2012**, *9*, 11–18. [CrossRef]
- 68. Trakhtenbrot, A.; Kadmon, R. Environmental cluster analysis as a tool for selecting complementary networks of conservation sites. *Ecol. Appl.* **2005**, *15*, 335–345. [CrossRef]
- 69. Ramachandra, T.V.; Setturu, B.; Chandran, S. Geospatial analysis of forest fragmentation in Uttara Kannada District, India. *For. Ecosyst.* **2016**, *3*, 10.



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